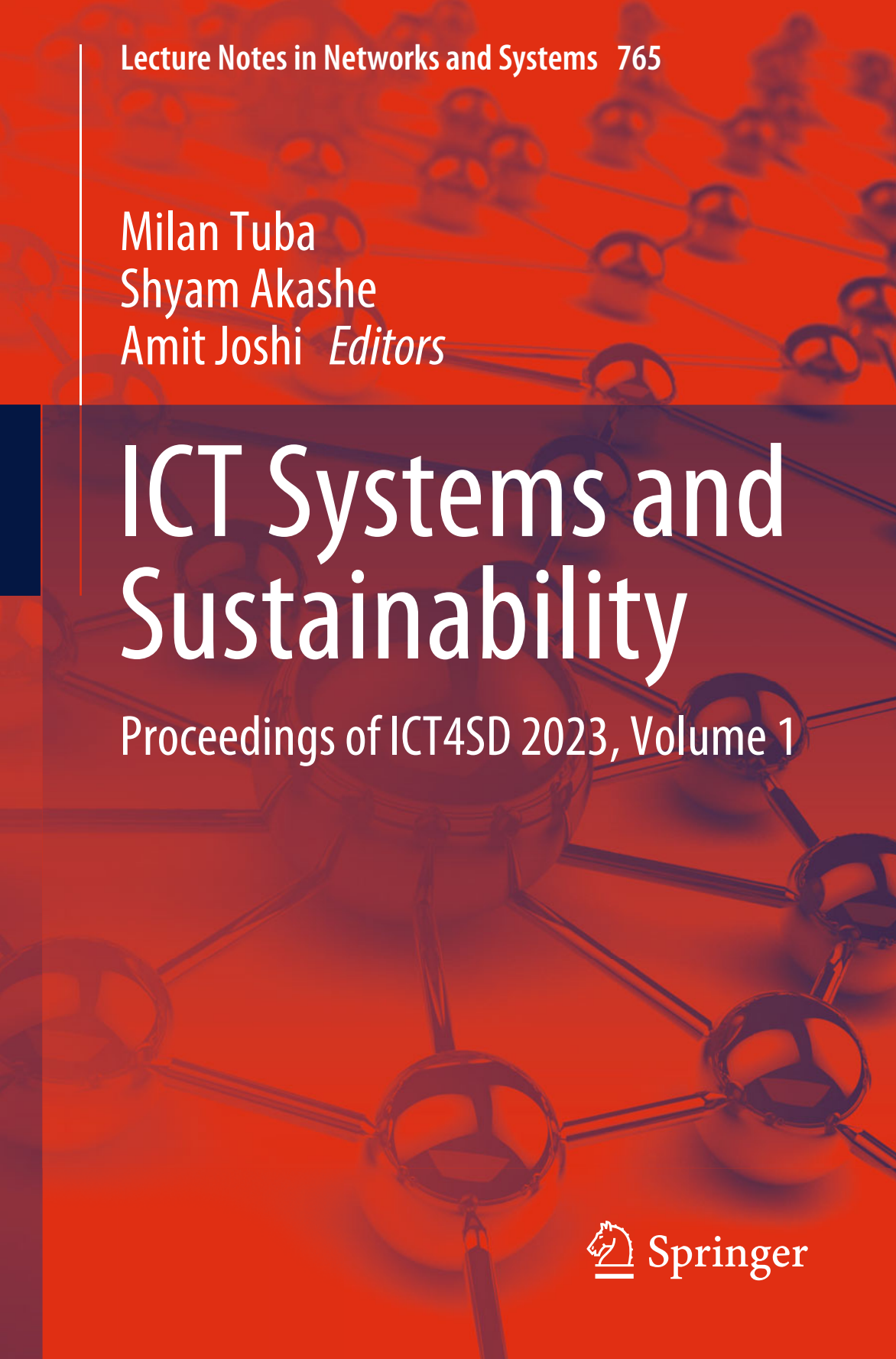
Lecture Notes in Networks and Systems 765

Milan Tuba
Shyam Akashe
Amit Joshi *Editors*

ICT Systems and Sustainability

Proceedings of ICT4SD 2023, Volume 1



Springer

Lecture Notes in Networks and Systems

Volume 765

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Milan Tuba · Shyam Akashe · Amit Joshi
Editors

ICT Systems and Sustainability

Proceedings of ICT4SD 2023, Volume 1

Springer

Editors
Milan Tuba
Singidunum University
Belgrade, Serbia

Shyam Akashe
ITM University
Gwalior, India

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-99-5651-7 ISBN 978-981-99-5652-4 (eBook)
https://doi.org/10.1007/978-981-99-5652-4

Preface

The eighth International Conference on ICT for Sustainable Development (ICT4SD 2023) targets theory, development, applications, experiences, and evaluation of interaction sciences with fellow students, researchers, and practitioners.

The conference is devoted to increasing the understanding role of technology issues and how engineering has day by day evolved to prepare human friendly technology. The conference will provide a platform for bringing forth significant research and literature across the field of ICT for Sustainable Development and provide an overview of the technologies awaiting unveiling. This interaction will be the focal point for leading experts to share their insights, provide guidance, and address participant's questions and concerns.

The conference will be held on August 3–4, 2023, at Hotel Vivanta by TAJ, Panaji, Goa. The conference is organized by Global Knowledge Research Foundation and Managed By: G R Scholastic LLP, State Chamber Partner Goa Chamber of Commerce and Industry and National Chamber Partner Knowledge Chamber of Commerce and Industry.

Research submissions in various advanced technology areas were received and after a rigorous peer-review process with the help of program committee members and 185 external reviewers for 1100+ papers from 16 different countries including Saudi Arabia, USA, Singapore, Denmark, Norway, Denmark, Ghana, United Arab Emirates, The Netherlands, Iraq, Bangladesh, Japan, Malaysia, and Finland, out of which 165 were accepted with an acceptance ratio of 0.16. These will be presented in 18 parallel sessions in two days and organized physical at Goa and virtual on Zoom including one inaugural and one keynote session.

Technology is the driving force of progress in this era of globalization. Information and Communication Technology (ICT) has become a functional requirement for the socioeconomic growth and sustained development of any country. The influence of information communications technology (ICT) in shaping the process of globalization, particularly in productivity, commercial, and financial spheres, is widely recognized. The ICT sector is undergoing a revolution that has momentous implications for the current and future social and economic situation of all the countries in the world. ICT plays a pivotal role in empowering people for self-efficacy and how

it can facilitate this mission to reach out to grassroots level. Finally, it is concluded that ICT is a significant contributor to the success of the ongoing initiative of start-up India.

In order to recognize and reward the extraordinary performance and achievements by ICT and allied sectors and promote universities, researchers and students through their research work adapting new scientific technologies and innovations, the two days conference had presentations from the researchers, scientists, academia, and students on the research work carried out by them in different sectors.

Belgrade, Serbia — Milan Tuba
Gwalior, India — Shyam Akashe
Ahmedabad, India — Amit Joshi

Contents

Editors and Contributors

About the Editors

Milan Tuba is Professor of Computer Science and Mathematics, Head of the Artificial Intelligence Project at Singidunum University and Vice-Rector of Research at Sinergia University, is included in both versions of the Stanford University list of 2% of the most influential scientists in the world in all disciplines, one for contribution during the entire career and other for contribution in the previous year (for years 2020–2022). He was Vice Rector for International Relations, Singidunum University, Belgrade, Head of the Department for Mathematical Sciences at State University of Novi Pazar and Dean of the Graduate School of Computer Science at John Naisbitt University. Professor Tuba is the author or coauthor of around 300 scientific papers (cited more than six, 600 times, h-index 49) and editor, coeditor or member of the editorial board or scientific committee of number of scientific journals, Springer books, congresses and international conferences. He was invited and delivered more than 90 keynote and inaugural lectures at international conferences.

He received B.S. in Mathematics, M.S. in Mathematics, M.S. in Computer Science, M.Ph. in Computer Science, Ph.D. in Computer Science from University of Belgrade and New York University. From 1983 to 1994 he was in the USA at Vanderbilt University in Nashville and Courant Institute of Mathematical Sciences, New York University and later as Assistant Professor of Electrical Engineering at Cooper Union School of Engineering, New York. During that time, he was the founder and director of Microprocessor Lab and VLSI Lab, leader of the NSF scientific projects and theses supervisor. He was the mentor of dozens of doctoral and master's dissertations at the Faculty of Mathematics University of Belgrade, Singidunum University, University of Sarajevo, State University of Novi Pazar, John Nesbitt University and University of East Sarajevo. He was teaching more than 20 graduate and undergraduate courses, from VLSI Design and Computer Architecture to Computer Networks, Operating Systems, Artificial Intelligence, Image Processing, Calculus and Queuing

Theory at numerous universities in Europe and the USA. Professor Tuba is a member of the National Agency for Accreditation of Universities of the Republic of Serbia.

His research interest includes Artificial Intelligence, Deep Learning, Neural Networks, Nature-inspired Optimization Algorithms, Image Processing, Computer Networks. Senior Member IEEE, ACM, AMS, SIAM, IFNA, Executive Board of IASEI.

Shyam Akashe is Professor at ITM University, Gwalior, Madhya Pradesh, India. He completed his Ph.D. at Thapar University, Punjab, and his M.Tech. in Electronics and Communication Engineering at the Institute of Technology and Management, Gwalior. He has authored 190 publications, including more than 50 papers in SCI-indexed journals. His main research focus is low-power system on chip (SoC) applications in which static random access memories (SRAMs) are omnipresent. He has authored two books entitled Moore's Law Alive: Gate-All-Around (GAA) Next Generation Transistor published by LAMBERT Academic Publishing, Germany and Low Power High Speed CMOS Multiplexer Design published by Nova Science Publishers, Inc., New York, USA. He has also published over 120 papers in leading national and international refereed journals of repute. Dr. Akashe has participated in numerous national and international conferences and presented over 100 papers.

Dr. Amit Joshi is currently the Director of Global Knowledge Research Foundation, also an entrepreneur and researcher who has completed his graduation (B.Tech.) in Information Technology and M.Tech. in Computer Science and Engineering and completed his research in the areas of Cloud Computing and Cryptography in Medical imaging with a focus on analysis of the current Government Strategies and World forums needs in different sectors on security purposes. He has an experience of around 12 years in academic and industry in prestigious organizations. He is an active member of ACM, IEEE, CSI, AMIE, IACSIT-Singapore, IDES, ACEEE, NPA and many other professional societies. Currently he is a Chairman of Computer Society of India (CSI) Udaipur Chapter and Secretary for Association of Computing Machinery (ACM) Udaipur Professional Chapter. Further currently he is also the International Chair of InterYIT at International Federation of Information Processing (IFIP, Austria), He has presented and published more than 50 papers in National and International Journals/Conferences of IEEE and ACM. He has also edited more than 20 books which are published by Springer, ACM and other reputed publishers. He has also organized more than 40 National and International Conferences and Workshops through ACM, Springer, IEEE across five countries including India, UK, Thailand, Europe. Apart from this academic involvement he is also associated with the Business and Industry Community across the globe. He is currently the Director of Emanant TechMedia (P) Limited which focuses on ICT and web based services. With Global Knowledge Research Foundation his major work area is focussed on building and making effective linkages among bureaucrats, industry associations, Academic Leaders and regulatory authorities for effective connectivity for taking up common issues related to research in different sectors and has organized more than 20 Industry forum Events and recognised and felicitated Awards Ceremonies

through his organizations in last 5 years. Liaisoning/communications with State as well as Central establishments along with corporate and Academics to take Industry and Government Sector together. He is majorly focusing on Academic Tie-ups for Medical, Engineering and management sectors, for the same he is representing the organization in many delegations to countries including USA, UK, Canada, Israel and visited more than 15 countries. He always looks forward and strives for Networking in various sectors with a focus on Education and Industry Sector.

Contributors

Sanjay Agal Dr. V R Godhania College of Engineering and Technology, Porbandar, Gujarat, India

Khushbu Agarwal PAHER University, Udaipur, Rajasthan, India

Hemel Sharker Akash Department of Computer Science and Engineering, Pabna University of Science and Technology, Pabna, Bangladesh

Nisha S. Amin Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India

Sandhya Arora Cummins College of Engineering for Women, Pune, India

M. Asha Rani Department of Electronics Communication and Engineering, JNTUH, Hyderabad, India

Geeta Babusingh Department of Applied Electronics, Gulbarga University, Kalaburgi, Karnataka, India

Yash Bagade Walchand Institute of Technology, Solapur, Maharashtra, India

Vishwanath P. Baligar K.L.E. Technological University, Hubballi, India

Chitrakant B. Banchhor Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

Janki Barot Silver Oak University, Ahmedabad, India

Santi Kumari Behera Department of CSE, VSSUT, Burla, India

Nisha Bharti Symbiosis Institute of International Business, Symbiosis International (Deemed University), Pune, India

Stuti Bhojak Nirma University, Ahmedabad, India

Pravin Kumar Bhoyar Symbiosis Institute of Management Studies (SIMS), A Constituent of Symbiosis International Deemed University, Pune, Maharashtra, India

Suneeta V. Budihal School of ECE, KLE Technological University, Hubballi, India

Saswati Chakraborty Symbiosis International (Deemed University), Lavale, Pune, Maharashtra, India

Sandeep Chaurasia Manipal University Jaipur, Jaipur, Rajasthan, India

Puja Chavan Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Vaishnavi Chillal School of ECE, KLE Technological University, Hubballi, India

Pradnya Vishwas Chitrao Symbiosis Institute of Management Studies (SIMS), A Constituent of Symbiosis International Deemed University, Pune, Maharashtra, India

Gaurav Choudhary DTU Compute, Technical University of Denmark, Kgs. Lyngby, Denmark;
DTU Compute, Technical University of Denmark, Copenhagen, Denmark

Vipul Chudasama Nirma University, Ahmedabad, India

Saroj S. Date Department of CS and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India

Mahipal Singh Deora Department of Computer Science, B. N. University, Udaipur, India

Darp Desai Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Gopal Deshmukh Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

Parth S. Deshmukh Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India

Riddhi Dhage Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India

Rajiv Divekar Symbiosis Institute of Management Studies (SIMS), A Constituent of Symbiosis International Deemed University, Pune, Maharashtra, India

Amit Dolas Vishwakarma Institute of Technology, Pune, India

Shital Dongare Vishwakarma Institute of Technology, Pune, India

Nishant Doshi Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Nicola Dragoni DTU Compute, Technical University of Denmark, Copenhagen, Denmark

Jayesh Gangrade Manipal University Jaipur, Jaipur, Rajasthan, India

Kumar Gaurav Department of Earth and Environmental Science, Indian Institute of Science Education and Research Bhopal, Bhopal, India

Piyush P. Gawali Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

Swapnil Gawali Vishwakarma Institute of Technology, Pune, India

Heena Kouser Gogi PDACE, Kalaburgi, Karnataka, India

Sanjay Gour Department of Computer Science and Engineering, Jaipur Engineering College and Research Centre, Jaipur, India

Divya Hiran Department of Home Science, Govt. Meera Girl's College, Udaipur, India

Sarvesh Hon Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India

Abhilash Preetham Jayanthi R.V. College of Engineering, Bangalore, India

P. Jayanthi Department of Information Technology, Anna University, Chennai, India

Varad Joshi LJ University, Ahmedabad, Gujarat, India

Vishab Julka Shah and Anchor Kutchhi Engineering College, Mumbai, India

Laxmi Kale Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India

Preeti Kale School of Computer Science Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

K. Karibasappa Department of Computer Science and Engineering, Graphic Era (Deemed to be) University, Dehradun, India

K. G. Karibasappa School of Computer Science and Engineering, KLE Technological University, Hubballi, Karnataka, India

Sai Kaustubh Jain (Deemed-to-Be University), Bengaluru, India

Shivprasad Kavathe Vishwakarma Institute of Technology, Pune, India

C. Kavipriya Department of Information Technology, Anna University, Chennai, India

Kumar Keshamoni Assistant Professor, Department of ECE, Vaagdevi Engineering College, Warangal, India

Vajid Khan Department of Computer Engineering, KJ College of Engineering and Management Research, SPPU Pune, Pune, India

Akshata S. Kori M. S. Ramaiah Institute of Technology, Bengaluru, Karnataka, India

Mahesh A. Kori KLS Gogte Institute of Technology, Belagavi, Karnataka, India

Sharada M. Kori KLS Gogte Institute of Technology, Belagavi, Karnataka, India

Hemant Kothari Pacific Academy of Higher Education and Research University, Udaipur, India

Y. Krishna Suhruth Department of Computer Science Engineering, CVR College of Engineering, Hyderabad, India

Prutha Kudtarkar Shah and Anchor Kutchhi Engineering College, Mumbai, India

P. Kumar Rajalakshmi Engineering College, Chennai, India

S. Vinodh Kumar Rajalakshmi Engineering College, Chennai, India

Sahil Kumar Indian Institute of Information Technology, Ranchi, Ranchi, India

Vaibhav Kumar Department of Data Science and Engineering, Indian Institute of Science Education and Research Bhopal, Bhopal, India

Sneha Kumari Symbiosis School of Economics, Symbiosis International (Deemed University), Pune, India

Vidya Kumbhar Symbiosis Institute of Geoinformatics (SIG), Symbiosis International (Deemed University), Pune, India

Hemlata Kunwar PAHER University, Udaipur, Rajasthan, India

Nidhi Kushwaha Indian Institute of Information Technology, Ranchi, Ranchi, India

R. Punitha Lakshmi Department of ECE, Puducherry Technological University, Puducherry, India

Parikshit N. Mahalle Professor, Vishwakarma Institute of Information Technology, Pune, India

Parikshit N. Mahalle Department of AI & DS, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

Vikas B. Maral Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

Rohini R. Mergu Walchand Institute of Technology, Solapur, Maharashtra, India

Abu Saleh Musa Miah School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Fukushima, Japan

Sumita Mishra KIIT School of Management, KIIT Deemed to Be University, Bhubaneswar, Odisha, India

Harsh Mody Thadomal Shahani Engineering College, Mumbai, India

Radhika Moondra Department of Science, Technology and Forensics, Sardar Patel University of Police, Security and Criminal Justice, Jodhpur, India

Zeeshan Mujawar School of Computer Science Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Rohini Mutalikdesai School of Computer Science and Engineering, KLE Technological University, Hubballi, Karnataka, India

P. Nagasundar Jain (Deemed-to-Be University), Bengaluru, India

Nabin Kumar Naik Department of Electronics, Sambalpur University, Burla, India

Savita Nandan Indian Institute of Information Technology, Ranchi, Ranchi, India

Suvarna Nandyal Department of CSE, PDACE, Kalaburgi, Karnataka, India

Apoorva Narode School of ECE, KLE Technological University, Hubballi, India

Anushk Naval Department of Earth and Environmental Science, Indian Institute of Science Education and Research Bhopal, Bhopal, India

Suyash Nehete Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India

K. A. Nethravathi R.V. College of Engineering, Bangalore, India

Saloni Nimgaonkar Vishwakarma Institute of Technology, Pune, Maharashtra, India

Chaitanya Nirfarake School of Computer Science Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Yuichi Okuyama School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Fukushima, Japan

Sanjay Pandey Career Launcher, Pune, India

Suruchi Pandey Symbiosis Institute of Management Studies, Symbiosis International (Deemed University), Pune, Maharashtra, India

Millie Panigrahi Department of ECE, Trident Academy of Technology, Bhubaneswar, India

Kanika Panwar Sardar Patel University of Police, Security and Criminal Justice, Jodhpur, India

Punita Panwar Manipal University Jaipur, Jaipur, Rajasthan, India

Harsh Parikh Thadomal Shahani Engineering College, Mumbai, India

Martin Parmar Chandubhai S. Patel Institute of Technology (CSPIT), Charotar University of Science and Technology (CHARUSAT), Anand, Gujarat, India

Tanuja Patankar Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India

Arpita Patel Charotar University of Science and Technology, Changa, Gujarat, India

Ishita Patel Shah and Anchor Kutchhi Engineering College, Mumbai, India

Premal Patel Department of Computer Engineering, College of Technology, Silver Oak University, Ahmedabad, Gujarat, India

Soham Patel Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Ashwini Patil Symbiosis Institute of Geoinformatics (SIG), Symbiosis International (Deemed University), Pune, India

Mayur Patil Vishwakarma Institute of Technology, Pune, Maharashtra, India

Neeraj Patil Thadomal Shahani Engineering College, Mumbai, India

Rushal Patil Vishwakarma Institute of Technology, Pune, Maharashtra, India

Suyog Patil Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Tanaya Patil Cummins College of Engineering for Women, Pune, India

Tanuja R. Patil K.L.E. Technological University, Hubballi, India

Priteshkumar Prajapati Chandubhai S. Patel Institute of Technology (CSPIT), Charotar University of Science and Technology (CHARUSAT), Changa, India

Sidhant Priyadarshi School of ECE, KLE Technological University, Hubballi, India

T. V. Raghavasimhan Department of Information Technology, Anna University, Chennai, India

Prasad Raghunath Mutkule Research Scholar, Vishwakarma Institute of Information Technology, Pune, India

Md. Abdur Rahim Department of Computer Science and Engineering, Pabna University of Science and Technology, Pabna, Bangladesh

R. L. Raibagkar Department of Applied Electronics, Gulbarga University, Kalaburgi, Karnataka, India

G. G. Rajput Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India

Jeet Ramoliya Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Prashil Ramteke Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Priyanka Ramteke Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Mahi Rastogi Jain (Deemed-to-Be University), Bengaluru, India

Prasad Rathod Vishwakarma Institute of Technology, Pune, India

Prashant Raut Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Dhruva R. Rinku Department of Electronics Communication and Engineering, CVR College of Engineering, Hyderabad, India

Nilesh P. Sable Associate Professor, Vishwakarma Institute of Information Technology, Pune, India

Malabika Sahoo KIIT School of Management, KIIT Deemed to Be University, Bhubaneswar, Odisha, India

Hemant Sahu JRN Rajasthan Vidyapeeth, Udaipur, India

Sachin R. Sakhare Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

N. Satyanarayana Department of Information Technology, TKR College of Engineering and Technology, Hyderabad, Telangana, India

Radha Senthilkumar Department of Information Technology, Anna University, Chennai, India

Prabira Kumar Sethy Department of Electronics, Sambalpur University, Burla, India

Chintan Shah Department of Computer Engineering, College of Technology, Silver Oak University, Ahmedabad, Gujarat, India

Parth Shah Chandubhai S. Patel Institute of Technology (CSPIT), Charotar University of Science and Technology (CHARUSAT), Anand, Gujarat, India

Pushpkant Shakdwipee PAHER University, Udaipur, Rajasthan, India

Ishita Sharma Shah and Anchor Kutchhi Engineering College, Mumbai, India

Neetu Sharma Department of Computer Science and Engineering, Engineering College Ajmer, Ajmer, India

Sonal Sharma Jain (Deemed-to-Be University), Bengaluru, India

Rupali J. Shelke Walchand Institute of Technology, Solapur, Maharashtra, India

Parth Shethji Vishwakarma Institute of Technology, Pune, Maharashtra, India

Pranav Shetkar Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Jungpil Shin School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Fukushima, Japan

Gitanjali R. Shinde Associate Professor, Vishwakarma Institute of Information Technology, Pune, India

Saroja V. Siddamal School of ECE, KLE Technological University, Hubballi, India

Vikas Sihag Department of Science, Technology and Forensics, Sardar Patel University of Police, Security and Criminal Justice, Jodhpur, India

Bharat Singh Indian Institute of Information Technology, Ranchi, Ranchi, India

Manish Singh Department of Computer Engineering, College of Technology, Silver Oak University, Ahmedabad, Gujarat, India

Shivoham Singh Pacific Academy of Higher Education and Research University, Udaipur, India;
Department of Management & Commerce, MRU, Faridabad, India;
PAHER University, Udaipur, Rajasthan, India;
Symbiosis Institute of Business Management (SIBM-H), Hyderabad, India

Deep Solanki Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Vikash Sonar School of ECE, KLE Technological University, Hubballi, India

Bhakti Sonawane Shah and Anchor Kutchhi Engineering College, Mumbai, India

Spurthi Srinivas Jain (Deemed-to-Be University), Bengaluru, India

V. Sudarsan Rao Department of Computer Science and Engineering, Sri Chaitanya Institute of Technology and Research, Khammam, Telangana, India

Anuroop Sundd Department of Computer Science, B. N. University, Udaipur, India

Dattatray G. Takale Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

Ajay Talele Vishwakarma Institute of Technology, Pune, Maharashtra, India

R. S. Tejeshwini Department of Information Technology, Anna University, Chennai, India

Priyanshi Thakkar Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Poonam Thanki Charotar University of Science and Technology, Changa, Gujarat, India

Shanthi Therese Thadomal Shahani Engineering College, Mumbai, India

Yoichi Tomioka School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Fukushima, Japan

Divam Trivedi R.V. College of Engineering, Bangalore, India

Srikanth Vadada LRDE, DRDO, Bangalore, India

Devendra Vashi Nirma University, Ahmedabad, India

Hrishikesh Vaze School of Computer Science Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Shanti Verma LJ University, Ahmedabad, Gujarat, India

V. Vijayalakshmi Department of ECE, Puducherry Technological University, Puducherry, India

Daiwat Vyas Nirma University, Ahmedabad, India

Atharva Wagh School of Computer Science Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Prajjwal Walchale Walchand Institute of Technology, Solapur, Maharashtra, India

Himshruti Yemul Walchand Institute of Technology, Solapur, Maharashtra, India

Parth Zanvar KLS Gogte Institute of Technology, Belagavi, Karnataka, India

Web Application for Monitoring Foundry Processes Using IoT

Sharada M. Kori, Mahesh A. Kori, Akshata S. Kori, and Parth Zanvar

Abstract Web application for monitoring Foundry Processes using IoT tracks all the manufacturing processes like pattern making, molding, melting, pouring, injection, cleaning, fettling, and inspection. Currently, the industries carry out the process of tracking by manual means of report maintenance, but by implementing the web application, all of the works related to the monitoring of the industrial process can be done on the phone/laptop itself. The customers who have placed order for any product can also track the status of their order and keep themselves updated. The owner/CEO of the company too can track what all activities/tasks are being carried out and can also check with the status of the products being manufactured. He/she need not go onsite for tracking but can track from his home or from anywhere in the world by just one click. Hence, at the end of the entire process of manufacturing, we also get to the success rate, i.e., the total products manufactured from the given raw materials and what was the wastage. We have also implemented a temperature monitoring system which monitors the temperature of the furnaces used during the melting process. As the temperature near the furnace is too high, it is very risky for the workers to go near it, and hence by implementing the temperature monitoring system using IoT, we can minimize the human involvement near the furnace.

Keywords Web application · Foundry Process · IoT · Automation · Monitoring system

S. M. Kori (✉) · M. A. Kori · P. Zanvar
KLS Gogte Institute of Technology, Belagavi, Karnataka, India
e-mail: kori.akshu@msrit.edu

A. S. Kori
M. S. Ramaiah Institute of Technology, Bengaluru, Karnataka, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_1

1 Introduction

1.1 Web Application

Web applications are computer programs that are hosted by a web server and accessed via a web browser over a network such as the Internet. They are used to perform various functions such as content management, customer relationship management, and e-commerce [1]. Web applications range from simple static web pages to complex dynamic web sites and web services. Web applications enable users to interact with a web server through a graphical user interface (GUI) or web page [2].

Web applications are software programs that are designed to be accessed over the internet through a web browser. They can be used for a wide variety of purposes, including communication, collaboration, data management, entertainment, and more.

Some common examples of web applications include webmail (such as Gmail or Outlook.com), online calculators, social media platforms (such as Facebook or Twitter), e-commerce shops (such as Amazon or eBay), and productivity tools (such as Google Docs or Trello).

While some web applications may be optimized for use with a specific browser, most are designed to be accessible from any web browser, including popular options like Google Chrome, Mozilla Firefox, Safari, and Microsoft Edge. This makes them highly flexible and widely accessible to users across different devices and platforms.

How Web Applications Work

Web applications are accessed through a web browser, which sends a request to the web server. The web server processes the request and sends back a response, which is displayed in the user's browser.

Behind the scenes, web applications typically consist of multiple layers. The web server manages the incoming requests and serves up the appropriate resources, such as HTML, CSS, and JavaScript files. The application server handles the logic of the application, processing user inputs and generating responses. The database stores any necessary data, such as user accounts, product listings, or transaction records [3].

Web applications are typically developed using a combination of front-end and back-end programming languages. HTML, CSS, and JavaScript are used for client-side programming, which creates the user interface and handles user interactions. Back-end programming languages like Python, Java, or Ruby are used to build the server-side logic, which processes data and generates responses.

One advantage of web applications is that they can be developed relatively quickly and easily by small teams. This is because they use standardized technologies and can be accessed from any web browser, making them highly portable and accessible to a wide audience [4].

1.2 JavaScript

JavaScript is a high-level programming language that is commonly used for creating dynamic and interactive web pages. It is a popular language for front-end development, which involves creating the user-facing parts of web applications such as forms, menus, and animations.

What JavaScript can be used for?

JavaScript is used to create interactive web pages, add functionality to web pages, validate forms, create cookies, detect browsers, create image rollovers, create calculators, and much more. JavaScript can be used to develop both client-side and server-side applications. It is most commonly used to create interactive web pages, but can also be used to create desktop applications, mobile apps, games, and more.

1.3 CSS

Cascading Style Sheet is used to control the style of a web document.

Resistance Temperature Detector

A resistance temperature detector (RTD) is a type of temperature sensor that is used to measure temperature in a wide range of applications. It works by measuring the resistance of a material, such as metal, as its temperature changes. RTDs are commonly used in industrial applications such as process control, power plants, and manufacturing facilities. They are also used in medical equipment, food and beverage processing, and other areas where accurate temperature measurements are necessary. Unlike thermocouples and thermistors, RTDs are not self-powered and require an external source of power to operate [5].

A Platinum Resistance Temperature Detector (PRTD) is a type of temperature sensor used for measuring temperatures of liquids and gases. It is a solid-state device consisting of a coil of fine resistance wire made from a platinum alloy, wound on an insulating ceramic core. The resistance of the wire increases linearly with temperature, allowing the PRTD to be used as a thermometer with a high degree of accuracy. The PRTD is a popular choice for industrial, medical, and scientific applications, due to its durability, accuracy, and low cost.

Types of resistance temperature detector

- Copper—0–180°.
- Nickel—220 to 300°.
- Tungsten—200 to 1000°.

1.4 JSX

JavaScript XML is a special kind of JavaScript object that can be used to represent and store XML data. It can be used to parse, traverse, and manipulate XML documents. JavaScript XML documents can also be converted to and from JSON, HTML, and other data types.

2 Methods and Materıal

To design and develop a web application for Monitoring Foundry Processes using web widgets on Mobile Phones and implementation of IoT.

2.1 Objectives

Input Objectives

1. Input design is a critical phase in the development of computer-based systems. It involves the process of creating a plan for how user input will be captured, validated, processed, and stored by a computer system. The input design process aims to ensure that the data entered into the system is accurate, complete, consistent, and relevant to the needs of the users and the organization. It involves understanding user requirements and designing user interfaces, data capture forms, and input validation routines that facilitate data entry and minimize errors.
2. By designing an effective data entry screen, users can enter data more efficiently, reduce errors, and increase productivity. The screen should be intuitive and easy to navigate, with clear instructions and guidance to assist users in entering accurate data. Additionally, the screen should allow for easy editing and correction of data, as well as provide a way to view and review the entered records.
3. When data is entered into an information system, it is important to ensure that the data is accurate, complete and conform to the required format. To achieve this, data validation checks are performed to ensure that the entered data is valid and meets the required criteria. If any errors or discrepancies are found, appropriate messages are displayed to prompt the user to correct the input.

Output Objectives

1. Designing computer output requires a systematic and organized approach, with a focus on meeting user requirements and providing a user-friendly experience. By following a structured design process and considering various factors, the resulting output can be effective, efficient, and meet the needs of the users.

2. Choose appropriate methods for presenting the information.
3. The primary objective of the output is to convey information about past activities, current status, or projections of the future. The output should provide users with a clear understanding of the state of the system, including its performance, trends, and potential issues. Additionally, the output should signal important events, opportunities, problems, or warnings to alert users to significant changes or issues.

2.2 Existing System

The existing system is a manual one in which the industry maintains and tracks all the industrial activities in the books and one can access the data by looking at the books itself, and there is no source through which the data can be accessed by means of online sources. Thus, maintenance of the data is also a tedious job and one has to maintain records of up to last 10 years, and thus, maintaining them in the books is unreliable.

Disadvantages of existing system

- Recording and maintaining become a tedious job.
- The furnace environment is very warm and there is risk involved if there is human intervention in and around the furnace; thus, by the implementation of the sensors, human interaction can be minimized and necessary protocols could be followed if the temperature is known well in advance.
- There is no existing research on industrial process management and recommendations.

Proposed System

- Here, we focus on helping the employees as well as the customers.
- The employees can utilize the web application and can monitor the entire manufacturing process and can keep track of the various processes like casting, molding, etc.
- The customers on the other hand can also keep a track of the status of the order they have placed and can make suitable decisions depending upon the status with respect to the future course of action.
- The other benefit of the system is that it also makes sure that there is safety involved while working in and around the furnace and can help with monitoring the temperature of the furnace.

2.3 Requirement Specifications

Requirement specifications are essential for any project. They provide a framework for the development process, that includes an understanding of the customer's

requirements, the software's capabilities, and the team's resources. The requirements are categorized into two types and are listed below.

2.4 Functional Requirements

Functional requirements are indeed an essential part of a requirement specification document. Their main purpose is to describe the specific functions that the system or software application should perform to satisfy the needs of its users or stakeholders. The depiction of the requirements of the system is presented below:

- The initial stage should be a login page for the company to access the portal.
- The company details and the services offered should be displayed.
- It should provide services like live temperature monitoring and live monitoring of the industrial activities.
- Live industrial activity monitoring should include the live status of the foundry manufacturing process.

2.5 Non-functional Requirements

Non-functional requirements are a critical part of a requirement specification document and describe the features, characteristics, and attributes of a system that are not directly related to its functional behavior. Based on these, the non-functional requirements are as follows (Fig. 1):

- The system should be secure and portable.
- The system should have an easy to use, user-friendly frontend.
- The system should perform efficiently in a short amount of time.
- The system should provide easy maintenance.
- The system should be easily scalable.

2.6 Hardware and Software Specifications

Hardware Specification

- Arduino UNO.
- 3-wire PT100 sensor.
- LCD with PCF8574.
- MAX 31865 IC.

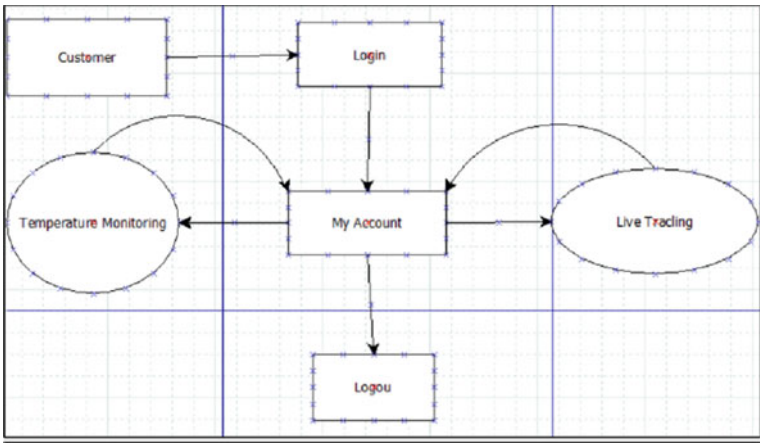


Fig. 1 Data flow diagram

- SSOP footprint.
- Resistance temperature detector.

Software Specification

Operating system: Windows/Linux/MacOS.

Programming language: JavaScript.

Web technologies: MERN Stack.

Mobile platforms: Android/iOS.

Cloud platforms used: GCP, AWS IoT.

Development IDE: Visual Studio Code.

3 Design Details

3.1 Data Flow Diagram

A data flow diagram (DFD) is a graphical representation of a system or a process that shows the flow of data through it. It is a visual representation of the inputs, outputs, processes, and data stores of a system, and it is used to describe how data moves through the system.

A DFD typically consists of four components.

1. Entities: Entities are the sources and destinations of data. They can be external to the system (such as customers or suppliers) or internal to the system (such as departments or subsystems).
2. Processes: Processes are the activities that transform data from one form to another. They represent the actions that take place within the system.
3. Data Flows: Data flows are the paths through which data moves from one entity or process to another. They represent the movement of data from one place to another.
4. Data Stores: Data stores are the places where data is stored within the system. They represent the repositories of data that are used by the system.

Once the customer accesses the portal, he/she needs to login; upon logging in, it will lead to the My Account section which in turn will give options such as Temperature Monitoring, Live Tracking, and the option of Logout. The Temperature Monitoring option will help the supplier to check the temperature in and around the furnace, and the Live Tracking option will help the customer to check the current status of the product he/she has ordered and the total success rate.

3.2 Use Case Diagram

A Use Case Diagram is a graphical representation of a system that depicts the interactions between actors (users or other systems) and the system itself. It is used to describe the functionality of a system from a user's perspective and to identify the different ways in which a user can interact with the system.

A Use Case Diagram typically consists of the following components:

Actors: Actors are the external entities that interact with the system. They represent the different roles that users or other systems play in the system.

Use Cases: Use Cases are the specific tasks or functions that the system can perform. They represent the features or services that the system provides to its users.

Relationships: Relationships are the connections between actors and use cases. They represent the interactions between the actors and the system.

System Boundary: The system boundary represents the scope of the system being modeled. It defines what is included in the system and what is outside of it.

Use Case Diagrams are commonly used in software engineering to capture the functional requirements of a system and to communicate the system's functionality to stakeholders. They are also useful for identifying potential errors or gaps in the system's design and for ensuring that the system meets the needs of its users (Fig. 2).

The Use Case Diagram above helps us to understand the basic functionalities of the proposed system which includes home page, Login, Temperature Monitoring, Live Tracking, and Logout which can be accessed by both customer as well as the admin. Other options like user data management, organizing sensor data can be accessed only by the admin.

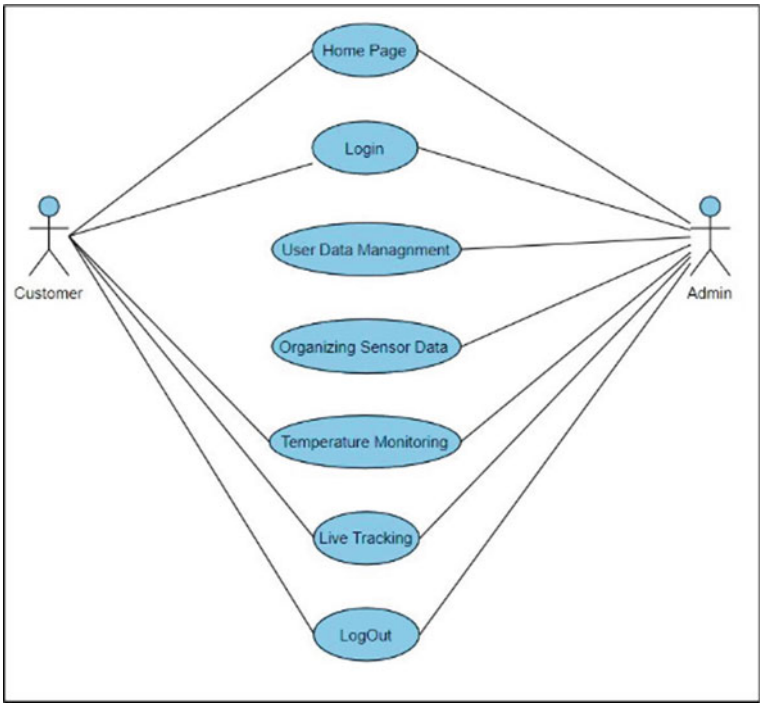


Fig. 2 Use case diagram

3.3 Activity Diagram

An Activity Diagram is a graphical representation of a system that depicts the flow of activities or processes within the system. It is used to describe the sequence of activities or steps that are required to complete a process and to show the different paths that can be taken through the system (Fig. 3).

Here, in the above Activity Diagram, we can understand the flow of the entire process, upon start the home page pops up which will ask for the login credentials, and for successful login, it will jump to My account which will have three options like Live Tracking, Temperature Monitoring, and Logout. If the login is unsuccessful, it will return to the home page.

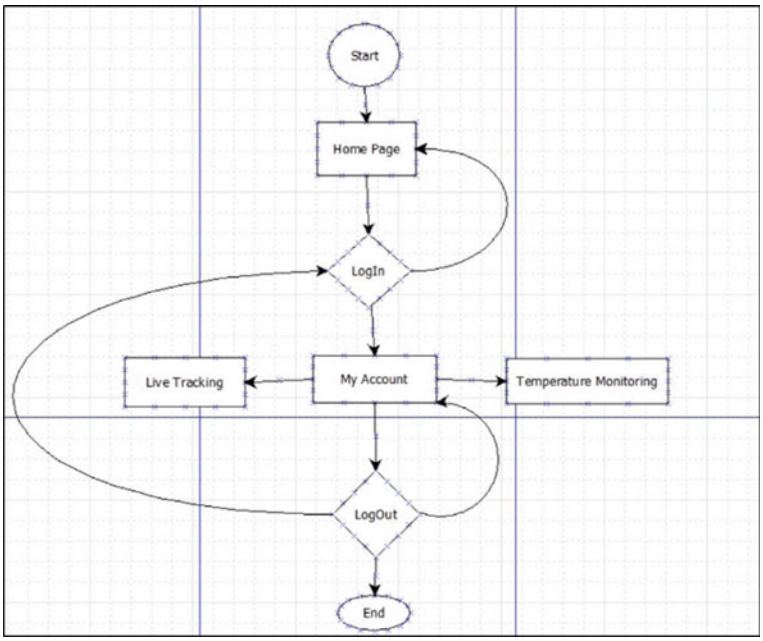


Fig. 3 Activity diagram

4 Results

4.1 System Design

See Figs. 4, 5, and 6.

5 Drawbacks in the Current System

The current system being used in the industries does not have the required facilities to implement the above-mentioned methodology of monitoring temperature and no such arrangement is also available in the actual scenario; hence below, we have discussed possible arrangements which could be made so as to implement the above-mentioned methodology feasible (Figs. 7, 8, 9, 10, and 11).

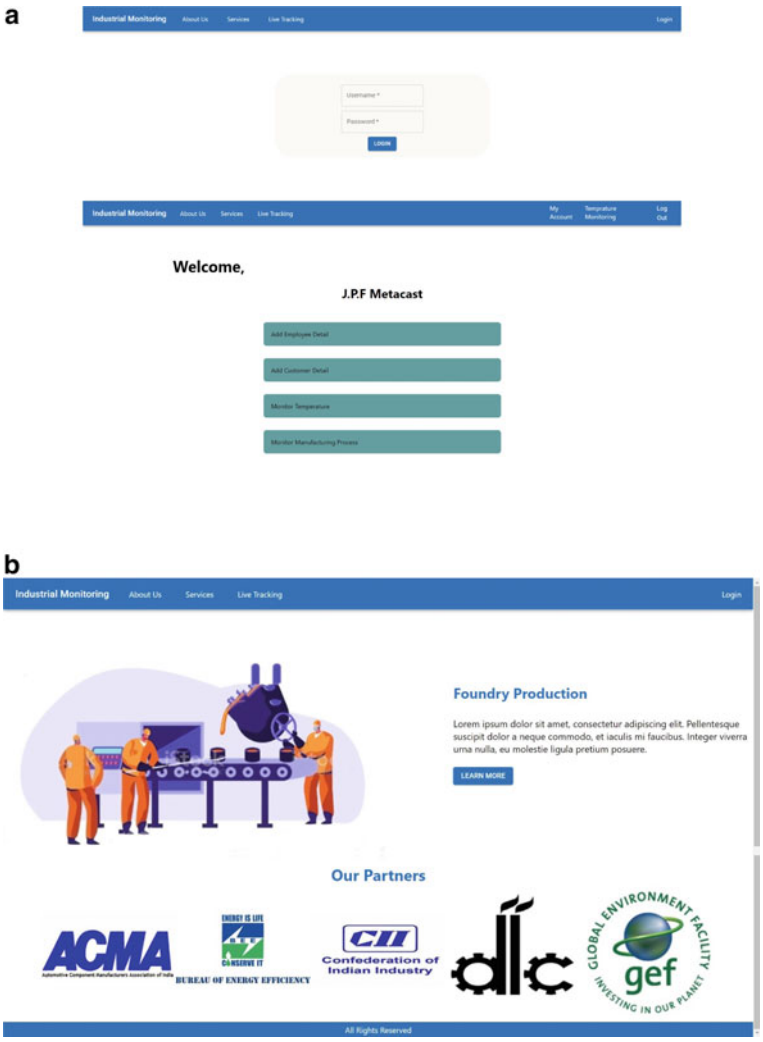


Fig. 4 **a** Login page; **b** Home page

6 Conclusion and Future Scope

The industrial monitoring system helps the employees as well as the CEO of the company to track the entire process of manufacturing, hence making it easy rather than the old traditional mechanisms of recording the data.

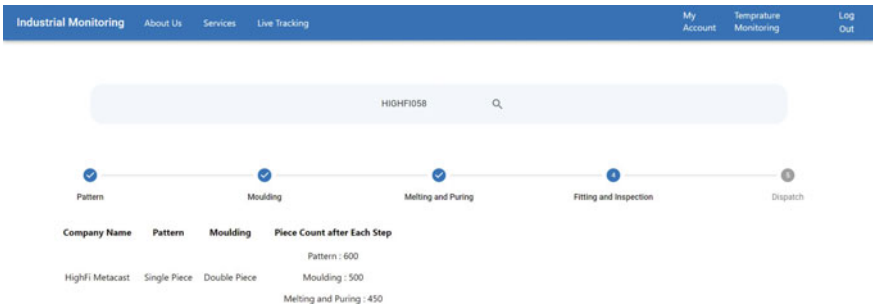


Fig. 5 Monitoring parameters

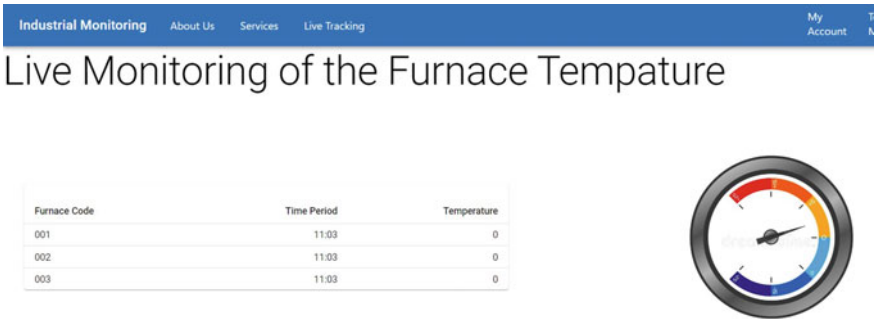


Fig. 6 Live monitoring of furnace temperature

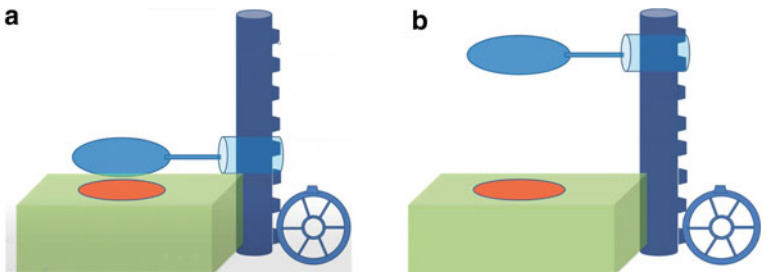


Fig. 7 **a**, **b** Design 1: with rack and pinion arrangement

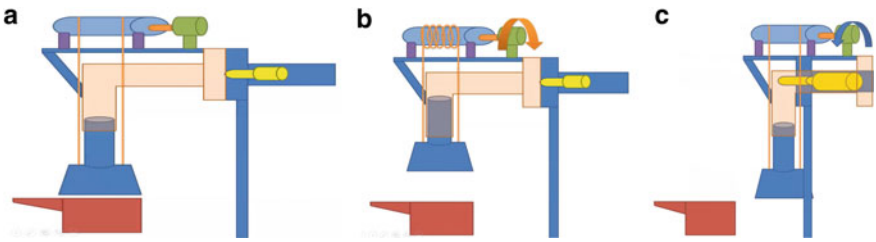


Fig. 8 **a–c** Design 2: with motorized and hydraulic arrangement

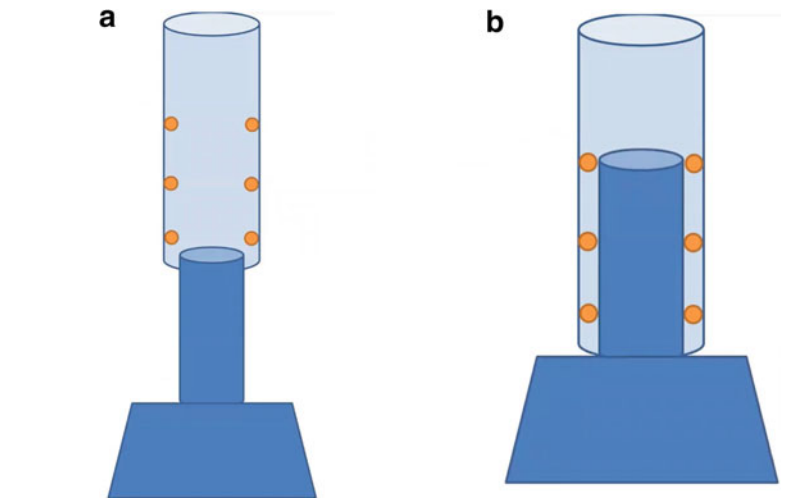


Fig. 9 **a**, **b** Draft tube arrangements

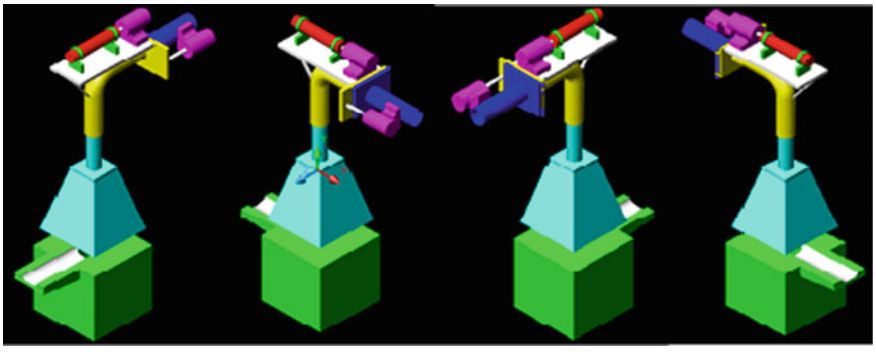


Fig. 10 Final optimized design

The monitoring of temperature using resistance temperature detector ensures safety of the workers and also helps in keeping track of the temperature around the furnace. In addition, it also helps to monitor the heat released out of the furnace and gives scope of how the same heat can be reused for various other industrial purposes.

Apart from just monitoring the temperature of the heat being released out of the furnace, the resistance temperature detectors can be also used to measure the temperature of the liquids that flow through various pipes within the industry and necessary actions can be taken by monitoring at which point the temperature drops and necessary steps can be taken to maintain the temperature.

Fig. 11 Final implemented design

Thus, by implementing the above-said techniques and methodologies, we can ease the entire process of industrial monitoring system and can also ensure safety of the workers by minimizing human involvement.

References

1. Wadhwa RS (2012) Flexibility in manufacturing automation: a living lab case study of Norwegian metalcasting SMEs. J Manuf Syst 31:444–454
2. Fisk G, Chandran R (1975) Tracing and recalling products. Harvard Bus Rev 53:90–96
3. Florence D, Queree C (1993) Traceability—problem or opportunity. Log Inform Manag 6(4):3–8
4. Vedel-Smith NK, Lenau TA (2012) Casting traceability with direct part marking using reconfigurable pin-type tooling based on paraffin-graphite actuators. J Manuf Syst 31:113–120
5. Santos EJP, Vasconcelos IB (2016) RTD-based smart temperature sensor. IEEE

BlueKei as an Example of Technology-Based Efficient and Cost-Effective Production Processes

Pradnya Vishwas Chitrao, Pravin Kumar Bhoyar, and Rajiv Divekar

Abstract Today everyone across the globe is attempting to implement an environment-friendly and a sustainable economic system. Businesses around the world are striving to infuse the SDGs in their processes and consequently get streamlined and more efficient. Currently, manufacturing units are utilizing systems engineering and systems design principles to survive the competition and make their product development processes more cost-effective and sustainable. The paper analyses the case of an enterprise, namely **BlueKei,** which was begun during the COVID-19 pandemic and which is a knowledge partner to organizations from different sectors, in order to learn how it contributes to sustainability through their services which are based on systems engineering and systems design. The study also seeks to find out how the venture managed to be profitable despite the corona lockdown.

Keywords Systems engineering · Systems design · Streamlined manufacturing process · Cost-effective · Efficient · Sustainable

1 Introduction

For daily functioning of businesses, leaders have to take many decisions both strategic and tactical. Time and again, external and internal disturbances affect these decisions, and organizations usually respond to them by adjusting business processes. Some factors that influence these decisions could be government policy around labour laws, implementation of emission norms, change in product demand cycle, etc. Sometimes crises like the corona pandemic can create for some time a total closure of businesses. Such crises create uncertainty and a very volatile environment that makes it difficult for businesses to re-start their operations. Leaders have to equip themselves with various new methods to tackle such situations and to keep the business going. Organizations are learning to respond to such tremors by looking beyond things like

P. V. Chitrao (✉) · P. K. Bhoyar · R. Divekar
Symbiosis Institute of Management Studies (SIMS), A Constituent of Symbiosis International Deemed University, Pune, Maharashtra, India
e-mail: pradnyac@sims.ed

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_2

balanced sheets and assessing the big picture and the entire ecosystem of operations. This big picture can be created accurately with the help of different Systems Thinking methodologies. Through the use of these Systems Thinking methodologies, business leaders gain deeper insights that enable them to reconstruct resilient operations, create a dynamic roadmap for the future and thus facilitate decision making. These methodologies support sustainability. Business or corporate sustainability involves achieving increased productivity while simultaneously reducing consumed resources without compromising on the quality and efficiency of a commodity or service or its profitability. Business feasibility is thus termed the triple bottom line (3BL) consisting of three critical aspects namely, financial (profit), social (people) and environmental (planet) [1].

2 Literature Review

The twenty-first century has brought to the forefront the difficulties experienced by every industry while bringing products to market. The corona pandemic has increased the number and the intensity of these difficulties. But, at the same time, history is proof of the fact that pandemics and similar crises can be catalysts for major social and economic transformations. But to overcome the adverse effects of a pandemic, one has to ensure fast innovation, optimized speed-to-market and scalability.

The present century is now regarded as the age of digitalization. This digitalization entails networked intelligence. Today information is very easily available, in the required framework, whenever one requires as a result of which organizations are acquiring new capacities to reframe business functions that yield faster and more precise service delivery.

During the entire span of a product or service, an elaborate Systems Thinking method pays attention to not just the intricate commodity and its various subsystems but also the larger systems of which it is a part. When it is a question of the systems level, the ones who use it, the surroundings, other commodities and procedures all have an effect on a product during its entire lifespan. Organizations who make preparations taking into consideration these points do perfectly well in their procedure. They thus put into practice “sustainable operations management”, which involves the integration of benefit and efficacy with the organization’s stakeholders and the resultant effects on the environment [2].

Manufacturers today desire to have a competitive edge and so achieve this by devising ways of successfully reducing the expenses in their cycle of commodity development, especially in the case of the creation of a new commodity. But generally, the cost of production (which covers the expenses of labour, of material and of processing cycle in production process) invariably comes to the basic acceptable level. So, businesses are constantly searching for new ways of curbing expenses by raising the competency level of design and by cutting down on waste in design. Most companies do not face so much waste on the assembly line as in the engineering section. Waste in aerospace and defence product development programmes comes

to ~ 60–90% of the charged time, with around 60% of all functions being no utilized at any given time [3].

In the process of product design, future-proofing concentrates on designing a product with the aim of minimizing technical outdatedness and other effects of risks at a later date. This task is carried out by product designers in every industry, where there is a requirement to have a holistic point of view and to consider every factor like electronic parts and availability of materials, novel technologies and methods, cross-industry learnings, new regulatory compliance requirements in future and sustainability aims.

Product design is a very intricate and unique process [4]. In product design, future-proofing covers designing of a product with the purpose of reducing to a minimum technical outdatedness and other effects of risks in future. This task is now very crucial for product designers in every realm, who are called upon to have a broader view point and to take into account everything from electronic parts and material availability, novel technologies and techniques, learnings across industries, to future regulatory adherence needs and sustainability aims.

Design teams have the capability of changing variables and noticing outcomes more quickly than they could with just a physical prototype by simply creating a real-time digital clone—or virtual representation—of the full production process. This minimizes costs and schedules, makes the most effective use of processes and enables one to iterate on designs more effectively and efficiently.

The use of Systems Thinking tenets in systems engineering is symbiotic that yields more efficient systems, products and designs. Systems Thinking is a package of instruments [5] that is based on the tenet that the equations between system parts and between the parts and the surroundings are as significant (with respect to system behaviour) as the parts themselves.

Systems engineering is an interdisciplinary strategy and a way of enabling the evolution of effective systems. It gives importance to spelling out customer requirements and needed functionality at an early stage in the development chain, putting on record client needs and then moving on to design, synthesis, checking, deployment, maintenance, modification and ensuing clearance of a system. Systems engineering synthesizes a huge array of engineering disciplines into a team effort and which utilizes a planned manufacturing procedure that involves moving from an earlier idea to manufacturing and operation of a system. It takes into consideration both the company and technical needs of all customers so as to come up with a quality product that satisfies all users' needs [6]. It is an ecological approach that combines the configuration, manufacturing and usage of products so as to minimize the use of resources. It comes up with a different approach that pays attention to the flows of matter and energy in the industrial setting and thinks of how a more environment-friendly and efficient procedure of manufacturing products and services could be designed [7]. Systems Thinking and systems engineering are interconnected, and when Systems Thinking concepts are applied to systems engineering, the end result is a superior system [8].

3 Research Objective

The researchers' aim was to study how entrepreneurs use technology for coming up with solutions for their client companies who are looking at ways of developing product processes that are lean, efficient, cost-effective and sustainable in order to overcome problems created by crisis time like the COVID-19 lockdown.

3.1 Research Methodology

The researchers did primary research by interviewing an entrepreneur who started a venture, namely **BlueKei Solutions Pvt., Ltd.,** to help business leaders solve specific operational or engineering issues, incorporate emerging technologies and acquire operational excellence through implementation of digital engineering methods. They also did secondary research by referring to research articles in reputed research journals.

3.2 Findings

3.2.1 BlueKei Solutions Pvt., Ltd.

Technology can have a critical role in coming up with streamlined and competent processes. It can guide you in reducing or removing replications and setbacks in the workflow, while at the same time enable you to accelerate by automating specific functions.

BlueKei Solutions offers novel approaches to empower decision-makers, transform businesses in technology adoption and integration in complex environments. It is a consultancy company based in Pune that specializes in creating Digital Engineering practices in organizations through systems engineering methodologies based on the Systems Thinking approach. It also offers capability development programmes in order to sensitize companies.

The company functions on the B2B and B2G business model. The Core Team consists of the two founders namely Ms Stueti Gupta and Mr. Ajit Mutalik. It was incorporated in May 2020. In October 2020, the company hired a Technical Director. In the same month, the venture entered into a business partnership with Equilibrant Force, South Asia. In December, the enterprise bagged its first successful business contract and it started generating revenue. In May 2021, the venture acquired its first international client.

The venture's client segment includes companies in aerospace and defence, automotive, mobility and transportation and Space Tech. Its target audience comprises

small, medium as well as large enterprises. Its clients are in India and in Asia Pacific areas.

3.2.2 The Genesis of the Venture

Ms Stueti Gupta, an ME in Design Engineering from Birla Institute of Technology and Science (BITS), Pilani, and an MS in mechanical engineering from Cornell University (USA) and also a certificate holder in Systems Design and Management from Massachusetts Institute of Technology (USA) was employed for many years in a well-known manufacturing company in India. The coronavirus pandemic which caused a shutdown of all manufacturing units gave her the idea of starting her own business. When employed, her responsibility was to utilize visual and mechanical models to project how a product will look three to four years down the line. She could also consequently predict how a function will behave. She had earlier worked in aerospace industry where many managers wanted to follow the digital way of functioning. She therefore collaborated with her mentor Mr Ajit Mutalick to start in March 2020 the venture. Though it was started in the middle of the pandemic, the venture survived.

3.2.3 Strategizing for Success

The founders quickly identified systems industry space senior people and got in touch with them in order to receive mentoring. This activity helped them to refine their business pitch. They did not stop networking with venture capitalists, and this effort paid off. In December 2020, they received the first proper training project from Bosch, and in January 2021, they got Atlas Copco training project.

Perseverance and mentoring were the key to the venture's success. Founders went to Bangalore and met mentors in coffee shops. Two of them have now become the venture's business partners, one for Australia and one for Singapore.

For a long time, the venture operated online. From January 2022, it started an office in Kharadi. It also started an office in Bangalore in December 2022. The Pune office has eleven people all from different engineering fields. There are three directors again from different engineering backgrounds. Industrial visits are made by engineers every month with each engineer spending on an average one week with the client company.

Systems engineering is not formally taught in India. So, the venture recruits fresh engineering graduates and grooms them. The venture is completely bootstrapped. It started with a capital investment of Rs. one lakh as it was going to be only services and which were online. Registration costs had to be borne. Today the company is thirty per cent plus profitable.

Today, some of its clients are St. Gobains, DRDO, Tata Advanced, Bosch, Tata Motors, to name but a few.

3.2.4 How the Venture Contributes Towards Sustainability

The venture indirectly contributes towards sustainability. It first comprehends the regulatory requirements for manufacturing units. It changes every technique to adapt to each organization's goals. In fact, sustainability is a major feature of the whole requirement set. The venture does the architectural work of bringing together the different functions together (like the embedded software function, the IT functions, etc.) which lessens at the very beginning itself different issues. The venture uses varied approaches for different companies like automotive, aerospace, etc., but the discipline adopted is the same. The venture first spells out the process as there are different ISO standards all of which are not necessary everywhere. It finds out the interventions required to lessen risks and to increase efficiency. Some companies already have the tools. But these tools have to be modified to suit each company's specific requirements. The venture helps a company in this tailoring process. It is thus a knowledge partner to its client aiding them in their efforts to improve their efficiency and achieve their sustainability goals.

4 Conclusion and Recommendations

Organizations today are in dire need of such services to reduce their costs, streamline their manufacturing processes, reduce their manufacturing waste, increase their efficiency and thereby become sustainable. For the attainment of their mandatory SDGs, companies require the services of such ventures. They should also invest in the trainings offered by such ventures.

References

1. Elkington J (1997) Cannibals with forks: the triple bottom line of 21st century business. Capstone Publishing Ltd., ManKato
2. Kleindorfer PR, Singhal K, Van Wassenhove LN (2005) Sustainable operations management. Product Oper Manag 14(4):482–492
3. Oppenheim BW (2004) Lean product development flow. Syst Eng 7(4):352–376
4. Panchal JH, Paredis CJJ, Allen JK, Mistree F (2009) Managing design-process complexity: a value-of-information based approach for scale and decision decoupling. J Comput Inform Sci Eng 9(2):633–647
5. Monat JP, Gannon TF (2015) What is systems thinking? A review of selected literature plus recommendations. Am J Syst Sci 4:2
6. Kossiakoff A, Sweet W, Seymour S, Biemer S (2011) Systems engineering: principles and practice, 2nd edn. Hoboken, Wiley
7. National Academic Press Linking Science and Technology to Society's Environmental Goals (1996) Chapter: use a systems engineering and ecological approach to reduce resource use. https://doi.org/10.17226/5409
8. Lawson H (2010) A journey through the systems landscape. College Publications, London

Plant Disease Detection with Learning-Based Models

Vaishnavi Chillal, Suneeta V. Budihal, and Saroja V. Siddamal

Abstract In this paper we try to assess and evaluate the accuracy of these models' performances. The VGG19 model is a 19-layer deep neural network that extracts information from input images using a series of convolutional and pooling layers, followed by fully connected layers. Its simplicity comes at the expense of increasing complexity and computing requirements. Similar convolutional and pooling layers are used by 14 Deep Convolutional Neural Network (DCNN), but with larger filter sizes and more complex layer arrangements. As a result, it is quicker and more effective than the VGG19 model, while it performs less accurately on specific datasets. Images of 7 different plant species are included in the dataset and experiment uses 19,589 different images. All the plant leaves are trained on the VGG19, 14DCNN, Inceptionv3, and AlexNet models. The accuracy of the considered models demonstrate that the 14DCNN model outperformed the others.

Keywords Plant disease · VGG19 · 14DCNN · Augmentation · ResNet

1 Introduction

Infections inhibit plant growth even though they can cultivate a wide variety of plants. One of the most important factors affecting crop fertility is plant diseases. The agricultural specialist's support takes a lot of time and effort to identify diseases. The next technique is to use insecticides, and excess of it could reduce plant quality while affecting plant growth. Applying additional pesticide to plants without determining how much of a specific pesticide is needed for that crop can be harmful to the environment and to human health. Automated disease detection models are becoming essential in a number of industries, including precision agriculture, high-throughput plant phenotyping, smart greenhouses, etc. In order to reduce output losses and promote agricultural sustainability, disease management and control operations must

V. Chillal (✉) · S. V. Budihal · S. V. Siddamal
School of ECE, KLE Technological University, Hubballi, India
e-mail: vaishnavichillal06@gmail.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_3

be carried out [1]. This emphasises the significance of crop monitoring together with early and accurate disease diagnosis. Computational algorithms can be used to quickly capture, analyse, and screen images of plants for detect diseases [2].

The current review aims to describe the state-of-the-art identification and analysis of plant disease detection techniques using Deep Learning (DL), which enhances traditional Artificial Neural Networks (ANN) by adding more "depth" to the network [3]. As a result, the focus of this survey is on key advancements in CNN and a variety of inventions that intended to enhance CNN's functionality and accurately diagnose diseases. There are now a number of automated plant disease diagnosis methods that use artificial intelligence methods with little human participation [4]. The most effective image identification method is the Deep Convolutional Neural Network (DCNN) [5].

1.1 Literature Survey

A machine learning technique called random forest is used to classify data because it solves the problem of overfitting [6]. The random forest provides greater recognition accuracy with fewer visual data [6]. 160 photos of papaya leaves are used in the experiment. The accuracy of random forest, SVM, K-mean neighbour, and logistic regression is 65.33%, 40.33%, 66.76%, and 70.14%, respectively.

A model for detecting diseases using soft computing and supervised machine learning [7] is used here. For the purpose of improving picture data, this model includes image pre-processing, segmentation, and feature extraction. Weight modification for classification based on supervised learning. The input parameters for soft computing must be fine-tuned. ML techniques SVM, random forest, and ANN have detection accuracy of 72.92%, 71.88%, and 59%, respectively. 26 diseases and 14 field crops were employed by the author [8]. There were 54,306 colour, grayscale, and leaf segmented photos used in total for the experiment using transfer learning and starting from scratch with various training and testing distributions. AlexNet and GoogleNet are the deep learning architectures that are employed. The GoogleNet algorithm's accuracy with an 80:20 distribution and transfer learning is 99.35%, while AlexNet's accuracy with the same attribute is 99.27%.

The author of [9] employs the following deep learning models: Single Shot Multibox Detector (SSD), Region-based Fully Convolutional Network (R-FCN), and Faster Region-based Convolutional Neural Network (Faster R-CNN). Based on the Intersection-over-Union [IoU] between the object suggestions and the ground-truth RPN's train, the Faster R-CNN is an object detection method. The R-FCN is another object detection system, however instead of using traditional ROI, it employs position sensitive cropping. As a result, R-FCN has accuracy that is comparable to Faster R-CNN. SSD is a feed-forward convolutional network that uses feature mapping and prediction to cope with objects of different sizes. A dataset is constructed using the images of tomato diseases and pests that were taken under field conditions. Anchor boxes are used by the YOLO v3 model [3] in Faster R-CNN. To expand the

number of feature maps, the enhanced YOLO v3 model uses feature fusion. Prior box dimension is calculated using the K-mean approach and utilised in large-scale feature maps. DarkNet-53 is used in image pyramid for feature extraction and feature visualisation. The YOLO v3 deep learning architecture was employed. To predit and enhances nine previous boxes and drill them into scale feature maps. Tomato plant leaves of 15,000 different images were utilised in the experiment. SSD, Faster R-CNN, YOLO v3, and Improved YOLO v3 accuracy scores are 84.32, 90.67, 88.31, and 92.39, respectively. It is evident that Improved YOLO v3 provides better accuracy.

The author [10] used 87,848 images of leaves from 58 different plant classes. The distribution of the test and training is 80/20. The average epoch for all deep learning models is 100. The accuracy of the VGG deep learning architecture and the AlexNetOWTBn model are respectively 99.53% and 99.49%. In the image-based plant disease diagnosis with deep learning, the author identified leaves via feature extraction, segmentation, and classification of patterns [11]. The accuracy of the Artificial Neural Network [ANN] architecture is 98.59%. Adam optimizers are used in conjunction with the ReLu piece-wise linear activation function. The methodology includes pre-processing, augmentation, and data extraction for the image dataset. For the purpose of implementing the disease detection, the ANN model is implemented from scratch. For an experiment, the Novel 14 layered 14DCNN [12] is applied to 147,500 plant leaf images coupled with the augmentation techniques Basic Image Manipulation [BIM], Deep Convolution Generative Adversarial Network [DCGAN], and Style transfer [NST]. Overall classification accuracy for the 14DCNN model was 99.9655; weighted average precision was 99.7999; weighted average recall was 99.7966; and weighted average $F1$-score was 99.7968. The accuracy of the Soybean Specific Plant Disease Detection [13] study utilising CNN algorithm is 99.32%. In the experiment, a total of 12,673 images of soybean are used. Three convolutional layers and three pooling layers compose up the algorithm. CNN excels at drawing out details from images [13]. The Adam optimizer with batch size 100 for 1000 epochs was employed by the CNN model. The ReLu activation function was employed in the last layer. More accurate than L2 regulation is the base model with dropout layer. The literature highlights the need for a framework to identify the plant diseases using the available DL models and evaluate the performance of VGG19 and 14DCNN model. The listed objectives to accomplish the task area as follows. To compare the DL models for their performance evaluation over various metrics and To identify a suitable DL model for plant disease detection with best performance.

2 Proposed Methodology

The block diagram of the novel 14DCNN model and the VGG19 model will be shown in this chapter. Briefly describe convolution, pooling, dropout, and other crucial layers for building neural networks. Each model presented in this paper must undergo the

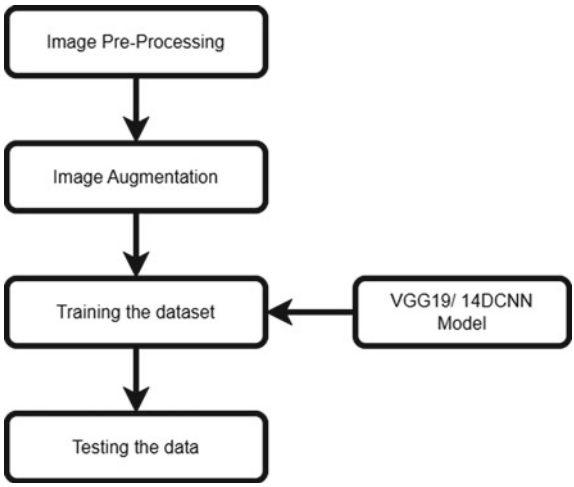


Fig. 1 Algorithm for the plant disease detection

procedure depicted in Fig. 1. The images in the dataset are first divided into training and testing datasets in an 80:20 ratio. The training images are then go through the following stages.

Image Pre-processing: Various pre-processing approaches are taken into consideration in order to reduce noise from images. Image clipping, or cropping, is the process of selecting the desired area of a leaf image. Image smoothing is done using the smoothing filter. **Image Augmentation**: It artificially generates training images by combining several types of augmentation, such as random rotation, shifts, shear, and flips. We could create images with rescale, shear range, zoom, and horizontal flip augmentation using the Image DataGenerator API. **Training and Testing Dataset**: Convolutional, pooling, and fully connected layers are among the layers that make up the DCNN, which can learn features from training data [14]. **Convolution Layer**: Equation 1 can be used to determine the convolution layer output's dimension. n_w is input width and n_h is input height. Additional f_w, f_h, f_c represent width, height, and channels of the kernel filter of the layer and s is stride number.

$$\text{Dimension}(\text{Convo}(n,k)) = \left(\left[\frac{n_w - f_w}{s} + 1, \frac{n_h - f_h}{s} + 1\right], f_c\right) \tag{1}$$

Pooling Layer: The phrase "max pooling" describes a pooling procedure that selects the biggest element from the feature map region that the filter encompasses. The dimension of pooling layer output can be calculated using Eq. 2. The parameter n_w, n_h, n_c is input width, height and channels. Remaining parameter f_w, f_h are width, height of kernel filter.

$$\text{Dimension}(\text{Pooling}(n,k)) = \left(\left[\frac{n_w - f_w}{s} + 1, \frac{n_h - f_h}{s} + 1\right], n_c\right) \tag{2}$$

Flatten Layer: A flattening layer is used to aggregate all of the 2-Dimensional arrays produced by pooling feature maps into a single, long continuous linear vector. **Dense Layer**: Here, z_j in Eq. 3: equation is output of first dense layer with ReLu as a activation function. x_i and w_i represent the value and weight of ith input of the jth output. The softmax activation function for the last dense layer output is shown in Eq. 4.

$$z_j = \text{ReLu}\left(0, \sum_{i}^{512} b_j + x_i w_i\right) \tag{3}$$

$$\text{softmax}(\sigma(z_i)) = \frac{e^{z_i}}{\sum_{j=1}^{59} e^{z_j}} \tag{4}$$

Dropout Layer: The Dropout layer serves as a mask, keeping all other neurons functional while removing some of their contributions to the layer below.

2.1 VGG19 Model

The trainable parameters in this VGG19 model are 6,423,555, whereas the non-trainable parameters are 20,024,384 (Fig. 2).

The trainable parameters are 6,423,555, and non-trainable parameters are 20,024,384. Convolutional output for the input dimension of 224, 224, 3 using Eq. 1 is 224, 224, 64. Using Eq. 2, the output layer dimension for the maximum pooling layer is 56, 56, 128 whereas the input is 128, 128, 128. The output of the input layer is multiplied to produce the flatten layer dimension, which for a (7, 7, 512) input results in (None, 25,088). Equation 3 can be used to compute the output of the first

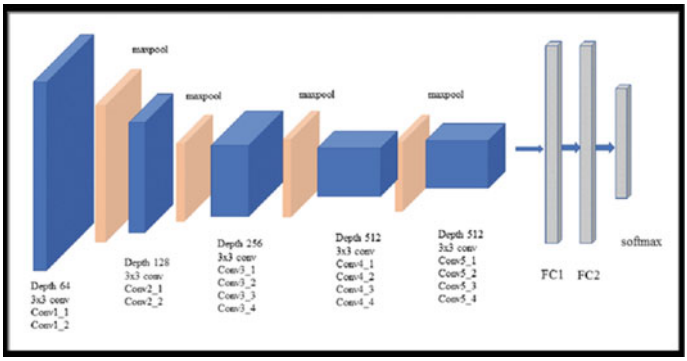


Fig. 2 Block diagram of VGG19 model

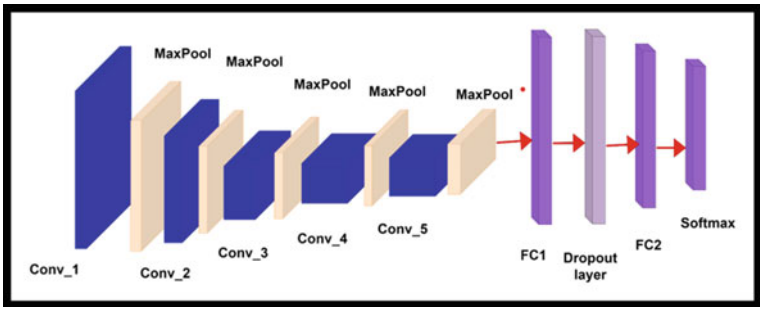


Fig. 3 Block diagram of 14DCNN model

dense layer, which is (None, 256), and Eq. 4 can be used to get the output of the last dense layer, which is (None, 3).

Specifications: The batch size for the optimised hyper parameters is 32. Optimizer used is Adam. Callbacks allow us to verify and adjust training behaviour. ReLu is used as the activation function in the convolution layer. Minimum learning rate for training is 0.000001. Softmax is used for last layer.

2.2 14DCNN Model

ReLu is used as the activation function in the convolution layer and the minimum learning rate for training is 0.000001. Loss is calculated using the categorical cross entropy with a dropout value of 0.2. Convolutional output for the input dimension of 128, 128, 3 using Eq. 1 is 126, 126, 4. Using Eq. 2, the output layer dimension for the maximum pooling layer is 63, 63, 4 whereas the input is 126, 126, 4. The output of the input layer is multiplied to produce the flatten layer dimension, which for a (2, 2, 128) input results in (None, 512). Equation 3 can be used to compute the output of the first dense layer, which is (None, 2048), and Eq. 4 can be used to get the output of the last dense layer, which is (None, 3) (Fig. 3).

Specifications: The batch size for the optimised hyperparameter is 32. Adam optimizer is used. ReLu is used as the activation function in the convolution layer. Minimum learning rate for training is 0.000001. Loss is calculated using the categorial cross entropy. Dropout layer value is 0.2.

3 Results and Discussion

This section includes dataset curation, data augmentation, model comparison, model accuracy, confusion matrix, and output format. Apple, Cherry, Corn, Grape, Peach,

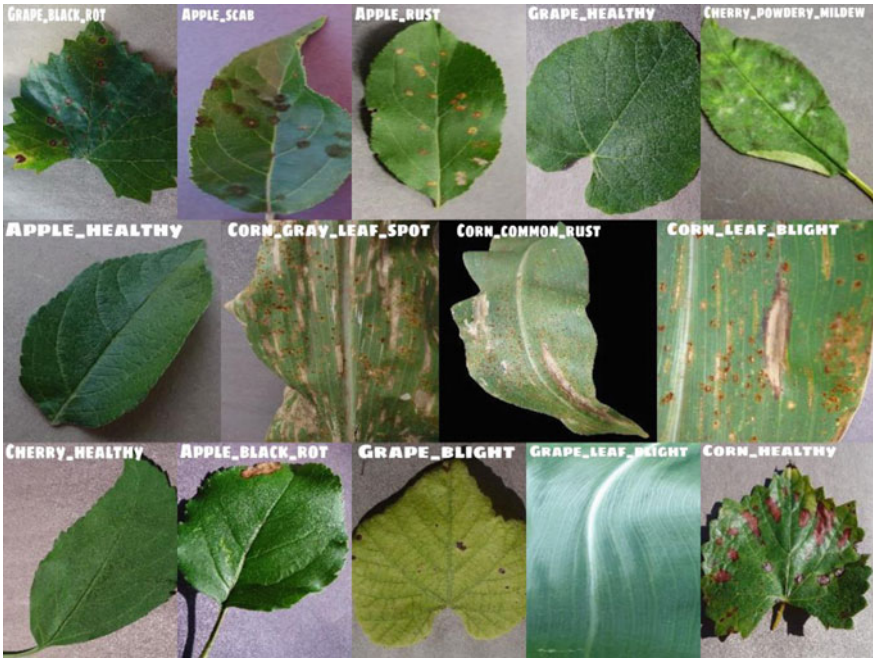


Fig. 4 Batch of shuffled plant dataset

Strawberry, and Potato plants utilised in the experiment are shown in Fig. 4 leading to 19,589 pictures.

3.1 Dataset Curation

The graphical distribution of the plant dataset is shown in Fig. 5, including healthy and diseased plants.

3.2 Data Augmentation

The response we obtained following argumentation is shown in Fig. 6. First column demonstrate the original image. Right to point out that rescale = 1/255 will change the pixels in the range [0, 255] to [0, 1], shown in second column of Fig. 6. If all of the images are scaled to the same range [0, 1], the loss will be scattered more uniformly [15]. The image will be sheared along an axis, as shown in the third column of

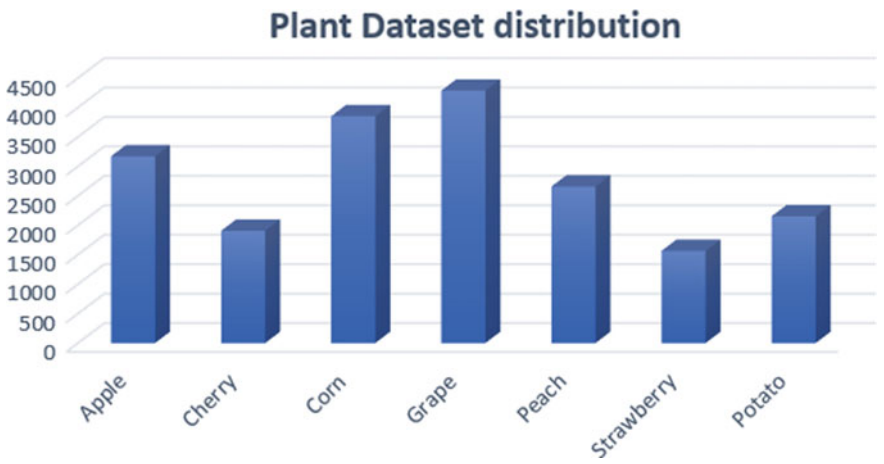


Fig. 5 Plant dataset distribution

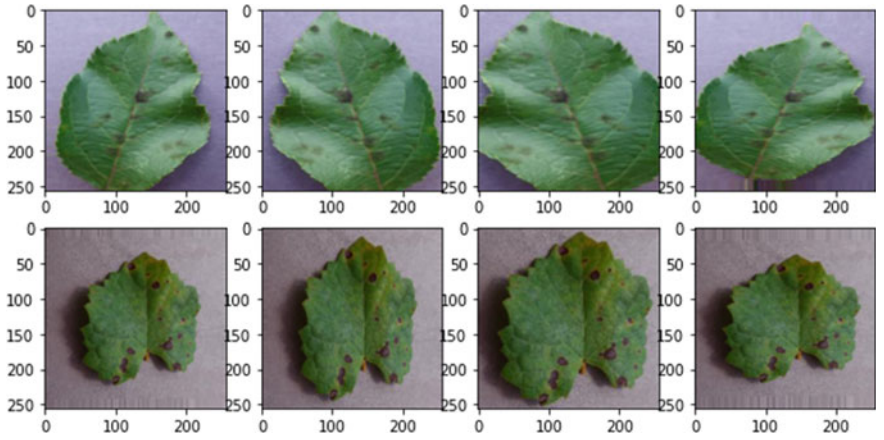
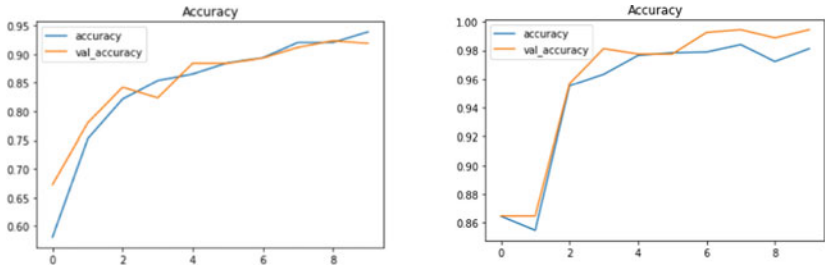


Fig. 6 Augmented plant images

Fig. 6, to produce or adjust the perception angles. By randomly rotate images across any angle between 0 and 360°, that is done by zooming, demonstrated in the fourth column of Fig. 6.

3.3 Model Accuracy

Figure 7a shows the accuracy of the VGG19 model, it uses its pre-trained weight for learning in the first epoch. The accuracy becomes stagnant after a certain epoch, which could lead to the model being overfitted. Figure 7b shows the accuracy of the



(a) Accuracy of the model VGG19 (b) Accuracy of the model 14DCNN

Fig. 7 Accuracy of the models

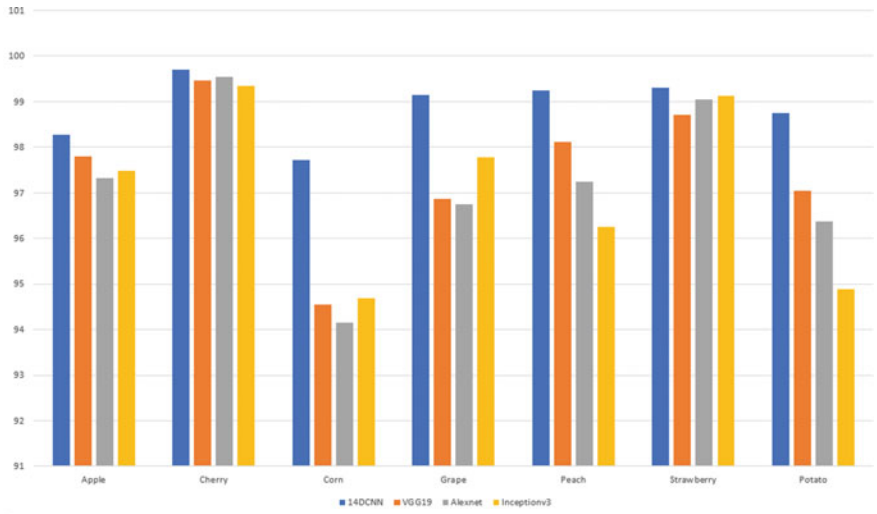


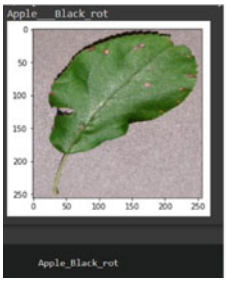
Fig. 8 Accuracy of different models for different plant dataset

14DCNN model, for training since the 14DCNN model is trained from scratch. To prevent overfitting of the model, there is no change in accuracy for the training or validation dataset after a predetermined number of epochs. The class-specific model accuracy percentage is displayed in Fig. 8. The disparate plant images and their class distribution are to bear responsibility for the irregular distribution. In addition to VGG19 and the novel 14DCNN model, we used state of art models like Inceptionv3 and AlexNet for comparison. The seven main classes of plants each have a unique distribution of accuracy. The output of the hierarchical model is shown in Fig. 8. By looking at each model's accuracy parameter, we can conclude that the 14DCNN model outperformed the others for every class dataset.

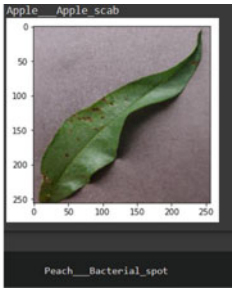
3.4 Training/Testing of Dataset

The results of the VGG19 model's subsequent image-based plant disease identification are shown in Fig. 9. After the model has been trained, the predict function is used to forecast the image with the proper disease. Figure 9a depicts the illness that has been accurately identified as apple black rot. The condition that was mistaken as a disease is shown in Fig. 9b.

Figure 10 displays the outcomes of the following image-based plant detection of diseases using the 14DCNN model. The predict function is used to anticipate the image with the correct disease after the model has been trained. The disease that has been correctly diagnosed as corn healthy is shown in Fig. 10a. Figure 10b depicts the condition that was misdiagnosed as a sickness. The crop should be resistant to bacterial apple black rust, according to the model. Comparison of VGG19 and 14DCNN model: In terms of accuracy, the 14DCNN model outperforms the VGG19 for small datasets. Since it uses transfer learning, the VGG19 takes longer to train and detect than the 14DCNN. Memory-wise, the VGG19 uses more memory to store transferred data than 14DCNN, which creates data from scratch.

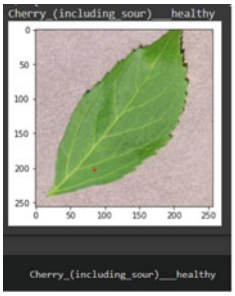


(a) Model correctly recognised the disease

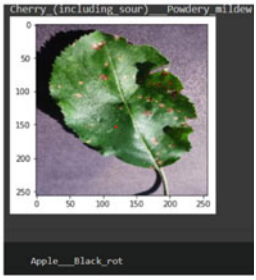


(b) Model misidentified the disease

Fig. 9 Results of the VGG19 model validation



(a) Model correctly recognised the disease



(b) Model misidentified the disease

Fig. 10 Results of the 14DCNN model validation

4 Conclusions

The VGG19 model and novel 14DCNN model methodology overview were presented in the project, and transfer learning was utilised with the VGG19 model. Image augmentation to prevent overfitting. Comparison of the most advanced models for plant disease identification, including VGG19, 14DCNN, Inceptionv3, and ALexNet. Even without transfer learning, the novel 14DCNN model outperformed the remaining models in terms of accuracy across all classes of plants.

References

1. Singh A, Budihal SV. Learning-based macronutrient detection through plant leaf. In: Advances in computing and network communications. Lecture notes in electrical engineering, vol 736. Springer, Singapore
2. Singh A, Budihal SV. Non-invasive techniques of nutrient detection in plants. In: Intelligent computing and applications. Advances in intelligent systems and computing, vol 1172. Springer, Singapore. https://doi.org/10.1007/978-981-15-5566-4_35
3. Liu J, Wang X (2020) Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network. Front Plant Sci 11:898
4. Rumpf T, Mahlein AK et al (2010) Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. Comput Electron Agric 74(1):91–99
5. Chen HC et al (2022) AlexNet convolutional neural network for disease detection and classification of tomato leaf. Electronics 11(6):951
6. Ramesh S et al (2018) Plant disease detection using machine learning. In: International conference on design innovations for 3Cs compute communicate control. IEEE Press, pp 41–45
7. Sehgal A, Mathur S (2019) Plant disease classification using soft computing supervised machine learning. In: 3rd international conference on electronics, communication and aerospace technology (ICECA). IEEE Press, pp 75–80
8. Mohanty SP, Hughes DP, Salathé M (2016) Using deep learning for image-based plant disease detection. Front Plant Sci 7:1419
9. Akila M, Deepan P (2018) Detection and classification of plant leaf diseases by using deep learning algorithm. IJERT 6(7):1–5
10. Ferentinos KP (2018) Deep learning models for plant disease detection and diagnosis. Comput Electron Agric 145:311–318
11. Dhakal A et al (2018) Image-based plant disease detection with deep learning. Int J Comput Trends Technol 61(1):26–29
12. Pandian JA et al (2022) Plant disease detection using deep convolutional neural network. Appl Sci 12(14):6982
13. Wallelign S, Polceanu M, Buche C (2018) Soybean plant disease identification using convolutional neural network. In: FLAIRS conference, pp 146–151
14. Leonardi C, Stagi F, Orsini G, Cavani A (2019) The architecture of trees. Princeton Architectural Press
15. Maniyar HM et al (2020) Plant disease detection: an augmented approach using CNN and generative adversarial network (GAN). In: Information, communication and computing technology (ICICCT), vol 1170. Springer, Singapore

A Bibliometric Analysis of Artificial Intelligence and Human Resource Management

Malabika Sahoo and Sumita Mishra

Abstract In the current business scenario, artificial intelligence (AI) is the trend and is becoming the new normal. There is an increasing application of AI in business and organizations in all sectors. The adoption of AI in the human resource department is not an exception. The different functions of human resource management such as recruitment, management of employees, and engagement have been changed by the application AI. The manual HR works are taken over by machines enabling HR managers to deal with more strategic roles. The aim of this work is to accomplish a bibliometric analysis which addresses the implementation of AI in human resource management. The bibliometric database Web of Science was chosen for the data extraction. The initial search yielded 3475 articles, out of which 210 were taken for final analysis. VOS viewer was used for producing the bibliometric information. The findings indicate the rising trend of using AI in HRM.

Keywords Artificial intelligence · Human resource management · Bibliometric analysis

1 Introduction

Industry 4.0 is marked by increased use of technology and artificial intelligence and its influence on different parts of the society [1]. The application of AI is increasing in organization for accuracy in work, better decisions, and enhanced proficiency. AI is defined as the application of digital technology to enable systems capable of performing tasks. Over the last decades, there have been significant advances and use of AI in different driplines [2]. AI is widely heralded as an advanced technology which will transform the world of work [3].

Human resource management is being reinvented critically by advanced innovative technologies [4]. The wider implementation of AI and other advanced technologies results in automation of varied managerial and routine tasks of HRM [5]. AI is

M. Sahoo (✉) · S. Mishra
KIIT School of Management, KIIT Deemed to Be University, Bhubaneswar, Odisha, India
e-mail: malabika.sahoo@ksom.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_4

reshaping the varied function of HRM. Different functions of HRM such as procurement, selection, performance management, training and development, employee engagement are effectively managed with AI applications [6].

The industry is attracted much by artificial intelligence bots and humanoid service robots [7]. It has been observed that these AI applications have transfigured the traditional HR functions resulting in better work management and better scope for HRM but at the same time posing many challenges [8]. In spite of the challenges, deep learning, machine learning, smart objects, internet of things (IOT) are useful for business as they foster enhanced coordination and co-operation [9]. Also, the digitalization of HR systems, i.e., electronic human resource information system (EHRIS) and additional advanced know-hows offers numerous prospects to increased efficiency and reduction of cost of varied HR aspects such as screening and evaluation of job applicants, employee performance appraisals [9, 10].

Owing to the increasing trend of AI application in HRM, the purpose of this research is to perform a bibliometric analysis on the rising trend of research in AI and HRM from 2013 to 2023. The work extends an organized and structured analysis of extant literature that draws linkage among AI, machine learning, advanced robotic technologies with HRM. In doing so, this will be an encouragement for more research work in this domain. The study revealed the use of AI in HRM and suggested how HR managers can overcome obstacles and do more accurate jobs with the help of AI.

2 Methodology

The research aims to scrutinize and present the extant literature on AI and human resource management. The methodology used was bibliometric analysis using VOS viewer. Bibliometric analysis refers to the quantitative synthesis of extant literature using bibliometric databases such as Web of Science (WOS) and Scopus that help researchers to understand the structure of knowledge in a particular area. Bibliometric analysis enables the researchers to present visual depiction of the data available in literature [11].

2.1 Search Criteria

The study used the WOS core collection bibliometric record for searching the literature and extraction of bibliometric data on 16.2.2023. WOS is considered as one of the best sources which contains quality articles. The database has more than 74.8 million scholarly data and data sets which covers around 254 varied areas [12]. WOS was used by researchers for performing bibliometric analysis containing quality publications [13]. The keywords used for data extraction are “AI” OR “Artificial Intelligence” OR “Machine Learning” OR “Deep Learning” OR “Neural Network” And “Human

resource management" or "Human resources" or "HRM" using the all-field option in WOS. Initial search yielded a total of 3475 numbers of documents. In terms of inclusion criteria, only journal papers published in English language from the year 2013 to 2023 were considered. Conference proceedings, books, and book chapters were excluded in the current study to warrant quality of research publication [14]. After screening for only journal publications in English language and time frame of 2013–2023, a total of 210 publications were retained.

3 Result and Analysis

The subsequent section discusses the outcome and findings of the research work, which aims to identify journals based on highest publications, year-wise publications, papers containing highest number of citations, most prolific authors, countries, etc. The result section will also analyze citation, co-citation, network analysis, intercountry, and keyword analysis in network maps. Figure 1 depicts the number of publications and citations from the year 2013 to 2023. The chart indicates a rise of publications and citations over the span of 10 years. The year 2021 is marked by highest number of citations.

3.1 Analysis of Citations (Highly Cited Papers and Authors)

Table 1 contains the data pertaining top five significant publications with highest citations. The paper Artificial Intelligence in Service is having highest citations (660) followed by Innovating through digital revolutions: the role of soft skills and big data in increasing firm performance (60).

Figure 2 illustrates the top ten highly cited authors in the domain of AI and HRM. Mostly, the top ten highly cited authors have published their papers during 2020–2023.

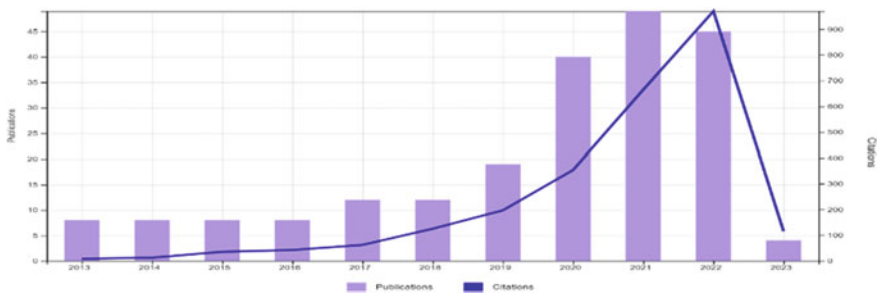


Fig. 1 Year-wise publications

Table 1 Highly cited papers

Authors	Title	Citations
Huang and Rust [15]	Artificial intelligence in service	660
Kheybari et al. [16]	Analytic network process: an overview of applications	60
Caputo et al. [17]	Innovating through digital revolutions: the role of soft skills and big data in increasing firm performance	60
Lukovac et al. [18]	Portfolio model for analyzing human resources: an approach based on neuro-fuzzy modeling and the simulated annealing algorithm	44
Pillai and Sivathanu [19]	Adoption of artificial intelligence (AI) for talent acquisition in IT/ITES organizations	35

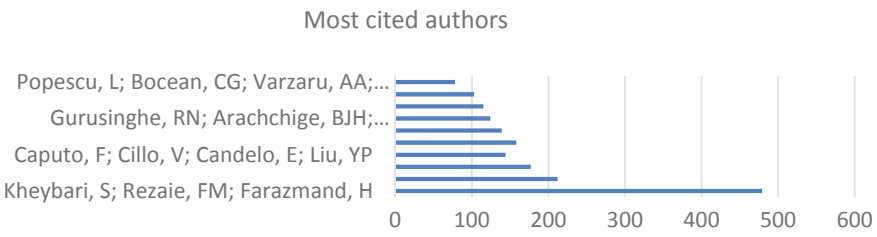


Fig. 2 Most cited authors

3.2 Co-Authorship Network Map

VOS viewer software was used to create network maps based on co-authorship, keywords, co-citation analysis, etc.

The visualization map of co-authorship (Fig. 3) indicated the state co-authorship in the data set. Out of 656 authors identified, only those researchers were considered who have co-authorship with one author, and hence, 20 met the criteria of being connected with each other. Ferreira is having the maximum connections followed by Martinez as the second highest.

3.3 Intercountry Co-authorship Map

Figure 4 represents the visualization inter-county network map. Sixty-three countries were identified in the map and 46 met the criteria of minimum number of two citations per country. From the chart, it is evident that authors of China, the USA, Australia, England, and India have co-authored with maximum numbers with authors of other countries.

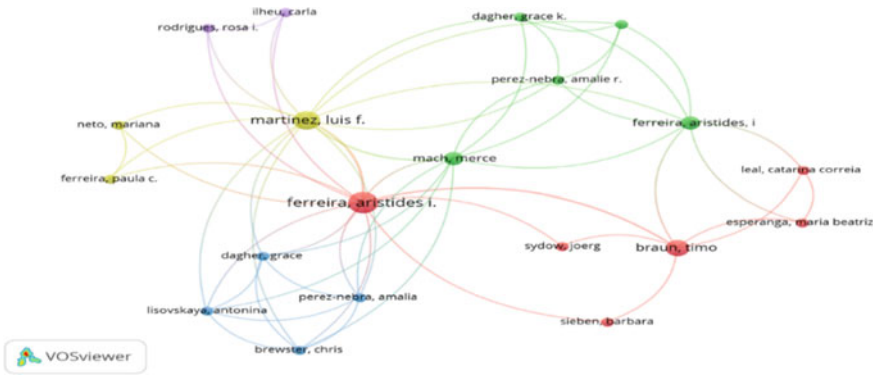


Fig. 3 Network map based on authors and co-authorship

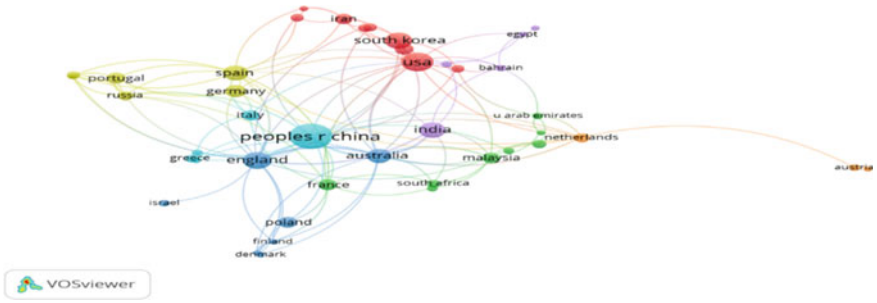


Fig. 4 Network map of intercountry co-author

4 Keyword Co-occurrence Map

The main aim of keyword co-occurrence visualization map is to identify the most prominent keyword used for maximum number of times. The keyword analysis helps the researchers to identify the research preference related to topic and context. From 1338 keywords, 38 keywords satisfied the conditions of six occurrence resulting in three different groups or clusters. The highest co-occurrence keywords related to application of AI in HRM are machine learning (28), artificial intelligence (26) human resource management (25), and performance (33). The keyword co-occurrence map is illustrated in Fig. 5. The different clusters are closely associated and partially integrated.

The co-word analysis of AI in HRM revealed three clusters such as Cluster 1 (blue color) with ten keywords which is labeled as machine learning and HRM, Cluster 2 (red color) with 15 keywords which is labeled as performance and HRM, and Cluster 3 (green color) with 12 keywords which is labeled as AI and its role in HRM. The density visualization (Fig. 6) indicates the impact factor of keywords.

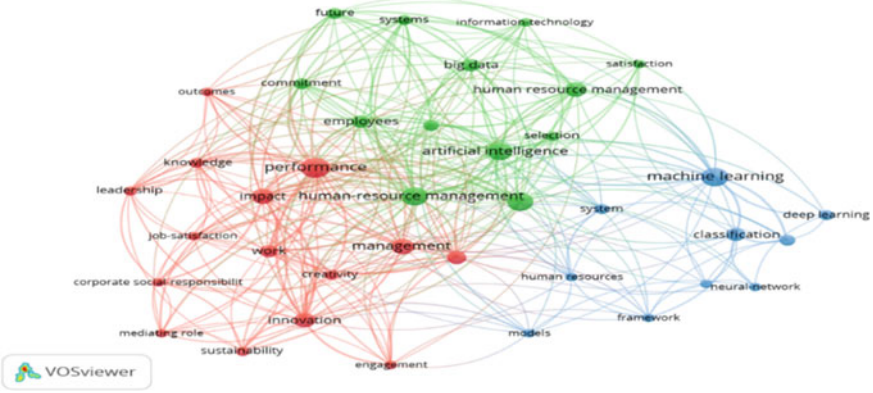


Fig. 5 Network map of keyword co-occurrence

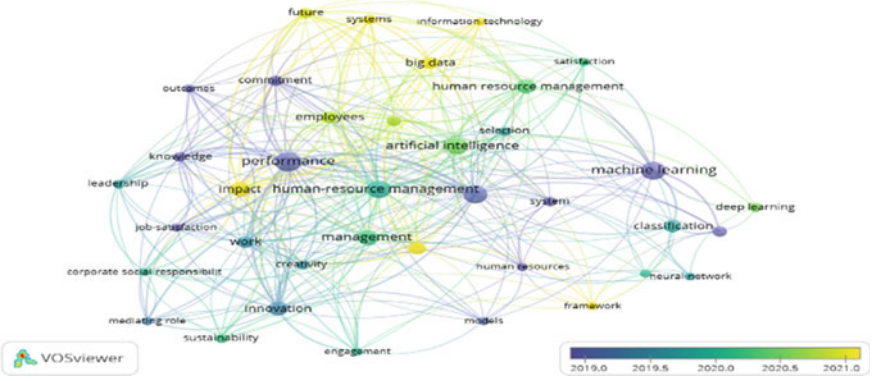


Fig. 6 Density visualization of keyword co-occurrence

5 Implications

The current review enriches the extant literature on AI and HRM through its contributions. In the competitive era, there is an uprise trend in the implementation of AI in HRM. The organizations are facing enormous challenges for adoption of AI in different domains of HRM. Thus, to understand the trend and role of AI in HRM in the last 10 years the study attempted bibliometric analysis to scrutinize the prominent research done in the domain, the prolific authors, country, etc. The review is an attempt to explain the research carried out on the use of AI in HRM that will enable increased performance by organizations by adopting AI. The study will be of great help to future researchers as a reference point for escalating and developing the content of this study.

The review also lends to some important implications for practice. The research evidence corroborates the increased adoption of AI in HRM over the years. The escalation of AI in different domains reduces the routine work of HR managers. Use of AI helps the organizations to take more informed decisions, enhanced performance, and reduction of human errors [20]. As mentioned by these authors, there is a need for managers in organizations to encourage the use of AI in different domains of HRM. The AI can be used in employee selection, learning and development, employee engagement, compensation, performance management, talent retention, etc. AI acts as a support function in human resource management and development. Most of the activities concerned with hiring, engaging, and retaining employees are professionally managed in a structured way with the help of AI.

In the current competitive business scenario, there are overwhelming adoption of technology, especially AI in HRM. The reasons owe to the enormous power of data and the decisive role for enhancing the performance of the individuals as well as organizations. The amalgamation of AI in the HRM domain helps the organization to make more accurate and better decisions. The use of AI results in enhanced employee productivity, reduced cost, improved organizational performance, automation of mundane human tasks, and predicting future outcomes [21]. Despite its significance, people have fear and negativity for the application of AI. However, in the long run, it is palpable that use of AI and technologies will no longer be options but will be the bare necessity for the organizations to remain competitive. Researchers have realized the importance of conducting research in this domain; however, there is dearth of ample empirical and experimental studies which opens the scope for future research in this domain.

6 Limitations of the Study

The study is not free from certain limitations. The review is restricted to using a lone database (Web of Science). Secondly, the study included only journal papers for maintaining quality [14]. The study could have included conference papers, book chapters, and books for wider collections. Thirdly, this work considered only bibliometric analysis, and future research could focus on combining detailed systematic literature review with bibliometric analysis in other domains for justifying the generalizability concern. However, despite its limitations, the study is an encouragement for researchers as we all practitioners in the domain of AI and human resource management.

References

1. Bolander T (2019) What do we loose when machines take the decisions? J Manag Govern 23:849–867
2. Vrontis D, Christofi M, Pereira V, Tarba S, Makrides A, Trichina E (2022) Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review. Int J Hum Resour Manag 33(6):1237–1266
3. Agrawal A, Gans J, Goldfarb A (2018) Prediction machines: the simple economics of artificial intelligence. Harvard Business Press, Brighton
4. Ancarani A, Di Mauro C, Mascali F (2019) Backshoring strategy and the adoption of industry 4.0: evidence from Europe. J World Bus 54(4):360–371
5. Larivière B, Bowen D, Andreassen TW, Kunz W, Sirianni NJ, Voss C et al (2017) Service encounter 2.0: an investigation into the roles of technology, employees and customers. J Bus Res 79:238–246
6. Kovach KA, Cathcart CE (1999) Human resource information systems (HRIS): providing business with rapid data access, information exchange and strategic advantage. Publ Person Manag 28(2):275–282
7. Thomaz F, Salge C, Karahanna E, Hulland J (2020) Learning from the Dark Web: leveraging conversational agents in the era of hyper-privacy to enhance marketing. J Acad Market Sci 48:43–63
8. Malik A, Budhwar P, Patel C, Srikanth NR (2022) May the bots be with you! Delivering HR cost-effectiveness and individualised employee experiences in an MNE. Int J Hum Resour Manag 33(6):1148–1178
9. Cooke FL, Veen A, Wood G (2017) What do we know about cross-country comparative studies in HRM? A critical review of literature in the period of 2000–2014. Int J Hum Resour Manag 28(1):196–233
10. Abraham M, Niessen C, Schnabel C, Lorek K, Grimm V, Möslein K, Wrede M (2019) Electronic monitoring at work: the role of attitudes, functions, and perceived control for the acceptance of tracking technologies. Hum Resour Manag J 29(4):657–675
11. Van Eck NJ, Waltman L (2014) Visualizing bibliometric networks. Measur Scholarly Impact Methods Pract 147:285–320
12. Singh VK, Singh P, Karmakar M, Leta J, Mayr P (2021) The journal coverage of web of science, scopus and dimensions: a comparative analysis. Scientometrics 126:5113–5142
13. Fauzi MA (2022) Knowledge hiding behavior in higher education institutions: a scientometric analysis and systematic literature review approach. J Knowl Manag 27:302–327
14. Mingers J, Macri F, Petrovici D (2012) Using the h-index to measure the quality of journals in the field of business and management. Inform Process Manag 48(2):234–241
15. Huang MH, Rust RT (2018) Artificial intelligence in service. J Serv Res 21(2):155–172
16. Kheybari S, Rezaie FM, Farazmand H (2020) Analytic network process: an overview of applications. Appl Math Comput 367:124780
17. Caputo F, Cillo V, Candelo E, Liu Y (2019) Innovating through digital revolution: the role of soft skills and big data in increasing firm performance. Manag Decis 57(8):2032–2051
18. Lukovac V, Pamučar D, Popović M, Đorović B (2017) Portfolio model for analyzing human resources: an approach based on neuro-fuzzy modeling and the simulated annealing algorithm. Exp Syst Appl 90:318–331
19. Pillai R, Sivathanu B (2020) Adoption of artificial intelligence (AI) for talent acquisition in IT/ITeS organizations. Benchmark Int J 27(9):2599–2629
20. Pan Y, Froese F, Liu N, Hu Y, Ye M (2022) The adoption of artificial intelligence in employee recruitment: the influence of contextual factors. Int J Hum Resour Manag 33(6):1125–1147
21. Nankervis A, Connell J, Cameron R, Montague A, Prikshat V (2021) Are we there yet? Australian HR professionals and the fourth industrial revolution. Asia Pacif J Hum Resour 59(1):3–19

A Review on Verifiable Image Licensing Approaches

Priyanshi Thakkar and Nishant Doshi

Abstract Image licensing involves granting permission to use an image while retaining ownership, rather than selling the full rights to it. This permission allows the user to edit, sell, and advertise the usage of the image, which is the standard practice in the industry. By licensing an image, an agreement is made that outlines the terms, classifications, and limitations for its use. However, the authenticity of the image and any modifications made to it cannot be determined using techniques like process of acme and perception ciphering. While cryptographic information has been gathered from previous studies, they have identified two key characteristics of a verifiable image licensing system: (1) authorized use, which means only licensed individuals or entities who agree to the usage conditions can create legitimate images, and (2) productivity, which means that the validation of a genuine image licensing system is fast and not impacted by the number of edits or image dimensions.

Keywords Photograph usage rights · Valid table photograph handling · Secure cipher operations · Digital authentications · Two-way pairings

1 Introduction

Photograph, a representation of the outside form of someone or element in artwork. Every photograph has its personal copyrights that the writer or the individual that is publishing has given. If anyone who wants to edit or change the photo, through public legislation, authorship rights are granted and infringement of those rights can result in penalties.

P. Thakkar (✉) · N. Doshi
Pandit Deendayal Energy University, Gandhinagar, Gujarat, India
e-mail: priyanshithakkar666@gmail.com

N. Doshi
e-mail: nishant.doshi@sot.pdpu.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_5

The copyright holder of a photograph can establish rules for its use by entering into a usage contract with the individual or entity seeking to use the photograph. Additionally, to a payment, a standard licensing agreement includes three key elements: (1) the authenticity of the image, which includes identifying the authorship right owner, (2) the permitted user who has the right to operate the photograph, and (3) the permissions granted for the image's use, including any restrictions or conditions on modifying the photo.

It is difficult for even experts to detect if a photograph has been altered using Photoshop if it is done intentionally. To prevent copyright infringement, technical solutions that match the photo licensing agreement are needed. There are various methods of authenticating photographs, such as watermarking [1], perception ciphering [2], and protected primitives. A digital watermark that is based on copyright protection is resistant to changes and provides a way to determine the originality of the photograph. The encoded information can be uncovered, even if the photograph has undergone changes. For example, a professional photograph with a "no-copy-allowed" watermark can resist typical image manipulations like JPEG compression, spinning, trimming, and adding disturbance. Watermarking is the process of embedding a digital code and is a well-established technique for identifying the source of a photograph without limiting the rights of the legal user or licensee.

A novel method for licensing images is suggested, which is based on the blockchain technology [3]. The proposed system involves embedding a digital signature [4–6] within the image itself, ensuring its authenticity. Furthermore, the system includes a marketplace for the buying and selling of licensed images, which could potentially provide a more convenient and lucrative way for digital content creators to protect their intellectual property and receive fair compensation.

1.1 Organization of the Paper

In Sect. 2 of the paper, we have provided an overview of different techniques that can be used for a verifiable image licensing system. In Sect. 3, we have presented an analysis of these approaches. The conclusion and references can be found at the end of the paper.

2 Literature

Traditional methods for authenticating photos, such as watermarking and perceptual hashing, do not allow users to independently verify the validity of photos because these techniques require the help of the copyright holder.

VILS is different in that it allows for traceable image originality, assignable image editors, and controllable editing behavior. An owner can give a unique code that allows access to their image and only authorized individuals can provide valid

evidence of the image. Additionally, the image licensor can set rules for how much editing is allowed. One of the biggest advantages of VILS is that it allows for verifiable proof of the ownership, which has not been addressed by other photo authentication systems. Furthermore, it only allows certain key holders to edit the photo and takes into consideration the quality and size of the image.

Traditional methods of authenticating photographs include techniques like watermarking and perceptual hashing, which provide effective ways to verify the integrity of an image. Semi-fragile watermarks, which are embedded in pictures, can detect malicious attempts to alter the image but are still able to withstand benign changes that do not affect the actual content of the image. Researchers have proposed several promising approaches to this problem.

Traditionally, authenticating photographs has relied on techniques such as watermarking and perceptual ciphering, which are effective in verifying the originality of an image. These semi-fragile watermarks embedded in images can detect any malicious changes made to the image, while remaining resilient to non-harmful changes that do not affect the image's content [7–11]. Researchers have proposed various promising solutions to address this issue. Perceptual hash: In this method of authenticating photographs, the image data is converted into a short sequence, known as a picture cipher, to confirm the reliability of the content. A photo cipher is a compressed illustration of an image created using a specific algorithm. The advantage of using image ciphering is that the cipher can be sent separately from the image.

The purpose of perceptual ciphers [12] is to create a unique "fingerprint" of an image for use in photo hunting and verifying photo information. Additionally, the cipher functions used in perceptual ciphering are designed to be robust to certain types of image manipulation. However, the image cipher cannot identify who made changes to an image and does not provide a way to identify the editor of a photo, similar to watermarking (Table 1).

2.1 Cryptography-Based Image Authentication

The use of virtual autographs and absence of information proofs are increasing in picture verification. This method involves signing the picture information with a standard digital signature scheme (DSS) and sending the mark as a validation code to client along with the photo [13]. The unforgeability feature of DSS ensures that no one, including users with good intentions, can alter retaining the photos' validation code. Virtual marks can also be used to spot alteration as it is sensitive to changes. However, this approach for image identification may not be robust, but it provides provable security and has limitations in detecting integrity issues.

Table 1 Comparison between VILS and other channels [3]

Primary	Permission goal	Achievement				
		Protection	Strength	Permission	Capability	Load
SFW	Detail verification	Disprovable	Yes	NO	≥ O(N)	0
PHF	Detail verification	Disprovable	Yes	NO	≥ O(N)	O(1)
DSS	Ethics identification	Verifiable	No	NO	O(1)	O(1)
DSS + CH	Random cutting	Verifiable	Restricted	NO	O(N)	O(N)
DSS + CS	Acceptable cutting	Verifiable	Restricted	NO	O(N)	O(1)
DSS + ACC	Detail withdrawal	Verifiable	Restricted	NO	O(N)	O(1)
DSS + SNARKs	Confirmable photo modifying	Verifiable	Flexible	NO	≥ O(N^2Log(N))	O(N)
DSS + ACC	Confirmable photo modifying	Verifiable	Flexible	NO	O(m)	O(1)
DSS + BLP	Confirmable and acceptable photo modifying	Verifiable	Flexible	Yes	O(m)	O(1)

2.2 Admissible Image Processing

Digital images are made up of two primary elements: snapshot parameters and snapshot message. Image data is the numerical values of the image's pixels stored in a computer, which sets computer-generated images apart from printed ones [2]. The image content, inversely, is the subject or meaning of the snapshot, which largely impacts its value in image-related applications. These operations allow users to achieve desired image content by editing the image data on a computer correctly (Fig. 1).

In this context, pi represents the specific processing parameters (such as a quality factor of 60) and n is the number of operations in P. For an image M, each operation results in a rare modified image Mi, represented as Mi = fi (M, pi). Approved image adjustment is defined as a method that processes a snapshot with modifications from set P and outputs a set of photos M, denoted as M ← process (MP).

System description: A new image identification is developed by them keeping in mind following things: (1) Photo Copyright Holder (a licensor), (2) Legal Photograph Editor (a licensee), (3)Verifier.

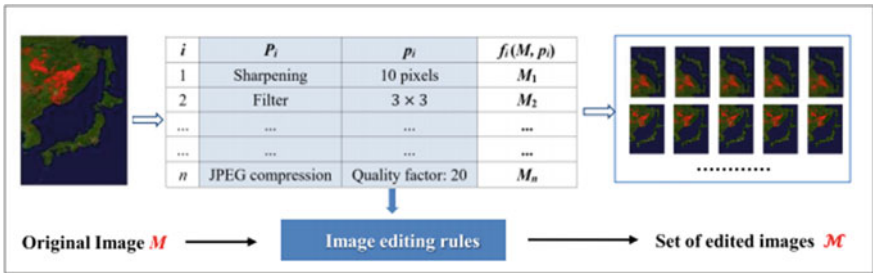


Fig. 1 Process algorithms [3]

Security: Safety is provided among competitor and an opponent [3]. The attempts to compromise the scheme by working with the competitor to reduce the security risk. The security is ensured if the chances of the opponent successfully winning the game are minimal. This method of assessment, which typically involves a security reduction, is commonly employed in encryption-based process.

Implementation and evaluation of their system: In the practical implementation of their system, Photoshop and similar software can be utilized to perform image manipulation procedures in an offline setting. As a result, the computational cost of the processing algorithm is not considered when evaluating this VILS. Consequently, the efficiency of the cryptographic techniques employed plays a significant role in determining the effectiveness of our VILS.

2.3 Implementation and Testing

To assess performance more specifically, they carried out all the methods using the Synchronized Encrypted data bank. To test the performance of the scheme more intuitively, they chose a set of images from the Kotak image set3 for testing. The Kotak image set comprises 28 lossless images, which they used to measure the effectiveness of system. VILS comprises four tedious procedures: Key-Gen, Validate, Edit-Prove, and Confirm [3]. In this test, the evaluating photos were converted into various forms from 200 to 2000 categories of processes for permission. In this process, the photos require to be pre-treated and ciphered to a set before authorization. The binary data of each image is processed and reduced to a compact representation. Therefore, ciphering a photo obtains fewer than 1.5 ms, and this method proves to be far more efficient in comparison to conventional image processing techniques. It also shows that the most remarkable achievement is the constant overhead of verification time. The verification time remains constant at 19 ms, regardless of the size of the image or the number of allowable operations. Cipher functions are utilized to transform photos of varying dimensions into fixed-size representations excluding duplication, and compression methods are employed to condense all cipher values into a concise and fixed-length validation code. As a result, the overhead for identification is consistent

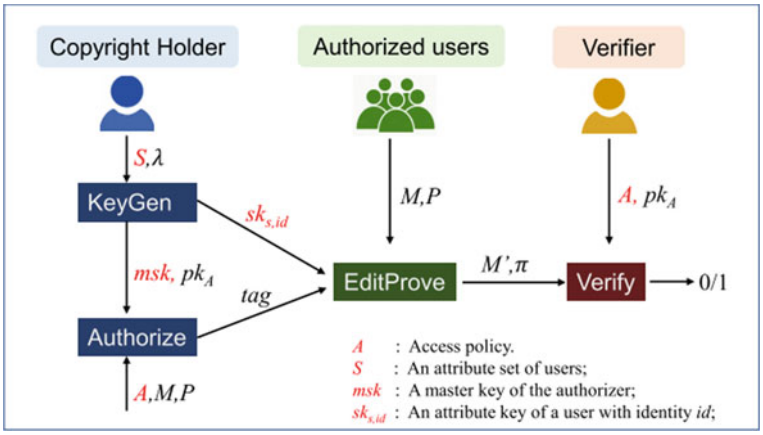


Fig. 2 Multi-user image authorization [3]

regardless of the photo dimension or count of processes. The VILS adeptness testifies to its capability in providing efficient image licensing verification (Fig. 2).

2.4 Scalability of Multi-person Authorization

The VILS suggests scientific assistance for agreements between a photograph author (rights holder) and a permitted operator (permitted party). While it is created to certify photo processes for a one person at a time, it can be adapted to support multiple users by implementing principles of ciphertext-policy attribute-based encryption (CP-ABE) [14–18]. In CP-ABE, decoding keys are issued based on user's attributes and the data proprietor specifies permission guidelines for the encryption text. By generating private keys of authorized users based on an attribute set S, the author (approver) can allow the photo(s) to many peoples with protocols A on S. Those whose characteristics meet A can modify the photo with their identifier keys. A photo receiver checks an image based on A. Other developed techniques like context-aware seal, multi-user seals, federation seal, secret splitting, etc. can also be used as alternatives for creating image validation method when collaborative permission is needed.

3 Analysis

From recent studies, we can say that only one person can use it at one time, and we need a system that provides multi-user technology to save time and space. We can provide license on the image, that would not allow any third person to use our image

Table 2 Comparison of some existing schemes

Snapshot dimension	Modification digits	Digit of entry	Time in seconds (s)			Memory (kB)	
			Origination	Validation	Affirmation	Size	Size of overhead
128*128	–	12,531,999	~ 367	~ 306	~ 0.5	2.6*10^6	2.67
Unrestricted	1000	–	~ 1	0.33	~ 0.025	< 13	< 2
Unrestricted	1000	–	~ 1	0.35	~ 0.018	< 15	< 2

which is called as trademark or say watermark. It also says if any person wants to crack the system or enters wrong password, the file will be automatically deleted. We also got information of how and what kind of attacks are seen on image (Table 2).

4 Conclusion

A verifiable image licensing system (VILS) has been created by them as an efficient solution for authenticating the originality of an image, the valid client, and the snapshot usage permission outlined in photo usage contracts. The system boasts advanced security features, such as tamper proofing and traceability, which greatly enhances its photo validation capabilities. Despite the increased number of elements to authenticate, the system's performance remains acceptable. The certifying process has a slight dip in performance, but the test process has seen a major upsurge of 40%. The team has come up with a design that enhances the VILS to include multi-user authorization utilizing the principles of ciphertext-policy attribute-based encryption (CP-ABE) [19]. They consider this approach as a possible solution for providing many people permission in cryptological-based photo identification plans.

Author Contributions We express our gratitude to the reviewers and organizers of the conference for their insightful feedback. Priyanshi is identified as the first author of the paper, and both Priyanshi and Dr. Nishant Doshi are the corresponding authors.

References

1. Van Schyndel RG, Tirkel AZ, Osborne CF (1994) A digital watermark. Proc Int Conf Image Process 2:86–90
2. Chen H, Wang S, Zhang H, Wu W (2018) Image authentication for permissible cropping. In: Proceedings of the 14th international conference on inscrypt, Fuzhou, China, pp 308–325
3. Chen H, Huang X, Ning J, Zhang F, Lin C (2022) VILS: a verifiable image licensing system. IEEE Trans Inform For Sec 17:1420–1434
4. Pointcheval D, Stern J (1996) Security proofs for signature schemes. In: Proceedings of the international conference on theory and applied cryptographic technology, vol 1070. Springer, Zaragoza, pp 387–398
5. Boneh D, Boyen X (2004) Short signatures without random oracles. In: Proceedings of the international conference on theory applied cryptographic technology, pp 56–73

6. Baric N, Pfitzmann B (1997) Collision-free accumulators and fail-stop signature schemes without trees. In: Proceedings of the international conference on theory applied cryptographic technology. Springer, Berlin, pp 480–494
7. Yao Y, Zhang W, Wang H, Zhou H, Yu N (2019) Content-adaptive reversible visible watermarking in encrypted images. Signal Process 164:386–401
8. Al-Otum HM (2014) Semi-fragile watermarking for grayscale image authentication and tamper detection based on an adjusted expanded-bit multiscale quantization-based technique. J Vis Commun Image Represent 25(5):1064–1081
9. Serra-Ruiz J, Qureshi A, Megías D (2019) Entropy-based semi-fragile watermarking of remote sensing images in the wavelet domain. Entropy 21(9):847
10. Peng F, Lin Z-X, Zhang X, Long M (2021) A semi-fragile reversible watermarking for authenticating 2D engineering graphics based on improved region nesting. IEEE Trans Circuits Syst Video Technol 31(1):411–424
11. Egorova A, Fedoseev V (2022) Semi-fragile watermarking for JPEG image licensing process: Google search. https://www.google.com/search?client=firefox-b-d&q=what+is+image+licensing
12. Venkatesan R, Koon SM, Jakubowski MH, Moulin P (2000) Robust image hashing. In: Proceedings of the international conference on image processing (ICIP), Vancouver, BC, Canada, pp 664–666
13. Shah SA, Khan IA, Kazmi SZH, Nasaruddin FHBM (2021) Semi-fragile watermarking scheme for relational database tamper detection. Malay J Comput Sci 34(1):1–12
14. Ning J, Cao Z, Dong X, Wei L, Lin X (2014) Large universe ciphertext-policy attribute-based encryption with white-box traceability. In: Proceedings of the European symposium on research computation security. Springer, Cham, pp 55–72
15. Goyal V, Pandey O, Sahai A, Waters B (2006) Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on computation communication security, pp 89–98
16. Maji HK, Prabhakaran M, Rosulek M (2011) Attribute-based signatures. In: Proceedings of the cryptographers track RSA Conference. Springer, Berlin, pp 376–392
17. Ning J, Cao Z, Dong X, Liang K, Ma H, Wei L (2018) Auditable σ-time outsourced attribute-based encryption for access control in cloud computing. IEEE Trans Inf For Sec 13(1):94–105
18. Ning J, Huang X, Susilo W, Liang K, Liu X, Zhang Y (2022) Dual access control for cloud-based data storage and sharing. IEEE Trans Depend Sec Comput 19(2):1036–1048. https://doi.org/10.1109/TDSC.2020.3011525
19. Chen H, Huang X, Wu W, Mu Y (2021) Privacy-aware image authentication from cryptographic primitives. Comput J 64(8):1178–1192

A Review on Multimodal Face Anti-Spoofing in Multi-layer Environment

Soham Patel and Nishant Doshi

Abstract With advancement of technology, the need of various threats also increases. One of the key threats is face spoofing, i.e. one can spoof the image of someone's face image and utilize for unethical purpose. In order to resolve this issue, researchers have suggested the face anti-spoofing techniques. Afterwards, many researchers have suggested and analyzed the techniques to improve upon the basic face anti-spoofing technique. These techniques utilize both local and long-range information. Recently, the researchers propose a new architecture called Conv-MLP that combines local features extracted by CNNs with global information from multi-layer perceptrons. Additionally, they suggest using a new loss function called ditch loss to improve the system's ability to generalize to new types of attacks. Thus, in this study we have reviews the available techniques under multi player environments for face anti-spoofing techniques and showcase the pros and cons of the same thing.

Keywords Face anti-spoofing · Multimodal · Inductive bias · Multimodal face anti-spoofing · Patch-based face anti-spoofing

1 Introduction

Anti-spoofing for face recognition systems is becoming more important in both academia and industry, but it can still be challenging to differentiate between genuine and counterfeit faces because of variety of spoofing methods, examples of attacks include print attacks, replay attacks, and masking attacks [1–11]. In the Fig. 1, there are some examples of true and fake faces to measure the effectiveness of our eyes in detecting attempts to deceive or "spoof" them. Out of the eight pictures shown, three are true and five are fake, and there are no clear and easy to discern indications to

S. Patel (✉) · N. Doshi
Pandit Deendayal Energy University, Gandhinagar, Gujarat, India
e-mail: patelsoham008@gmail.com

N. Doshi
e-mail: nishant.doshi@sot.pdpu.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_6

differentiate between genuine faces and counterfeit ones. Recently, researchers are devising more generalized and anti-spoofing face discriminant functions like LBP, HOG, LBP-TOP, DoG, etc. [12–19]. On average, these properties are known as handmade elements as they are created by individual. To the extent of our knowledge, this is a pioneering effort. In relation to manual characteristics, characteristics obtained from CNN are able to capture more discriminative stimuli based on the data way. It is more important as per the experiment results, capable of acquiring more general characteristics for various forms of deception.

The research introduces a new architecture called Conv-MLP [20] to prevent face fraud, which combines localized convolution and global multilayer perceptron. The approach efficiently utilizes both local data and distant dependencies to detect fake faces and increase the distinction between authentic and counterfeit faces. The paper [20] also proposes a new method called moat loss, which is specifically designed to prevent face fraud. This approach helps to create more distinct representations and improves the ability to detect new types of spoofing attacks. Conv-MLP was evaluated on six datasets that include both single and multiple modes, and it performed competitively compared with other cutting-edge algorithms in terms of precision and efficiency of computation.

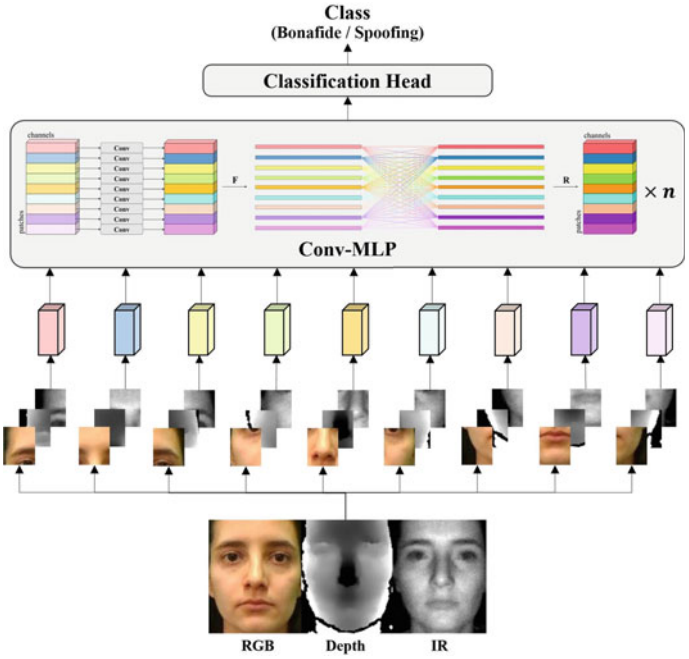


Fig. 1 Categorization head, cross-patch global blending via multilayer perceptron, and per-patch local convolution [20]

1.1 Organization of the Paper

Within the Sect. 2, we have given the various available techniques for the face anti-spoofing approaches. In Sect. 3, we have given the analysis based on the approaches. Conclusion and references are at the end.

2 Literature

The structure of the literature review is as follows: (1) Face anti-spoofing (2) Multi-modal face anti-spoofing (3) Patch-based face anti-spoofing (4) Conv-MLP (5) Patch generation and multimodal fusion (6) Moat loss.

2.1 Face Anti-Spoofing

The face recognition community is faced with the crucial yet difficult subject of face anti-spoofing. It has several real-world uses in face authentication, security verification, and access management. This task entails determining whether a face was photographed during spoof attacks [1–9, 21, 22] such as those involving printed faces, digital replays of face videos, masks, etc. Face anti-spoofing is therefore extremely essential to the safety of face recognition technology. Recently, jobs involving face anti-spoofing have seen widespread use of CNN. In certain, common models are adjusted for various tasks. To obtain the fine variations in facial movements, a VGG model [23] which is pre-trained is utilized. Additionally, AlexNet model [1, 2, 21, 22, 24–27] uses both losses due to cross-entropy and focalization. However, simple fine-tuning only results in a modest performance boost. The majority of methods discussed in the body of prior research make an effort to address the problem using visible spectrum datasets. As the methods used for spoofing change over time, it is becoming increasingly difficult to detect them using only visible light. To address this issue, researchers are now exploring the use of additional types of data, including details on depth and infrared, to enhance the effectiveness of face anti-spoofing methods.

2.2 Multi-modal Face Anti-Spoofing

Insufficient studies have been conducted on this subject so far. In their three-stream network proposal by Zhang et al. who use ResNet18 [20] as the framework, they use RGB, Depth, and IR face photos as the inputs for each stream. The final two residual blocks receive a concatenated version of these properties. Researchers also take

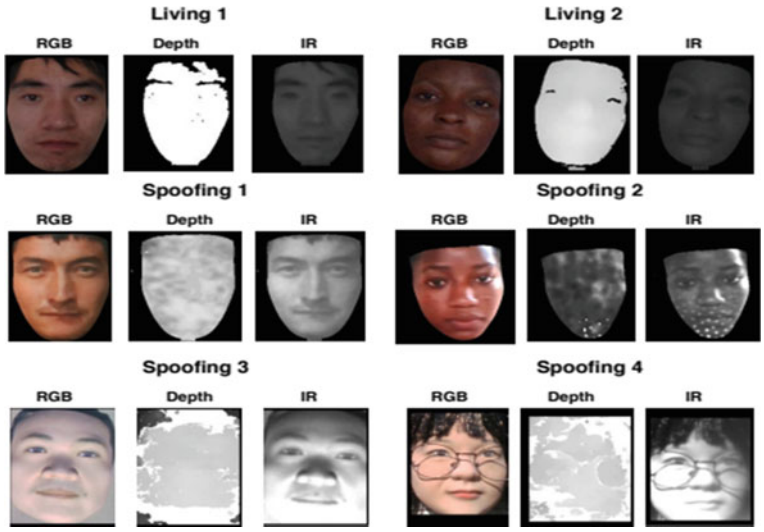


Fig. 2 Utilizing multi-modal techniques for distinguishing authentic and counterfeit face images [20]

into account the three-stream combination network. The backbone is chosen to be ResNet34, and all residual blocks are fused using multi-scale features. FaceBagNet is a multi-stream CNN architecture presented by Tao et al. prior techniques only took into account. The conventional structure (ResNet) with layered basic convolutions for various modes may not be sufficient to precisely distinguish the distinctions between authentic and counterfeit faces as shown in the Fig. 2 below.

2.3 Patch-Based Face Anti-spoofing

Previous studies have shown that information that encompasses both nearby and far-reaching aspects are important for detecting fake faces, according to Cai et al. Researchers such as Wang et al. [20] have used random cropping of face regions to extract local information, while Yang et al. developed a patch-based method that uses a custom architecture for each type of information from multiple modes to optimize utilization.

In this particular CNN [9], there are several reasons to utilize patches rather than a whole face. The first step is to enhance the training samples available for CNN learning. It should be noted that there are just a few samples available for training for each anti-spoofing dataset that is currently accessible. Due to different face picture resolutions, classic CNN must resize faces when full face images are utilized as input. This scale adjustment could result in the loss of discriminative information. Third, regardless of the patch position, inputting at the patch-level can encourage the CNN

to uncover spoof-specific discriminatory information if it is present in a specific area of the entire face region. Compared with utilizing the entire face image, this learning process is more constrained or challenging.

The CNN examines information affiliated to face spoofing from small patches, while the RNN [28] gradually learns to understand the overall information from the patches that have been studied. Due to the fact that RNN computations are fundamentally sequential, parallelization is also not viable.

2.4 Conv-MLP

Convolutional-multilayer perceptron (Conv-MLP) [20] can learn spatial feature hierarchies from input data in an automatic and adaptive manner. By combining convolutional and fully connected layers, the Conv-MLP architecture allows a model to learn spatial and temporal features, making it well-suited for tasks involving images and time-series data. Conv-MLP architectures have become a popular choice for image classification and other computer vision tasks due to their high accuracy and ability to capture both local and global features. Further research is needed to extend this architecture to other domains and tasks, and to improve its performance.

2.5 Patch Generation and Multimodal Fusion

This layer-wise fusion method [20] is independent of the specific network used for a given modality, preserving its structure while uncovering underlying correlations between the modalities. Additionally, by using this method in various layers [4–11, 21, 28, 29], the model can take advantage of correlations between different modalities using both low-level and high-level features.

2.6 Moat Loss

The moat loss is a method [20] created when it comes to detecting fake or manipulated facial images, with the goal of improving the ability to detect fake faces and increasing the generalization of the system to new types of spoofing attacks that it has not seen before. The commonly used method for training anti-spoofing models is to use the BCE loss. Recently, attention has been directed towards utilizing more intricate loss functions to enhance the similarities within the same sample class and/or amplify the differences between classes as shown in Fig. 3 below.

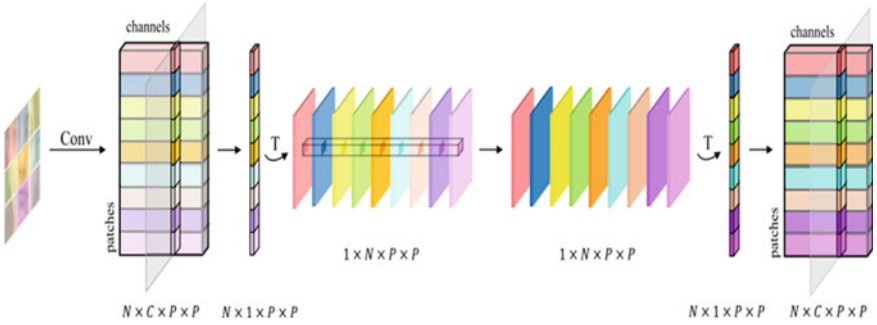


Fig. 3 CNN MLP combines N patches' characteristics [20]

3 Analysis

Face anti-spoofing is an important task in the field of machine learning, which aims to prevent impersonation attacks using fake facial biometrics. In recent years, various datasets have been created to support the research and development of face anti-spoofing algorithms. There are different types of face anti-spoofing datasets, including 2D and 3D face datasets, and datasets containing real and synthetic face data. Frequently employed datasets for face anti-spoofing include the Replay-Attack dataset [3], the Oulu-NPU [30] dataset, and the CASIA-MFSD [28] dataset. The Replay-Attack dataset contains 2D face data collected from real and synthetic (printed and video replay) attacks. The Oulu-NPU dataset includes 2D and 3D face data, with real and synthetic face data recorded under different lighting conditions. The CASIA-MFSD dataset contains 2D face data collected from real and synthetic face attacks, including print and video attacks. Every one of these datasets possesses distinct attributes and difficulties, which makes them useful for testing and evaluating different face anti-spoofing algorithms. For example, the Replay-Attack dataset is considered a benchmark dataset for 2D face anti-spoofing, while the Oulu-NPU dataset is useful for testing 3D face anti-spoofing algorithms. In general, utilization of these datasets has aided in creating stronger and more efficient solutions in face anti-spoofing algorithms. The results obtained from these datasets have shown that deep learning algorithms, such as ConvNets and RNNs are highly effective in detecting facial spoof attacks. As indicated in the Table 1, the analysis of different datasets on different models revealed significant variations in the accuracy and performance of each model.

Table 1 Analysis of different datasets on different models

Parameters	APCER (%)	BPCER (%)	ACER (%)
Precision of the CEFA dataset	1.33 ± 0.30	1.42 ± 0.26	1.37 ± 0.27
Precision of the CASIA-SURF dataset	1.5	1.8	1.6
Precision of the most extensive protocol of the WMCA dataset	0.8	1.0	0.9
Precision of the most extensive protocol of the WMCA dataset	1.62	1.29	1.45
Precision of the SIW dataset	0.28	1.06	0.67

4 Conclusion

In conclusion, the use of different face anti-spoofing datasets is critical for the development and evaluation of face anti-spoofing algorithms. By offering a means to evaluate and contrast the effectiveness of various algorithms, these datasets enable the ongoing improvement of accuracy and resilience in face anti-spoofing systems.

Author Contributions We are thankful to anonymous reviewers and conference organizers for thoughtful suggestions. Soham is first author in the paper. Soham and Dr. Nishant are the corresponding authors in the paper.

References

1. Yang Q, Zhu X, Fwu JK, Ye Y, You G, Zhu Y (2020) PipeNet: selective modal pipeline of fusion network for multi-modal face anti-spoofing. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 2739–2747
2. Shen T, Huang Y, Tong Z (2019) FaceBagNet: bag-of-local-features model for multi-modal face anti-spoofing. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 1611–1616
3. George A, Mostaani Z, Geissenbuhler D, Nikisins O, Anjos A, Marcel S (2020) Biometric face presentation attack detection with multichannel convolutional neural network. IEEE Trans Inf For Sec 15:42–55
4. George A, Marcel S (2021) Cross modal focal loss for RGBD face anti-spoofing. In: Proceedings of the IEEE computation on society of conference on computation visual pattern recognition (CVPR), pp 7878–7887
5. Cai R, Li H, Wang S, Chen C, Kot AC (2021) DRL-FAS: a novel framework based on deep reinforcement learning for face anti-spoofing. IEEE Trans Inf For Sec 16:937–951
6. Atoum Y, Liu Y, Jourabloo A, Liu X (2017) Face anti-spoofing using patch and depth-based CNNs. In: Proceedings of the IEEE international joint conference on biometrics (IJCB), pp 319–328
7. Battaglia PW et al (2018) Relational inductive biases, deep learning, and graph networks. arXiv:1806.01261
8. Wang X, Girshick R, Gupta A, He K (2018) Non-local neural networks. In: Proceedings of the IEEE/CVF conference on computation on visual pattern recognition, pp 7794–7803
9. Dosovitskiy A et al (2021) An image is worth 16×16 words: transformers for image recognition at scale. In: Proceedings of the international conference on learning representation (ICLR), pp 1–22

10. George A, Marcel A (2021) On the effectiveness of vision transformers for zero-shot face anti-spoofing. In: Proceedings of the IEEE international joint conference on biometrics (IJCB), pp 1–8
11. Liu A et al (2021) Face anti-spoofing via adversarial cross-modality translation. IEEE Trans Inf For Sec 16:2759–2772
12. Chingovska I, Anjos A, Marcel S (2012) On the effectiveness of local binary patterns in face anti-spoofing. In: Proceedings of the international conference on biometrics special interest group (BIOSIG), pp 1–7
13. Boulkenafet Z, Komulainen J, Hadid A (2015) Face anti-spoofing based on color texture analysis. In: Proceedings of the IEEE international conference on image processing (ICIP), pp 2636–2640
14. Ojala T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
15. Patel K, Han H, Jain A (2016) Secure face unlock: spoof detection on smartphones. IEEE Trans Inf For Sec 11(10):2268–2283
16. Boulkenafet Z, Komulainen J, Hadid A (2017) Face antispoofing using speeded-up robust features and Fisher vector encoding. IEEE Signal Process Lett 24(2):141–145
17. Komulainen J, Hadid A, Pietikäinen M (2013) Context based face anti-spoofing. In: Proceedings of the IEEE international conference on biomedical theory in applied system (BTAS), pp 1–8
18. Peixoto B, Michelassi C, Rocha A (2011) Face liveness detection under bad illumination conditions. In: Proceedings of the IEEE international conference on image processing (ICIP), pp 3557–3560
19. Liu Y, Jourabloo A, Liu X (2018) Learning deep models for face anti-spoofing: Binary or auxiliary supervision. In: Proceedings of the IEEE computing society of conference computation visual pattern recognition (CVPR), pp 389–398
20. Wang W, Wen F, Zheng H, Ying R, Liu P (2022) Conv-MLP: a convolution and MLP mixed model for multimodal face anti-spoofing. IEEE Trans Inform For Sec 17:2284–2297. https://doi.org/10.1109/TIFS.2022.3183398
21. Zhang S et al (2020) CASIA-SURF: a large-scale multi-modal benchmark for face anti-spoofing. IEEE Trans Biometr Behav Identity Sci 2(2):182–193
22. Liu A, Tan Z, Wan J, Escalera S, Guo G, Li SZ (2021) CASIA-SURF CeFA: a benchmark for multi-modal cross-ethnicity face anti-spoofing. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, pp 1178–1186
23. Yu Z, Li X, Niu X, Shi J, Zhao G (2020) Face anti-spoofing with human material perception. In: Vedaldi A, Bischof H, Brox T, Frahm J (eds) Proceedings of the European conference on computer visual (ECCV), in lecture notes in computer science, vol 12352. Springer, Cham, pp 557–575
24. Yu Z et al (2020) Searching central difference convolutional networks for face anti-spoofing. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5294–5304
25. Wang Z et al (2020) Deep spatial gradient and temporal depth learning for face anti-spoofing. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5041–5050
26. Yang X et al (2019) Face anti-spoofing: model matters, so does data. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 3502–3511
27. George A, Marcel S (2021) Learning one class representations for face presentation attack detection using multi-channel convolutional neural networks. IEEE Trans Inf For Sec 16:361–375
28. Schroff F, Kalenichenko D, Philbin J (2015) Facenet: a unified embedding for face recognition and clustering. In: Proceedings of the IEEE computation society of conference on computer visual pattern recognition (CVPR), pp 815–823
29. Yang X et al (2019) Face anti-spoofing: model matters, so does data. In: Proceedings of the IEEE computation society of conference on computer visual pattern recognition (CVPR), pp 3502–3511

30. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-CAM: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision, pp 618–626
31. Shao R, Lan X, Li J, Yuen PC (2019) Multi-adversarial discriminative deep domain generalization for face presentation attack detection. In: Proceedings of the IEEE computing society of the conference on computation visual pattern recognition (CVPR), pp 10015–10023

A Review on Quantum Blockchain Based Consensus Mechanism Approaches

Jeet Ramoliya and Nishant Doshi

Abstract Blockchain is dependent on consensus mechanism, hash rate, digital signature, and quantum protocols or algorithms. Indeed, quantum computers are powerful to break some consensus mechanism blockchains hash rate. Some quantum blockchains are proposed but not efficient for applications because of lackness due to quantum blockchain consensus mechanism. To eliminate the threat of quantum computer for future blockchain consensus mechanism, in research many quantum algorithms have been proposed and analyzed. Quantum entangled is used to form the chain to block using single quantum state. Quantum-based consensus mechanism is more secure and efficient than normal blockchain consensus mechanism for future applications. Thus, in this paper we have to survey the recent approaches based on the blockchain-based consensus algorithm in quantum world.

Keywords Blockchain · Quantum blockchain · Quantum cryptography · Consensus mechanisms · Quantum entanglement

1 Introduction

We can see in past years, distributed ledger technology has gained significant attention as a means of storing and validating data in a decentralized manner. One of the most promising distributed ledger technology (DLT) platforms is the blockchain. It utilizes a distributed database and cryptographic approaches to ensure the ethicality and security of data. However, traditional blockchains face several challenges, including low scalability, slow transaction speeds, and high energy consumption. Blockchain is a distributed decentralized immutable database. In peer-to-peer blockchain network, there is no central authority controlling the network, and each node has equal power

J. Ramoliya (✉) · N. Doshi
Pandit Deendayal Energy University, Gandhinagar, Gujarat, India
e-mail: jeetramoliya10@gmail.com

N. Doshi
e-mail: nishant.doshi@sot.pdpu.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_7

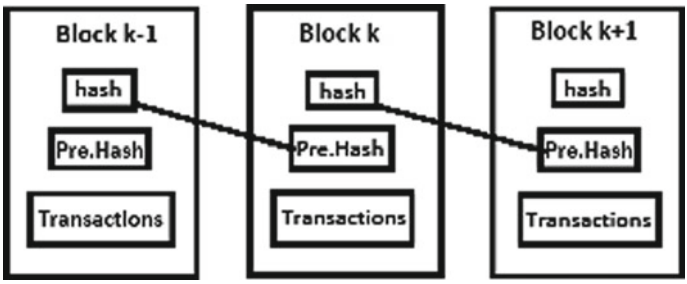


Fig. 1 Blockchain structure [3]

and copy of the entire blockchain, so they can validate and add new transactions to the chain. This allows for a more decentralized and democratic system, as no single entity has control over the network or the assets stored on it. Every blockchain concepts are combined of cryptography and consensus mechanism. Blockchains are used in various fields like e-voting, logistics, smart health, economics, [1] and personal data management [2]. There is multiple consensus mechanism for different blockchain requirements. Blockchain users securely interact by using asymmetric cryptography. Hash functions are used for generating digital signature and linking the blocks of blockchain. Main problem is that basic consensus mechanism and algorithms of a normal blockchain are vulnerable because of the power of quantum computers (Fig. 1).

Quantum delegated proof of stake (QD-PoS) [3] is a new type of distributed ledger technology that integrate the power of quantum computing with a unique consensus algorithm. It aims to provide faster, more secure, and more scalable solutions for various industries including finance, supply chain, and healthcare. At the core of quantum blockchain is a quantum-resistant proof of stake (QR-PoS) [3] algorithm, which allows network participants to validate transactions and reach consensus in a decentralized manner. QR-PoS uses complex mathematical computations and quantum-resistant cryptographic techniques to ensure the integrity and security of its network. The best advantages of quantum blockchain are its expertise to handle big transaction speeds and large amounts of data, with help of computational power of quantum computers. This makes it ideal for use cases that require real-time processing, such as high-frequency trading and online marketplaces. Additionally, quantum blockchain is designed to be energy-efficient, using a novel algorithm to reduce the energy consumption of the network. This makes it more environmentally friendly and cost-effective compared with other blockchain technologies. Overall, quantum blockchain has the potential to revolutionize various industries by providing a secure, fast, and scalable solution for storing and processing data.

Blockchain consensus mechanism like proof of work (PoW) and proof of stake (PoS) is dependent on hash rate. Shor's quantum algorithm easily attacks the blockchain [4]. Delegated proof of stake (DPoS) [4] consensus mechanism does not depend upon hash rate. That is why in quantum blockchain consensus mechanisms are proposed for future security.

1.1 Organization of Paper

In Sect. 2, we have given the literature survey of the existing approaches in the quantum-based blockchain approaches. In Sect. 3, we have given analysis of the approaches as mentioned in the Sect. 2. Conclusion and references are at the end.

2 Literature

The subsequent are framework of the literature review: (1) Brief info about blockchain and their application. (2) Brief info about quantum blockchain.

2.1 Brief Info about Blockchain

2.1.1 Normal Blockchain

Blockchain is normally a distributed decentralized immutable data source. Blockchain is proposed by Satoshi Nakamoto in 2008 [5] as peer-to-peer secure communication. Blockchain contains the consensus mechanism for validating the blocks without using third parties. Bitcoin is the first blockchain technology that used proof of work consensus mechanism. After bitcoin, Ethereum proposed by vitalik buterin in 2013–2014 for development purpose and that use proof of work consensus mechanism in past but now they shift on proof of stake consensus mechanism on 15 September 2022 for some benefits purpose. Ethereum is working with smart contract to develop the decentralized application. That project is known as Web3.0. Blockchain has some concepts for future work like DAO, DeFi, and DAPP. Every consensus mechanism has its own qualities.

Mainly, blockchain relies on consensus mechanism, hash function, and digital signature. Digital signature is used a public key cryptography. That is why blockchain users communicate securely. Hash function is also important in blockchain and mainly SHA256 is used as hash function.

2.1.2 Blockchain Application

Blockchain can be used to create a tamper-proof record of the movement of goods from the manufacturer to the consumer. Each block in the chain can contain information such as the origin of the goods, the parties involved in the transaction, and the location of the goods at a specific point in time. This can help increase transparency and trust in the supply chain, as well as improve efficiency by reducing the need for intermediaries.

Smart contracts, a key feature of blockchain technology, is another classical application, it allows for the automation of complex processes and eliminates the need for intermediaries, which can save time and costs. And it is widely used in many industries such as: insurance, real estate, finance, etc. As blockchains use peer-to-peer transaction system, they do not need any third parties for verification. That is why blockchain is used in various fields.

2.2 Brief Info about Quantum Blockchain

2.2.1 Quantum Blockchain

Quantum blockchain is a combination of quantum computing concepts and blockchain concepts. Quantum computer has good computation. Quantum mechanics and quantum algorithm easily break the normal blockchain concepts. Quantum blockchain is secure for future quantum attacks or different attacks.

In a traditional DPoS system, token holders vote to elect a group of validators, called witnesses, who are liable for maintaining an integrity of blockchain by adding new blocks to the chain. In a quantum DPoS system, the addition of quantum computing is proposed to be used to increase the security of this process (Fig. 2).

Quantum blockchain refers to a concept of using quantum computing to improve a security and functionality of blockchain technology. One suggested method for achieving this is by using quantum entanglement [3], a phenomenon in which two quantum particles can become linked such that the state of one particle is dependent on the state of the other, even when separated by large distances. Entanglement could be used to create a shared, secure key that is used to encrypt and decrypt transactions. The idea is that, because the key is based on the state of entangled particles, it would be impossible for an attacker to intercept and decode a key without being detected,

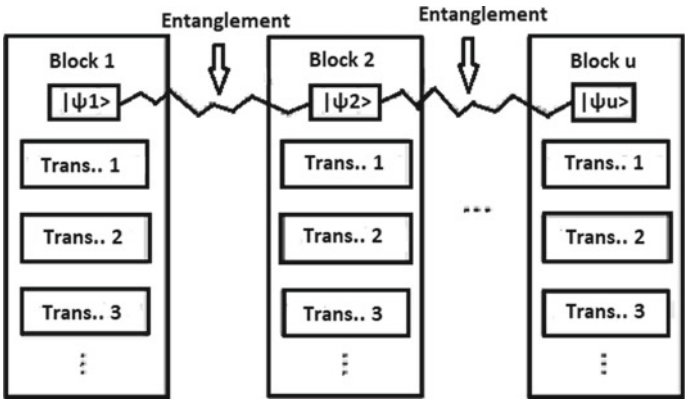


Fig. 2 Quantum blocks link via entanglement operations [3]

providing a high level of security. QD-PoS mechanism use quantum entangled to chain blocks. Single qubit tack data and that data is encrypted by QKD and stores in block using quantum entangled. This consensus mechanism gave efficient quantum computation than proposed quantum blockchain concepts.

In quantum blockchain whenever the transaction initiate that broadcast as peer-to-peer on network. After that transaction verified using quantum digital signature. Verified transaction using quantum delegated proof of stake (QD-PoS) consensus mechanism generates a quantum block by representative node and that reveals information on quantum blocks. That is how information block adds into a quantum blockchain.

In past, quantum secure blockchain [6] and quantum blockchain using entanglement in time [7] scheme are proposed. Both proposed schemes use two-way communication for transaction and that perform in $O(n^2)$ time. [6] scheme uses Byzantine agreement consensus mechanism and [7] scheme uses theta-protocol consensus mechanism. [6] consensus mechanism has $O(n^{(f+1)})$ complexity and [7] consensus mechanism has $O(n^2)$ complexity. [6] scheme needs classical bits for generating a block and [7] scheme needs a one two-qubit bell state for generating a block. The structure of the chain of [7] scheme is classical chain and [8] scheme is GHZ state. In [46], the authors have recently proposed the scheme for one-way communication that is used for transaction and that performs in $O(n^2)$ time. One single qubit state need for generating a block that scheme is resisting on hash rate attacks. QDPoS has a $O(n)$ time complexity. In that consensus mechanism quantum resources are not lost between processes.

3 Analysis

Quantum blockchain is more secure than normal blockchain when compared with cryptography and digital signature. In Table 1, we have given the summary of the normal blockchain against quantum blockchain based on the various parameters. Based on our analysis from this table, we can say that the normal blockchain is safe due to compromise of hash rate, whereas quantum blockchain is not vulnerable against same attack. In addition, quantum blockchain requires higher computation power as to normal blockchain-based approaches.

4 Conclusion

As we can see comparison among quantum blockchain and normal blockchain consensus mechanisms, where some quantum blockchain consensus mechanism generates a quantum block with single quantum state and chains are formed with quantum entangled. In future, quantum cryptography or quantum digital signature

Table 1 Comparison between blockchain and quantum blockchain

Parameters	Normal blockchain	Quantum blockchain
Consensus mechanism	1. Proof of work 2. Proof of stake 3. Delegated proof of stake 4. Proof of history 5. Proof of burn 6. Proof of capacity	1. Byzantine agreement 2. Theta-protocol 3. Relative phase
Cryptography algorithms	1. Symmetric-key cryptography 2. Asymmetric-key cryptography	1. Quantum-key distribution 2. Mistrustful-quantum cryptography 3. Bounded and noisy-quantum-storage model
Hashing function based on their type	1. 256SHA 2. 512SHA	1. Use quantum entangled instead of hash function
Break blockchain by any algorithms	1. Yes (Shor's algorithm and other quantum algorithms)	1. No
Structure	1. Merkle tree	1. Weighted graph state 2. Weighted hypergraph state 3. GHZ state

scheme will be replaced with other schemes for making this quantum blockchain consensus mechanism further secure.

Authors Contributions We express our gratitude to the anonymous reviewers and conference organizers for their valuable suggestions. The first author of this paper is Jeet, and the corresponding authors are Jeet and Dr. Nishant.

References

1. Catalini C, Gans JS (2020) Some simple economics of the blockchain. Commun ACM 63(7):80–90
2. Truong NB, Sun K, Lee GM, Guo Y (2020) GDPR-compliant personal data management: a blockchain-based solution. IEEE Trans Inform For Sec 15:1746–1761
3. Li Q, Wu J, Quan J, Shi J, Zhang S (2022) Efficient quantum blockchain with a consensus mechanism QDPoS. IEEE Trans Inform For Sec 17:3264–3276. https://doi.org/10.1109/TIFS.2022.3203316
4. Shor PW (1994) Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings of the 35th annual symposium on foundations of computer science, pp 124–134
5. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. Decentr Bus Rev 2:21260
6. Kiktenko EO, Pozhar NO, Anufriev MN, Trushechkin AS, Yunusov RR, Kurochkin YV, Lvovsky A, Fedorov AK (2018) Quantum secured blockchain. Quant Sci Technol 3(3):035004
7. Rajan D, Visser M (2019) Quantum blockchain using entanglement in time. Quant Rep 1(1):3–11
8. Banerjee S, Mukherjee A, Panigrahi PK (2020) Quantum blockchain using weighted hypergraph states. Phys Rev Res 2(1):013322

A Review on Poligraph: System for False News Detection

Darp Desai and Nishant Doshi

Abstract Poligraph is a solution to the issue of identifying false news and information that takes into account the various challenges posed by this issue. It aims to build a practical and long-lasting platform that can effectively detect fake news by utilizing both machine learning techniques and human expert assessments. The quick but less accurate results from machine learning-based reviews are balanced with the more accurate but slower results from human reviews, particularly expert reviews. This balance is made possible by the general agreement built into Poligraph that allows for the seamless integration of these two approaches. The performance of Poligraph was assessed through evaluations utilizing reports from online media and publications on the Amazon Elastic Compute Cloud.

Keywords Reliability · Fault tolerance · Machine learning (ML) · Deep learning (DL) · Human review (HR)

1 Introduction

Polygraph is a term commonly used to describe a lie detector test, which is used to assess whether an individual is telling the truth or not. In the context of human examination, a polygraph machine is used to measure physiological responses such as blood pressure, pulse, respiration, and skin conductivity while a person is being questioned [1]. This is used in physical world. But Poligraph is used in virtual world.

The significance of language and the meanings it conveys in influencing public opinion is immense, especially in the realm of news media and politics. Fabricated information, known as fake news, can be easily mistaken for legitimate news due to its similar form, but lacks the editorial standards, accuracy, and credibility of

D. Desai (✉) · N. Doshi
Pandit Deendayal Energy University, Gandhinagar, Gujarat, India
e-mail: darpdesai@yahoo.com

N. Doshi
e-mail: nishant.doshi@sot.pdpu.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_8

actual news. The distribution of false messages can have serious repercussion for both individuals and companies as it frequently misleads the public by presenting information out of context. Unfortunately, people's ability to distinguish truth from deception is only marginally better than random, with an average accuracy of 54% [2].

The spread of misleading information regarding COVID-19 on social network is becoming increasingly concerning. In light of this, the ability to recognize fake news is more critical than ever [2].

To determine the accuracy of detecting false news, researchers have approached the subject in various ways. Categorization can be observed in four distinct areas: experimental studies, modeling fake news, and the implementation of both ML and DL techniques.

Participants are involved by investigators in experimental studies to evaluate the validity of news and how their opinions and beliefs are influenced by false information.

The identification of fabricated news requires the creation of models that document and analyze the evolution and dissemination of false articles. Various message features have been utilized by researchers and have trained numerous machine learning models using standard techniques such as regression and classification to determine the likelihood of articles being fake. Furthermore, in recent times, there has been an exploration of utilizing deep learning methodologies for spotting false news, complementing conventional machine learning approaches [3, 4].

Facebook is exploring the use of crypto currency as a means to combat fake news and bots [5]. While there have been numerous advancements in fake news detection methods and some single-provider services, such as Amazon Mechanical Turk [6], exist, there is currently no publicly available system that is tamper-proof and can guarantee reliability, security, and accuracy in detecting fake news. This underscores the importance of decentralization and protection against both Byzantine failures and malicious attacks [3, 7–20].

The identified challenges and requirements of the construction tamper resistant fake news detection system are as follow:

- **Requirement for ML and HR Integration**: Compared to DL, ML boasts a lower latency, leading to a more efficient system with higher processing speed.
- **Reproducibility**: It's crucial to make sure that the prediction of a message can be duplicated by anyone and remains consistent over time. However, the utilization of exclusive datasets by many current ML systems renders the prediction not capable of being duplicated.
- **Detection of False Information Online Versus Offline**: For various applications, the ability to deliver immediate/real time online predictions is important, though the accuracy of the outcomes may be diminished or compromised. Conversely, offline systems have the advantage of producing more accurate results.

1.1 Organization of the Paper

In Sect. 2, we have given the literature survey of the existing approaches in the Poligraph: system for false news detection. In Sect. 3, we have given analysis. Conclusion and references are at the end.

2 Literature

The subsequent are framework of the Literature review. 2.1 Detection of False News. 2.2 Byzantine Fault Tolerance. 2.3 Blockchains.

2.1 Detection of False News

The investigation into the elimination of false news has been a focus of research over the past two decades. This research can be categorized into four areas: experimentation studies, modeling of fake news, and the use of machine learning and deep learning techniques for detection [21–23].

One way to investigate the cognitive processes behind fake news detection are through behavioral experiments. Pennycook et al. conducted an experiment using Amazon Mechanical Turk (AMT) to determine how prior exposure to fake news affects a person's perception of its accuracy [24–26].

The models used in detecting fake news are built based on the unique characteristics of messages, forming the basis of false news modeling approaches.

The aim of using machine learning in detecting false news is to increase accuracy. To achieve this, different machine learning models are utilized that build on previous research and incorporate new, innovative features to improve the ability to identify fake news [3, 7–12, 27, 28].

Recently, the use of deep learning techniques in false message detection has also been explored. In this pursuit, a variety of deep learning models such as convolutional neural networks (CNNs) [2], long short-term memory (LSTM) networks, bidirectional LSTMs, adversarial neural networks, and even transfer learning methods have been employed to detect false data [13, 21, 22, 29].

2.2 Byzantine Fault Tolerance

The objective of BFT general agreement is for the right copies to agree on the arrangement of client requests. It enables a decentralized network to reach consensus

even when a small number of nodes are unresponsive or provide incorrect data [30–32].

2.3 Blockchains

Blockchains can be divided into two categories: permission-less and permissioned. Permission-less blockchain allows for dynamic membership through the use of mechanisms that guard against Sybil attacks, whereas in a permissioned blockchain, servers must identify themselves but not necessarily trust each other [2, 5, 33].

3 Analysis

Figure 1 shows: An evaluation of fake news detection systems reveals that Poligraph outperforms its competitors as it has support, unlike others.

Figure 2 shows: Human evaluations tend to produce higher results for all measurement criteria, but with a drawback of extended review time. An analysis of Poligraph was carried out on Amazon Elastic Compute Cloud, utilizing a maximum of 35 general-purpose nodes, each equipped with 2 virtual CPUs and 4 GB of memory. The system achieved a latency of only 0.05 s and a throughput of over 5000 transactions per second, which is only slightly (4–7%) slower compared to a non-replicated version on a single server.

Poligraph demonstrates stable peak performance as the number of servers (n) increases, ranging from 4 to 10. The decrease in performance from one experiment to the next, as n increases, is minimal, hovering around 4 to 5% (Fig. 3).

Advantage and disadvantage of implementing polygraph (Table 1):

A. Limitation of implementing Poligraph in practice

   a. Implementing Poligraph could pose limitations such as financial challenge due to the significant costs required for development, maintenance, and training.

fake news detection system	decentralized	Byzantine failure	Human review	ML	Reproducibility	growing data set
ML based online detection[30,39]	○	○	○	●	○	○
ML detection models[11,12,43]	○	○	○	●	○	○
Human determination[33,34]	○	○	●	○	○	○
Amazon Mechanical Turk[42]	○	○	●	○	○	○
Poligraph[40]	●	●	●	●	●	●

Fig. 1 Comparison of systems for detecting false news [2]

Approach	Accuracy	Precision	Recall	F1-score	Latency
Non-expert review (Group 1)	0.625	0.663	0.529	0.551	14 min
Expert review (Group 2)	1	1	1	1	3.3 min
Human review (Combined)	0.652	0.709	0.58	0.601	7 min
ML review	0.637	0.688	0.784	0.732	6.5s
T-stat	2.756**	3.534***	-4.891***	-0.931	NA

Fig. 2 Evaluating outcome of HR and ML reviews [2]

n	Peak Throughput (tx/s)	Degradation(%)
4	4358.75	-
7	4171.5	(4.29%)
10	3955.41	(5.18%)

Fig. 3 Scalability of poligraph [2]

Table 1 Advantage and disadvantage

Advantage	Disadvantage
Utilize ML and expert review	Require financial resources
Defense against tamper, attack, and failures	Technical complexity
Tackles different detection issues	Language limitations
Provides self-defense mechanism	Detect only certain type of false information
Importance of decentralized and tamper-proof systems	Privacy and ethical issues

b. Additionally, the platform's technical complexity may limit its accessibility to some users, particularly those without the necessary technical skills.
c. Furthermore, the system may be ineffective in detecting false information in certain languages, thereby reducing its usefulness in some regions.
d. Moreover, the Poligraph system may only be capable of identifying specific types of false information, and more sophisticated forms of false information may go undetected.

B. Challenges of Implementing Poligraph in Practice:

a. Implementing Poligraph could pose challenges such as privacy concerns due to the personal data involved in the Poligraph system's use of machine learning algorithms and human assessments.
b. Additionally, the use of human assessments may be subject to bias, and machine learning algorithms may produce false positives resulting in rise of ethical concerns.

c. Furthermore, to keep up with evolving tactics used by bad actors in spreading disinformation, the Poligraph system may need to be adjusted over time.
d. Moreover, the effectiveness of the Poligraph system may depend on user engagement and adoption, which can be challenging to achieve.

4 Conclusion

The Poligraph system utilizes a combination of ML and expert insights to detect false information with high accuracy. It is built with safeguards against tampering, malicious attacks, and Byzantine failures, and provides a self-defense mechanism. The two-layer consensus approach employed by Poligraph enables it to tackle diverse challenges in detection of false information, including offline compromises, secure parallel evaluations, and efficient transfer of information. The paper highlights the importance of decentralized and tamper-proof systems and acknowledges the challenges of constructing such systems. The future prospects of Poligraph include exploring further integration of machine learning and human evaluations and expanding its application to more languages and social media platforms. It represents a significant improvement in fake news detection methods and holds great potential for future development.

Author Contributions We are thankful to anonymous reviewers and conference organizers for thoughtful suggestions. Darp Desai is first author in the paper. Darp Desai and Nishant Doshi are the corresponding authors in the paper.

References

1. Polygraph, Wikipedia, 21 Nov 2022 [Online]. Available https://en.wikipedia.org/wiki/Polygraph. Accessed 22 Nov 2022
2. Shan G, Zhao B, Clavin JR, Zhang H, Duan S (2022) Poligraph: intrusion-tolerant and distributed fake news detection system. IEEE Trans Inf Forensics Secur 17:28–41. https://doi.org/10.1109/TIFS.2021.3131026
3. Mokhtar MS, Jusoh YY, Admodisastro N, Pa N, Amruddin AY (2019) Fakebuster: fake news detection system using logistic regression technique in machine learning. Int J Eng Adv Technol 9(1):2407–2410
4. Duan S, Zhang H (2016) Practical state machine replication with confidentiality. In: Proceedings of IEEE 35th symposium on reliable distributed systems (SRDS), pp 187–196
5. Mearian L (2019) Facebook's blockchain cryptocurrency could mean big money and kill fake news. Facebook, CA, USA, Tech Rep
6. Mason W, Suri S (2012) Conducting behavioral research on Amazon's mechanical Turk. Behav Res Methods 44(1):1–23
7. Nasir JA, Khan OS, Varlamis I (2021) Fake news detection: a hybrid CNN-RNN based deep learning approach. Int J Inf Manage Data Insights 1(1), Art. no. 100007
8. Aslam N, Khan IU, Alotaibi FS, Aldaej LA, Aldubaikil AK (2021) Fake detect: a deep learning ensemble model for fake news detection. Complexity 2021:1–8

9. Vedova MLD, Tacchini E, Moret S, Ballarin G, Pierro M, de Alfaro L (2018) Automatic online fake news detection combining content and social signals. In: Proceedings of 22nd conference of open innovations association (FRUCT), pp 272–279.
10. Wang Y et al (2018) EANN: Event adversarial neural networks for multi-modal fake news detection. In: Proceedings of SIGKDD, pp 849–857
11. Khattar D, Goud JS, Gupta M, Varma V (2019) MVAE: multimodal variational autoencoder for fake news detection. In: Proceedings of WWW, pp 2915–2921
12. Singhal S, Shah RR, Chakraborty T, Kumaraguru P, Satoh S (2019) SpotFake: a multi-modal framework for fake news detection. In: Proceedings of BigMM, pp 39–47
13. Pérez-Rosas V, Kleinberg B, Lefevre A, Mihalcea R (2017) Automatic detection of fake news. arXiv:1708.07104
14. Rashkin H, Choi E, Jang JY, Volkova S, Choi Y (2017) Truth of varying shades: analyzing language in fake news and political factchecking. In: Proceedings of EMNLP, pp 2931–2937
15. Lazer DMJ et al (2018) The science of fake news. Science 359(6380):1094–1096
16. Pennycook G, Cannon T, Rand DG (2018) Prior exposure increases perceived accuracy of fake news. J Experim Psychol Gen 147(2):1865
17. Tandoc EC, Ling R, Westlund O, Duffy A, Goh D, Wei LZ (2018) Audiences' acts of authentication in the age of fake news: a conceptual framework. New Media Soc 20(8):2745–2763
18. Sethi RJ (2017) Crowdsourcing the verification of fake news and alternative facts. In: Proceedings of HT, pp 315–316
19. Vicario MD, Quattrociocchi W, Scala A, Zollo F (2019) Polarization and fake news: early warning of potential misinformation targets. ACM Trans Web 13(2):1–22
20. Zellers R et al (2019) Defending against neural fake news. In: Proceedings of NIPS, pp 9051–9062
21. Lin J, Tremblay-Taylor G, Mou G, You D, Lee K (2019) Detecting fake news articles. In: Proceedings of big data, pp 3021–3025
22. Pinnaparaju N, Indurthi V, Varma V (2020) Identifying fake news spreaders in social media. In: Proceedings of CLEF, pp 1–5
23. Zhou X, Wu J, Zafarani R (2020) Safe: similarity-aware multi-modal fake news detection. In: Proceedings of PAKDD, vol 12085, p 354
24. Rubin VL, Chen Y, Conroy NJ (2015) Deception detection for news: three types of fakes. In: Proceedings of ASIST, p 83
25. Rubin VL, Conroy NJ, Chen Y (2015) Towards news verification: deception detection methods for news discourse. In: Proceedings of HICSS, pp 1–8
26. Tschiatschek S, Singla A, Rodriguez MG, Merchant A, Krause A (2018) Fake news detection in social networks via crowd signals. In: Proceedings of Companion Web Conference, pp 517–524
27. Singhal S, Kabra A, Sharma M, Shah RR, Chakraborty T, Kumaraguru P (2020) SpotFake+: a multimodal framework for fake news detection via transfer learning (student abstract). Proc AAAI 34(10):13915–13916
28. Cui L, Wang S, Lee D (2019) Same: sentiment-aware multimodal embedding for detecting fake news. In: Proceedings of ASONAM, pp 41–48
29. Bhutani B, Rastogi N, Sehgal P, Purwar A (2019) Fake news detection using sentiment analysis. In: Proceedings of IC, pp 1–5
30. Duan S, Peisert S, Levitt KN (2015) HBFT: speculative byzantine fault tolerance with minimum cost. IEEE Trans Dependable Secure Comput 12(1):58–70
31. Androulaki E, Cachin C, Dobre D, Vukoli´c M (2014) Erasure-coded Byzantine storage with separate metadata. In: Principles of distributed systems. Springer, Germany, pp 76–90
32. Liu S, Viotti P, Cachin C, Quéma V, Vukolic M (2016) XFT: practical fault tolerance beyond crashes. In: Proceedings of OSDI, pp 485–500
33. Androulaki E et al (2018) Hyperledger fabric: a distributed operating system for permissioned blockchains. In: Proceedings of EuroSys, pp 1–15

A Secure and Practical Scientific Computation Outsourcing Application in Cloud Computing

V. Sudarsan Rao and N. Satyanarayana

Abstract Cloud computing permits customers with constrained computational capability and communication bandwidth to economically and feasibly outsource their massive-sized computations and operations to a public-oriented cloud having large-scale computation capacity. Cloud possesses the large storage, computational capability as well as software program which may be utilized by customers for the reduction of their overhead in terms of computing. Though, in the case scenario of the outsourcing, the privacy of customer's private data must be preserved. In this research work, we developed a methodology for the outsourcing big-scale eigen count value computational threat to a malevolent-natured cloud that gives input/output information security, verifiability of results, and empirical performance. As far as the calculation approach to discover all the eigenvectors is of computationally high cost for huge-dimensional input data, we have utilized energy-iterative procedure for finding out the biggest eigenvalues well as the relatively corresponding eigen set vector of a certain matrix. In order to save the primitiveness, certain modifications are carried out to the certain input matrix in order to obtain its encrypted matrix that is transferred to the virtual cloud and further perform decryption of the result which is obtained back from the specific cloud entity for obtaining the accurate answer of eigen cost problem. We also provided end-result verification scientific mechanism to detect dishonest entities and supplied theoretical evaluation as well as experimental evaluation finding that explains excessive efficiency, feasibility, correctness, protection, and robust dishonest impending of the given protocol.

Keywords Outsourcing of computation · Cloud computing · Paillier cryptosystem · Data security · Power procedure · Eigenvalue computational problem

V. Sudarsan Rao (✉)
Department of Computer Science and Engineering, Sri Chaitanya Institute of Technology and Research, Khammam, Telangana, India
e-mail: sudharshan.cse2008@gmail.com

N. Satyanarayana
Department of Information Technology, TKR College of Engineering and Technology, Hyderabad, Telangana, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_9

1 Introduction

Distributed computing bears the cost of an enormous assortment of administrations to customers. Among them, outsourcing gigantic scope calculations is transforming into much well-known nowadays. With regards to customers that are computationally illiterate, outsourcing (or gadgets), limited via the sluggish handling memory, performance, and other factors requirements can re-appropriate their calculations and certain private records to cloud which may be prepared with the enormous computational assets.

Although having astounding facilities, as clients' clouds are not included inside the identical depended on domain, there can exist several safety concerns as well as challenges toward this concept of outsourcing [1]. In this, the primary objective is to guard the security of customer's private records. Customer's confidential information can also comprise private records. For preserving records protection, the client requires to encrypt their personal records before the outsourcing and finally decrypt back-end result from the cloud. The second process involves confirming whether the back result is accurate or incorrect; this is important because the cloud system can also produce false outcomes. Also, there can be certain software-based bugs or the hardware-oriented failures in the specific cloud server that might additionally bring about a wrong result. The third undertaking is customer's performance. The version should offer higher efficiency to consumer in perspectives of the computational measure, cost measure, and time.

Putting focus at the engineering as well as clinical computing issues, eigenvalue hassle is a simple and feasible computational problem and in fact possesses a number of use cases. It is utilized in information extraction [2] as well as Google exploits the eigenvector despite of the largest eigen set value of the certain Google matrix in order to decide the specific rank of a web page to seek [3]. Eigenvectors are generally the fundamental key elements to the primary components-oriented analysis that is usually exploited for the dimensionality minimization in face value and different gadget gaining knowledge of programs [4]. Eigenvalues as well as eigenvectors, as entities, are also utilized in the image and processing domain. The certain matrices from those packages are generally massive. When dealing with a large matrix in an eigenvalue problem, outsourcing the computation to an efficient cloud service can be a cost-effective solution. The cloud is a cost-effective option infusion. Again, for the clients as the battery-constrained mobile phones as well as portable appliances, or immersed playing cards, an efficient outsourcing is desired despite the fact that the statistics is in a certain adequate scale [5].

Subsequently, we are administered to layout a protocol behavior that outsources eigenvalue problem set to a public structured cloud and gives performance, security, and strong dishonest entity recognition to the patron. The direct scientific computation procedure for locating all eigenvectors takes $O(n^3)$ time-oriented complexity, that becomes very steeply priced for massive dimension data. So, in this research work, we have constructed the core problem of secretly outsourcing the very large-sized Eigenvalue problem through iterative approach, i.e., power-based method [6,

7] that concurrently fulfills the specific goals of excessive performance, correctness, safety, and sturdy dishonest entity detection [5]. The mechanism results in computational cost reduction. While fixing domestically, massive-scale eigenvalue set problem results in $O(n^2)$. Cost in line with generation, our proposed method incurs O(n) amount of computational complexity for the patron in every new release [8]. From the experimental evaluation and obtained results, we have depicted that as the matrix size becomes larger, user achieves a better performance.

The remaining paper is structured as follows: Sect. 2 discusses theoretical history of the eigenvalue problem. Section 3 represents system model. In Sect. 4, we have got portrayed the model we suggested as a remedy for the eigenvalue problem. Theoretical and the experimental evaluation end results are furnished in Sect. 5. Finally, in Sect. 6, conclusions of the paper are given.

2 Theoretical Background

2.1 Power-Iterative Method

Limit value issues, for example, investigation of vibrating frameworks, structure examination, and electric circuit framework examination decrease to an arrangement of conditions of the structure $Ax + \Lambda x$. Such problems are known as eigenvalue problems, where there is a scalar constant denoted as the eigenvalue, and x is denoted as the eigenvector. Because of simpler execution and lesser computational resource utilization of force iterative strategy, a few applications exploit it often times. Nonetheless, it will observe hands down the biggest eigen esteem and comparing eigenvector. The power emphasis calculation begins with an irregular vector or a guess to the prevailing eigenvector x_0. The iterative condition can be addressed as:

$$x^{k+1} = \frac{A \cdot x^k}{\left\| A \cdot x^k \right\|}$$

Thus, at each cycle, the vector $n \times n$ (non-solitary) coefficient network is A and $A \cdot x^k$ boundary is the scalability factor, which is the biggest computational power of $A \cdot x^k$. At the point when the cycle unites, we obtain the biggest eigenvalue and the associated eigenvector. This strategy can be more effective for a few computational undertakings, in spite of the fact that it observes simply the biggest eigenvalue. For instance, Google utilizes it to compute the PageRank [3] of the archives in their web crawler.

2.2 Homomorphic Encryption

Homomorphic-oriented encryption is a type of mechanism of encryption that permits calculations to be completed on ciphertext, subsequently creating an encoded outcome which, when decoded, matches the consequence of activities carried out on the plaintext [9]. We utilized Paillier cryptosystem [10], that is, to some extent homomorphic encryption framework and have added substance homomorphic-oriented property giving preferable effectiveness over the completely homomorphic method [11]. It is a type of asymmetric key cryptography that utilizes public key procedure.

Assuming that two ciphertexts are getting multiplied, encoded type of expansion of the certain plaintexts is framed. Decoding this, we obtain the amount of two plaintexts (Fig. 1).

$$D(E(x1, r1)) * \left(E(x2, r2) \bmod n^2\right) = x1 + x2 \bmod n \tag{1}$$

If a ciphertext and furthermore a plaintext raise to g are increased, scrambled type of expansion of the plaintexts is shaped.

Decoding this, we obtain the addition of two individual plaintexts.

$$D\left(E(x1, r1) * g^{x2} \bmod n^2\right) = x1 + x2 \bmod n \tag{2}$$

Practically, the Paillier cryptosystem satisfies the following constraints and identities [12]:

$$D(E(x1)) * \left(E(\square x2) \bmod n^2\right) = x1 + x2 \bmod n \tag{3}$$

$$D\left(E(x)^k \bmod n^2\right) = kx \bmod n \tag{4}$$

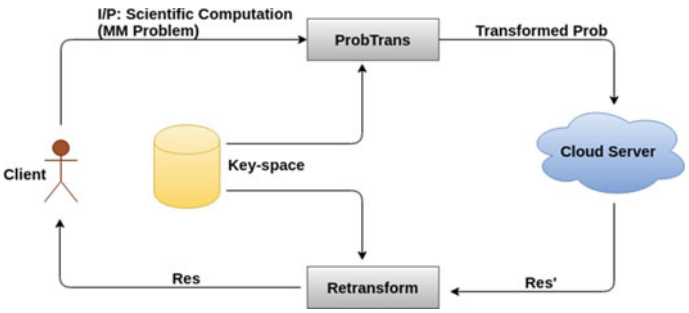


Fig. 1 Client cloud: computational flow of the framework

$$D(E(x1)) * \left(g^{x2} \bmod n^2\right) = x_1 + x_2 \bmod n \tag{5}$$

$$D\left(E(x1)^{x2} \bmod n^2\right) = x_1 x_2 n \tag{6}$$

2.3 Case Study: Usefulness Example for Parallel Computation in Cloud

The parallel execution of the loop in lattice augmentation optimization is dealt with as instance below.

```
#pragmaomp_parallel
Define type as integer thread_numer = openmp_get_thread_number(),
Num_of_threads = openmp_getnumber_threads();
Define type as integer low = N*threadnumber/number of threads,
high = N*(threadnum+1)/number of threads;
apply looping till: for (i=min; i< max; i++)
// procedural code
```

#Pragma omp parallel in addition to *cuBLAS* functional libraries is exploited here in the method (Fig. 2). The application apportioned the necessary lattices and vectors in the GPU memory space, filled them with input information, call the grouping of wanted equal capacities natives, and afterward transferred the outcomes from the GPU memory space back to the host. It is a lightweight library that adds adaptability in grid information formats, input types, register types, and furthermore in picking the algorithmic executions and heuristics through boundary programmability.

Current GPUs are astounding, immensely multi-hung, and with prevalent capacity. With the progression of re-appropriating estimation, it is useful for clients with confined handling resources for re-proper overpowering computational tasks to the cloud server and accordingly quiet massive load of the client. As perpetual thought of assignment lately, requirements of safety, and capability are seriously worried beyond question, especially for securely reallocated network augmentation.

Figure 3 illustrates the parallel computation of network ASCII values across various matrix blocks, represented in blue and orange to indicate a multi-core architecture. There are ordinarily three sorts of data get to skeletons for instance A Square, The Column, and Row—and two skeletons with square access for instance a Warp-by-Warp and Square-by-Block. Each "skeleton" provides a parallelism and unrivaled utilization structure that fragments data shows into data squares and a while later underscores over those particular squares. Data obtain to skeletons bundle n data

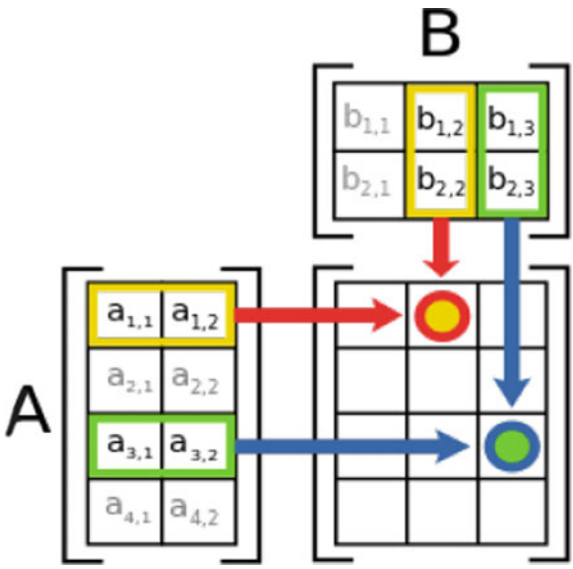


Fig. 2 Scenario of shared memory access

parts into the m super segments size data squares. This m data squares be next allotted crossways p string squares using a 1D or, more than likely 2D configuration plan. Ordinary programming strategies are practical to the decent size data squares toward enabling execution tests intended for different kinds of the parallelism for instance direction stage parallelism (DSP), data stage parallelism (DSP), close by string stage parallelism (CSSP). GPUs achieve unrivaled by gigantic parallelism. Present-day illustrations processors, for example, GTX 1650 V100, even incorporate a great many handling centers. GPU engineering upholds various types of parallelisms: specifically, the parallelisms depict guidance-level parallelism, string-level errand parallelism, and information parallelism. ILP searches in order to remove numerous self-overseeing requests of the successive guidance a stream of data that may then be performed in equal on various handling stages inside the whole processing core.

Figure 4 shows the functionality differences between CPU and GPU architectures. Scientific classification bunches Flynn imitation series and equal into four gatherings, at least one along two tomahawks (guidelines and information): Single-Single-guidance Data (SISD), single guidance multi-information (MISD), multi-directions multi-information (MIMD), and single guidance numerous information (SIMD). These groups are valuable to arrange equal equipment, programming, and formats.

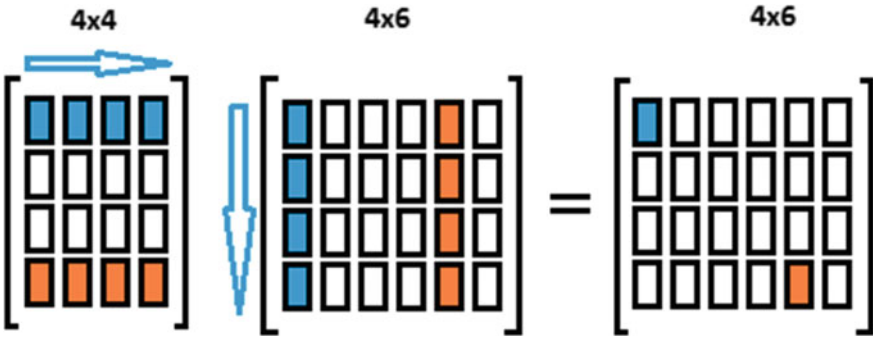


Fig. 3 Parallel matrix multiplication scenario

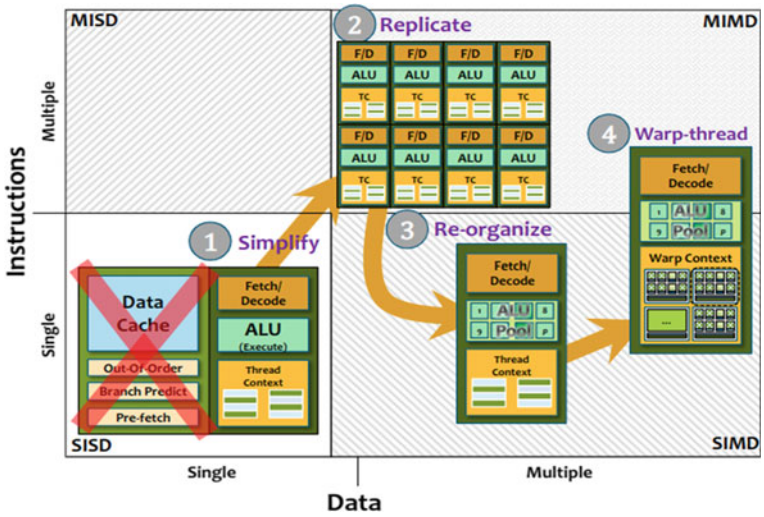


Fig. 4 Functionality differences between CPU and GPU architectures

3 System Prototype Model

Figure 1 shows the framework model for safely rethinking eigen esteem problem. Customer having a limited computational capacity re-appropriates the issue to a certain cloud server possessing immense computational-oriented power and extraordinary programming. To secure information protection, the customer initially encodes the first issue utilizing a private key K to get scrambled issue. Then, at that point, the encoded issue is given to the cloud for an outcome. When the cloud gets scrambled issue, the calculation is done with programming. Then, at that point, cloud returns back the outcome to customer. In the wake of getting the returned outcome, the customer then, at that point, unscrambles the outcome utilizing the mystery key

K and gets unique outcome. Then, at that point, the customer checks the outcome whether or not it is right. On the off chance that right customer acknowledges it and in any case dismisses it.

4 Proposed Method

As tackling eigenvalue computational problem by utilizing direct strategy is tedious for huge lattice, it probably will not be a decent choice for asset-restricted customers. Besides, mathematical strategies for the approximating foundations of polynomial conditions of serious level are much sensitive to adjusting estimate mistakes [13]. Consequently, we have utilized power-iterative strategy that mainly observes the eigenvalue which was considered as the biggest in outright worth and relating eigenvector. We then constructed after presumptions: the coefficient framework A has a predominant eigenvalue with relating prevailing eigenvectors [13]. That implies inside all the eigenvalues corresponding to A, one specific eigenvalue is completely more noteworthy than different eigenvalues in extent. The underlying estimate of the beginning vector x^0 should be a vector with a nonzero value in R_n.

Our proposed methodology possesses mainly three phases, i.e., problem space transformation, problem solving along with the verification of results.

4.1 Problem Transformation

The iterative variety equation can then be depicted as:

$$X^{k+1} = \frac{A \cdot x^k}{\left\| A \cdot x^k \right\|} \tag{7}$$

In above equation, the most costly calculation is matrix vector multipThen, at that pointlication, $A \cdot x^k$. It takes $O\,(n^2)$ time complexity.

For this reason, this estimation is completed by the cloud. Yet, as cloud server can be vindictive, customer cannot send coefficient network An and x straightforwardly to the cloud. For keeping up with the secrecy of information, customer scrambles A by Paillier encryption, in this way gets encoded An and afterward forwards Encrypted (A) toward the cloud. Then, at that point, at every emphasis, customer covers x by an arbitrary vector $r \in R\,n$ and obtains z, where

$$z^k = x^k + r \tag{8}$$

Then, at that point, customer forwards z^k to the cloud. Cloud carries out procedure on z^k and returns encoded rendition of $A \cdot z^k$ to the customer. Customer decodes and

get

$$A \cdot z^k = A\left(x^k + r\right) \tag{9}$$

From (9), for obtaining, client further subtracts c from the $A.z^k$, where $c = Ar$.So,

$$A \cdot x^k = A \cdot z^k - c \tag{10}$$

Then, at that point, customer refreshes x by specific (7) and afterward again mask the X by (8) for the following emphasis and transfers the data to the cloud once more. This interaction proceeds until the data convergence.

4.2 Problem Solving

In critical thinking stage, we have thought about the general strides for tackling eigenvalue issue. Cloud processes grid vector augmentation $A \cdot z^k$. in every cycle. The purpose is to illustrate the first round of the recursive cycle as a way to describe the entire problem-solving process.

1. Initially, the client estimates the eigenvector's first approximation, i.e., x^0, $x^0 = {x_1}^0, {x_2}^0, \ldots\ldots {x_n}^0$. The client then creates a random vector r, z^0 by $z^0 = x^0 + r$. Client does encrypts A by using Paillier encryption, and encrypted (A) and z^0 are sent to the cloud. After that, it creates c by c which is equal to Ar.
2. Encrypted (A) and z^0 on the encrypted matrix using the additive homomorphic characteristic of the Paillier encryption, the cloud computes the value Encrypted $(A.z^0)$.
3. When Encrypted $(A.z^0)$ is received from the cloud, the client decrypts it with the private key and returns $A.z^0$. The client then uses (10) to acquire $A.x^0$ and then uses (12) to update the next estimate of x (7).

This interaction proceeds until intermingling happens. Customer checks for union in every emphasis. In kth emphasis, customer sends k th guess z^k to the specific cloud and then cloud ascertains Encrypted $(A \cdot z^k)$ and sends it back to the customer. Then, at that point, customer refreshes next guess $z^k + 1$ and then transmits it to cloud. Customer forwards Encrypted (A) first cycle.

4.3 Result Verification

Cloud server might be faithless and provide the erroneous outcome to the customer. Along these lines, customer needs to check the outcome for distinguishing cheating. At the point when cycle joins, customer performs result check by the accompanying condition,

$$Ax = \lambda x \tag{11}$$

Where the maximum eigenvalue is denoted by λ, and x is the corresponding eigenvector. If the result of Ax is equal to λx, then the customer accepts the result as correct. Otherwise, the result is rejected.

A genuinely malevolent cloud server can return subjective responses in every cycle. Thus, a circumstance can happen when the emphasis could never merge, squandering the assets of the customer [14]. To forestall such wastage of assets, customer can preset an adequately huge worth so that program will naturally combine after specific emphases. After combination, customer utilizes (11) to actually look at the accuracy of the outcome. Once more, customer can utilize numerous cloud suppliers to moderate the dangers related with the noxious cloud.

5 Theoretical Examination and Experimental Results

5.1 Theoretical Analysis

(1) *Convergence Analysis*: When managing iterative methods, it is must to decide if and when the cycle will meet [14]. Customer characterizes a mistake limit and decides intermingling if,

$$x^k - x^{k+1} \leq \in \tag{12}$$

where worth of ∈ is tiny, near nothing.

(2) *Input Security Evaluation*: We can claim that what we have offered is a good idea, i.e., convention secures customer's feedback information protection in the event that the cloud cannot get the first lattice A from the encoded framework Encrypted (A) along with unique vector x from the scrambled vector z in every emphasis [5]. We scramble A by Paillier-type cryptosystem, so the cloud server cannot recover A from the Encrypted (A). In every emphasis of our given strategy, the cloud just observes the plaintext of z^k and the encoded variant of lattice Encrypted (A). Regardless of whether we transmit z to the cloud in plaintext, it is encrypted and protected on the grounds that unique vector x is arbitrarily covered by an specific vector r. No data of x would be spilled as long as r remains mystery by the customer [14]. For the high security, every component of arbitrary vector r ought to be 128 pieces or even more. At present, on the off chance that a supercomputer could actually take a look at a billion (1018) keys each second, breaking a 128-cycle key space by animal power needs around 3 × 1012 years [15]. For improving the protection of input information, we can additionally utilize an arbitrary scaling factor $a^k \in Z$ for every cycle to break the connection capacity of two successive emphases of the convention [14]. In particular, rather than sending z^k to the cloud server, the customer can deploy

a^k. Z^k for the k th cycle. Subsequently, the cloud server particularly can never build up direct conditions got from $a^k \cdot z^k$ and $a^{k+1} \cdot z^{k+1}$ [14].

(3) *Security Analysis of the Output*: We can claim that what we have offered is a good idea, i.e., convention secures yield information protection in the event that the cloud cannot get the right $A \cdot x^k$ from the Encrypted $(A \cdot z^k)$ in every iteration. If the client forwards $a^k \cdot z^k$ for the k th cycle to the particular cloud, the cloud servers close back the scrambled outcome Encrypted $(a^k \cdot A.z^k)$. Then the customer essentially unscrambles the Encrypted $(a^k \cdot A \cdot z^k)$ to obtain $a^k \cdot A.z^k$ and partitions every part with a^k. For the following emphasis, one more arbitrary scaling factor a^{k+1} is increased to z^{k+1} preceding shipped off to the cloud server [14].

(4) *Efficiency Examination and Analysis*: If customer needs to take care of the opening issue locally, client requires to figure network vector augmentation $A \cdot x^k$ in every emphasis, which takes $O(n^2)$ intricacy. This exorbitant calculation in every emphasis brings about gigantic weight for the customer. Then again, in our proposed convention, in the underlying arrangement stage, customer needs to scramble A by Paillier-oriented cryptosystem that takes $O(n^2)$ intricacy and produces c by $c = Ar$ which necessitates complication. The client must play out these two $O(n^2)$ calculations just a single time. Then, at that point, in every one of the cycles, customer just requires to frame z^k, decode to obtain $A \cdot z^k$, use the vector of Encrypted $(A \cdot z^k)$. x^k from $A \cdot z^k$, and update next estimation. Every one of these tasks takes just $O(n)$ intricacy. After intermingling, result confirmation is carried out by customer that takes $O(n^2)$ intricacy. Costly lattice vector duplication in every emphasis is completed by the cloud all things considered. Therefore, computational weight diminishes for the customer and productivity increments. Once more, as the encryption on every component of the coefficient grid An is autonomous, it tends to be effectively parallelized [14]. Hence, it is not obligatory for the customer to stack the entire coefficient network in the memory in any case and customer's computational weight further abatements.

5.2 Experimental Evaluation and Obtained Results

In our test, we have utilized a similar machine arrangement both for the cloud and customer. Then, at that point, we have investigated the exhibition of the customer side. We have seen how much proficiency is accomplished by the customer by utilizing our proposed convention and how much computational weight is decreased due to re-appropriating. We have disregarded the correspondence inertness between the customers. Since the application's inception, the cloud has been used, and calculation overwhelms the execution time as displayed in our examination [5]. The test is done utilizing eclipse pre-installed workstation. In our examination, we have executed power-iterative strategy to take care of eigen esteem issue where the most costly computation lattice vector augmentation is utilized by the cloud.

Our primary objective is to give proficiency to the customer by utilizing rethinking. For this reason, effectiveness gain is determined. Effectiveness gain can be named as client improved amount speedup. Here, the client-oriented speedup is the proportion of the specific time that is required for customer assuming the calculation is done on a local level, and the amount of time it takes is required by the customer's calculation assuming that re-evaluating is picked [5]. Meaning of certain boundaries utilized in our investigation is given in Table 1.

Our trial result is depicted in Table 2. We can see that customer speedup increments as lattice aspect n gets bigger. In every emphasis, $O(n^2)$ calculation is carried out by cloud and $O(n)$ calculation is carried out by customer. Thus, customer gets wanted computational investment funds when n tends adequately huge. We have likewise seen that the worth of client speedup is more noteworthy than 1 that implies we have accomplished wanted execution gain.

Table 1 Definition of notations

Notation	Description
t_{original}	Time needed for certain client to resolve the original issue on a local level if outsourcing is not used
t_{cloud}	The amount of Time it needed for the cloud to calculate the data calculations that are sent to the cloud
t_{client}	The amount of time it takes for the client to encrypt data, decryption, and result verification if outsourcing is used
Client Speedup	$\frac{t_{\text{original}}}{t_{\text{client}}}$

Table 2 Eigenvalue problem experimental result

Data	Original (base) problem	Encrypted problem		Client's speedup
Dimension vector, n	t_{original}	t_{cloud}	t_{client}	$t_{\text{plain}}/t_{\text{client}}$
52	0.3232	0.3231	0.1804	1.6364
104	1.2820	1.2814	0.6136	2.1304
210	7.3113	7.3115	2.2362	3.2699
307	21.6119	21.6715	4.9385	4.3894
410	47.5579	47.5862	9.1376	5.2224
522	84.9727	84.9722	13.3156	6.3771
756	273.7267	273.7263	31.7612	8.6109
1016	485.7187	485.7183	51.0104	9.5337

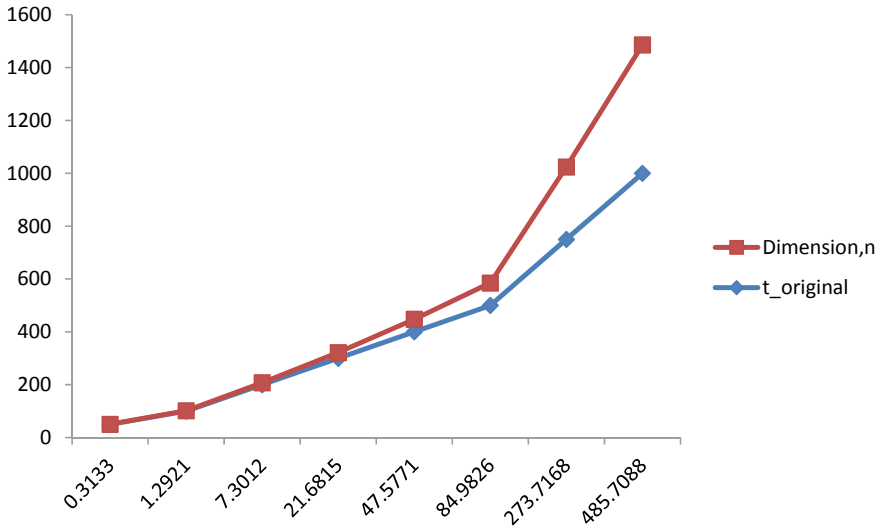


Fig. 5 Time required for the client to resolve the original issue on a local level if outsourcing is not used

As we have executed the examination in similar workstation, that implies we have taken into consideration customer's machine as both customer and cloud, we possess that the time needed for cloud is adequately enormous. Time needed for the cloud is provided in the third section. As the genuine cloud is furnished with superior execution equal processing and virtualization, network vector duplication can be separated into sub-undertakings checkerboard block [16, 16]. Thus, in the event that genuine cloud is utilized, the amount of time it takes for the cloud to execute network vector duplications would be a lot more modest contrasting the outcome accomplished in our analysis as we have thought about customer's machine as cloud. All things considered, we would have the option to change the condition for client speedup as,

$$\text{Client\_ Speedup} = \frac{t_{\text{original}}}{t_{\text{client}}} \tag{13}$$

If there should be an occurrence of genuine cloud, on the off chance that we employ (16) to compute client speedup, we can accept that we would in any case accomplish huge speedup as t_{cloud} would be tiny for genuine cloud (Figs. 5, 6, 7, and 8).

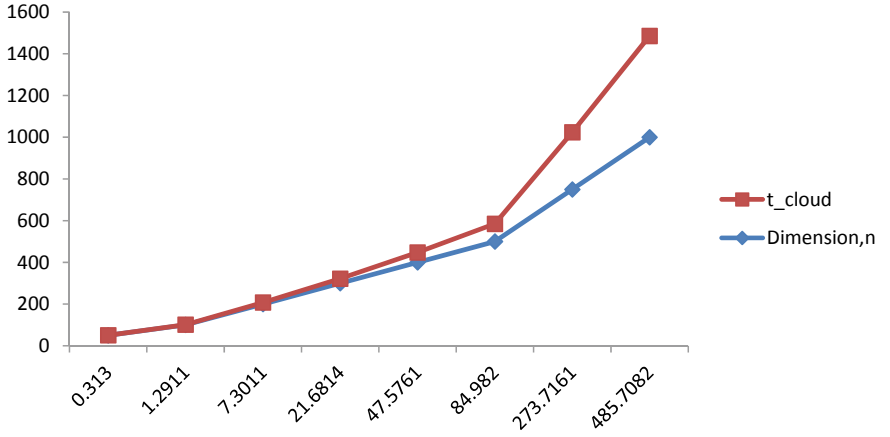


Fig. 6 Amount of time it takes for the cloud to calculate the data calculations that are sent to the cloud

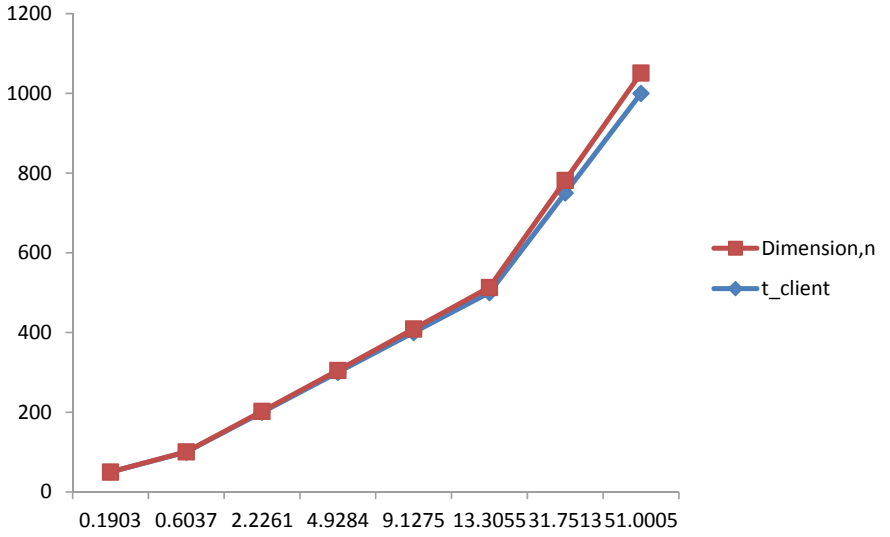


Fig. 7 Client encrypts the original problem

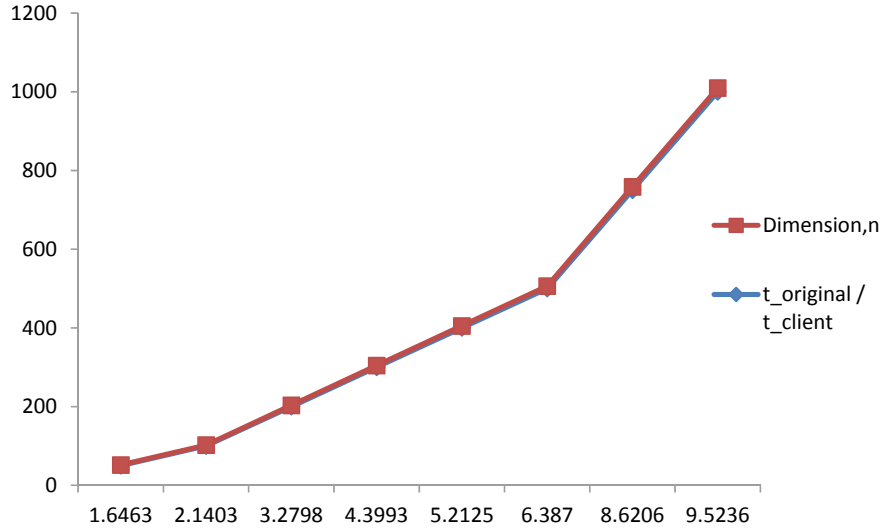


Fig. 8 Client speedup in the ratio

6 Conclusion

This paper provides a convention for taking care of eigenvalue issue utilizing outsourcing. Worried with regards to customer's weight in the event of taking care of huge scope eigenvalue issue, we further proposed a thought of re-appropriating that would build the customer's proficiency and brings down the weight of customer. In corresponding with giving effectiveness to customer, this convention likewise gives security to customer's private information and result check. This convention diminishes customer's computational weight and gives better effectiveness. Bearings for additional exploration include: (1) implementing of this convention utilizing genuine cloud. (2) Identifying a superior outcome confirmation strategy that would certainly perform result check at each and every iteration. (3) Identifying the new significant logical and designing computational undertakings and afterward planning conventions to tackle them [5]. (4) Adding the result check for a few existing conventions that do not have the result confirmation channel.

References

1. Cloud Security Alliance (2009) Security guidance for critical areas of focus in cloud computing. http://www.cloudsecurityalliance.org
2. Deerwester SC, Dumais ST, Landauer TK, Furnas GW, Harshman RA (1990) Indexing by latent semantic analysis. JASIS 41(6):391–407

3. Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. In: International conference on world wide web
4. Jolliffe IT (1986) Principal component analysis. Springer, New York
5. Lei X, Liao X, Huang T, Li H, Hu C (2013) Outsourcing large matrix inversion computation to a public cloud. IEEE Trans Cloud Compt 1(1):1–11
6. Arnoldi WE (1951) The principle of minimized iterations in the solution of the matrix eigenvalue problem. Quart Appl Math 9:1729
7. Cullum JK, Willoughby RA (1985) Lanczos algorithms for large symmetric eigenvalue computations. Cambridge University Press, Cambridge
8. Carpentieri B (2002) Sparse preconditioners for dense linear systems from electromagnetic applications. PhD dissertation, CERFACS, Toulouse, France
9. (2015) Homomorphic encryption [Online]. Available: https://en.wikipedia.org/wiki/Homomorphic_encryption.
10. Paillier P (1999) Public-key cryptosystems based on composite degree residuosity classes. In: EUROCRYPT, Springer Verlag, pp 223–238
11. Gentry C (2009) Fully homomorphic encryption using ideal lattices. In: Annual ACM symposium on theory of computing, ACM, New York, NY, USA, pp 169–178
12. (2015) Paillier cryptosystem [Online]. https://en.wikipedia.org/wiki/Paillier_cryptosystem
13. Chen X, Huang X, Li J, Ma J, Lou W, Wong DS (2014) New algorithms for secure outsourcing of large-scale systems of linear equations. IEEE Trans Inf Forensics Secur 10(1):69–78
14. Wang C, Ren K, Wang J, Wang Q (2013) Harnessing the cloud for securely out sourcing large-scale systems of linear equations. IEEE Trans Parallel Distrib Syst 24(6):1172–1181
15. (2015) Parallel methods for matrix-vector multiplication [Online]. https://www.hpcc.unn.ru/mskurs/ENG/DOC/pp07.pdf
16. Velkoski G, Ristov S, Gusev M (2013) Various algorithms' behavior in the cloud. In: The 10th conference for informatics and information technology (CIIT 2013)

Navigating the Complexities of Cryptography: Trends, Problems, and Solutions

Varad Joshi and Shanti Verma

Abstract Cryptography is a complex field that has evolved significantly over the years. It is essential for the secure communication of sensitive information online. However, with the increase in cyber-attacks, keeping up with the latest trends, and identifying potential problems has become more important than ever. In this paper, authors try to explore the latest trends in cryptography, highlighting the problems that come with them, and discussing potential solutions for each. From quantum computing to blockchain technology, authors will delve into the complexities of cryptography, providing you with the knowledge and tools to navigate through this essential field with confidence.

Keywords Cryptography · Blockchain · Quantum computing · Homomorphic encryption · Post-quantum cryptography · Secret key cryptography

1 Introduction

Cryptography is the practice of securing communication from third-party entities or adversaries. It involves ciphering messages, documents or even data files to protect them from unauthorized access or usage. Cryptography has been in use for centuries, from ancient Greece to modern-day technology. It has evolved from simple substitution ciphers to complex mathematical algorithms that are almost impossible to crack. The need for cryptography has increased as technology advancements have led to new forms of communication and data storage. Cryptography is used in secure communication channels such as that used in online transactions, protecting sensitive information such as passwords and securing data storage [1]. Cryptography is

V. Joshi (✉) · S. Verma
LJ University, Ahmedabad, Gujarat, India
e-mail: varadjoshi13@gmail.com

S. Verma
e-mail: vermashanti@gmail.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_10

also used by governments, military and intelligence agencies to secure classified information.

The cryptography is divided into two different types: symmetric and asymmetric. Symmetric cryptography uses the same key to encrypt and decrypt messages, while asymmetric cryptography uses two keys for the same [2]. Cryptography is not perfect and there are still vulnerabilities that attackers can exploit. Therefore, it's essential to understand cryptography's trends, problems, and solutions to navigate its complexities.

2 Trends

Cryptography has come a long way since its beginning in 1900 BC and the latest developments in this field are shaping the future of digital security. One major trend in cryptography is the use of blockchain technology. Blockchain is a decentralized ledger system that uses cryptography to protect the data stored within it from tampering or unauthorized access. Blockchain technology has emerged as a new era in cryptography with its decentralized and secure nature. It is essentially a digital ledger of transactions that is distributed across a network of computers. Each block in the chain contains a unique cryptographic code, making it almost impossible to tamper with the data stored within [3]. Blockchain utilizes the following encryption building blocks: (1) Public Key Cryptography: This type of cryptography is used for digital identities and encryption. (2) Zero-Knowledge Proof: Show that you are aware of something without disclosing it. (3) Hash Functions: One-way pseudo-random algebraic expressions. Merkle trees used the hash function as one component of the block heading [4].

Another trend is the development of quantum-resistant cryptography. With the emergence of quantum computers, traditional cryptographic systems may become vulnerable, and quantum-resistant cryptography is being developed in response to this threat [5].

Homomorphic encryption is another area of cryptography that is seeing rapid development. Homomorphic encryption allows data to be processed in an encrypted form, which increases security and privacy, while still allowing for the data to be used for analysis and computation [6]. The goal of homomorphic processes is to increase the reuse of key-switch clues, which is essential for reducing data movement from sender to receiver [7].

The post-quantum cryptography is another sector that is gaining a lot of attention. This type of cryptography is designed to be water-proof to attacks by quantum computers, which are expected to be developed in the near future [8, 9]. Post-Quantum Cryptography (PQC) is a distinct encryption scheme that can protect against quantum computer assaults. The National Institute of Standards and Technology (NIST) hopes to develop a standard for quantum-resistant cryptography by 2022 by implementing

this contest, and the target algorithms include the signature Key Exchange Mechanism. (KEM). It is worth noting that the symmetric encryption technique is not listed as a target. NIST chose 15 algorithms as candidates for three rounds in July 2020. Seven algorithms were chosen as final finalists from a pool of fifteen. There are now five forms of quantum-resistant encryption which are Lattice, Code, Hash, Isogeny, and Multivariate [10].

3 Problems

While cryptography is an essential tool for securing data, there are still some problems that need to be addressed to ensure that it remains effective in the future.

Some of the most significant challenges in key management include safe key transit (in a symmetric key setup) and the high overhead required, which makes many of these systems unfeasible. Furthermore, no standard key management infrastructure exists, and each scheme must be adapted to match the network and security needs of distinct systems [11].

Another notable research challenge found with the hardware-oriented authentication approach for Advanced Metering Infrastructure (AMI) was the recommendation of a novel key management mechanism (KMS) for the AMI system. This type of key management will employ the key graph and key management modes such as unicast, broadcast, and multicast. Because of intrusive memory, this key management is prone to spoofing and tampering attacks. The most recent relevant research challenge was a proposal to integrate physical key generation (PKG) and physically unclonable functions (PUFs) across a wireless channel to safeguard the link between end users and original equipment manufacturer servers. The basic difficulty is that no explanation for the real-time demands in computer and secure communication subsystems has been identified. It is also vulnerable to man-in-the-middle attacks [11].

These approaches were tested in order to create a key distribution mechanism for Smart Grids (SGs), so that Smart Meters (SMs) and service providers may authenticate and communicate securely using a session key. The problem is that this technique is vulnerable to ephemeral secret leaks and privileged-insider attacks, and it provides inadequate secrecy for end-user credentials in SMs. Because SMs can contribute to a wide range of applications, adequate mitigation of SM-related attack manipulations, while taking into account modern technology, can limit the devastating consequences, whether on the level of human life or infrastructure [11].

Another significant problem is the issue of key management. The more keys that are in use, the more difficult it becomes to manage them effectively. This can lead to errors, which in turn can compromise the security of the data. It is essential to have secure and efficient key management systems that can handle large numbers of keys [12].

Another challenge is the issue of standardization. Cryptography is a complex field, and there are many different algorithms and protocols in use today. This can make it

difficult to ensure that all systems are compatible and interoperable. Greater efforts need to be made to standardize the use of cryptography, to ensure that it is deployed effectively and that all systems can communicate with each other securely.

Finally, there is the issue of user education. Cryptography can be complex and difficult to understand, and many users may not be aware of the risks associated with insecure systems. It is important to provide adequate education and training to users, to ensure that they understand the importance of strong encryption and how to use it effectively. By addressing these problems, we can ensure that cryptography remains an effective and essential tool for securing data in the years to come.

4 Solutions

Cryptography is a complex field with several challenges and issues that need to be addressed. Fortunately, there are several strategies and solutions that can be employed to overcome these challenges and ensure cybersecurity.

One of the key solutions in cryptography is the use of multi-factor authentication. This involves requiring users to provide more than one form of identification before accessing a system or service. This could be a combination of something the user knows, such as a password or PIN, something the user has, such as a smart card or token, and something the user is, such as a biometric identifier like a fingerprint or facial recognition.

Secret Key Cryptography is another form of solution. It is often classified as either a stream cipher or a block cipher. Stream ciphers work on a single bit at a time and use some kind of feedback mechanism to keep the key changing. A block cipher gets its name from the fact that it encrypts one block of data at a time using the same key on each block. In general, while using the same key, the same plaintext block will always encrypt to the same ciphertext in a block cipher, but the same plaintext will encrypt to a different ciphertext in a stream cipher [1].

Another solution is Public Key Cryptography (PKC). In this a person cannot readily discern the other key even though the two keys used by generic PKC that are mathematically connected. When encrypting plaintext, one key is used, while decrypting ciphertext uses the other key. Here, it's crucial to remember that both keys must be used for the procedure to proceed; it doesn't matter which key is applied first. This strategy is also known as asymmetric cryptography as two keys are needed [1].

One of the keys in PKC is labeled the public key and may be distributed as widely as the owner desires in terms of advertising. The other key, known as the private key, is never made available to another one. Sent messages using this system are simple to understand [1].

Another alternative is the hash function. In a sense, no key algorithms include hash functions, commonly known as message digests and one-way encryption. The plaintext is instead used to calculate a fixed-length hash value, making it impossible to determine the plaintext's length or contents.

Several reports assert that no two files can have the exact same hash value, which is a common misconception about hash functions. In actuality, this is incorrect. Think about a hash function that generates a value that is 128 bits long. Naturally, there are 2^{128} potential hash values. Yet there are infinitely many potential files, and 2^{128} is just one of them. Consequently, there must be more than one file that may have the same 128-bit hash value. In fact, there must be an endless number of files [1].

Authentication in Smart grids (SGs), Smart meters (SMs) are significant entities since there is a requirement for data networking reliability in SG security. There are two possible solutions to this problem [11].

First, the ring oscillator physically unclonable functions (ROPUFs) have proposed a hardware-oriented authentication method for AMI. By creating keys from the configuration of the integrated circuit or field programmable gate array (FPGA) chip within the SM, ROPUFs can safeguard AMI systems from attacks to data integrity and confidentiality. The other option is to use a three-factor authentication mechanism to boost the efficiency of renewable energy-based SGs [11].

Another way of solution is Chaos-Based Solution. There are two ways to implement symmetric chaos encryption: stream-based and block-based. A random-like (pseudo-random) number obtained from the chaos map is used by stream-based ciphers to encrypt plaintext. The plaintext is broken up into blocks in block-based encryption, which uses encryption techniques including substitution and shift rotation. The basic core of the block-based cipher is thought to be substitution boxes, which are extensively employed in all common block encryptions like Data Encryption Standard (DES) and Advanced Encryption Standard (AES). As non-linear transformations, these boxes are quite important. Algorithm resistance against external assaults like statistical or differential attacks is prompted by appropriate substitution boxes [13].

Moreover, cryptographic key management is another area that requires attention. The management of cryptographic keys must be done in a secure manner, and the keys must be stored and distributed carefully. Solutions such as key escrow and key rotation can be employed to ensure that keys are not compromised. Finally, the use of blockchain technology is gaining momentum in the field of cryptography. Blockchain provides a decentralized and secure way of storing and sharing data, and its applications in cryptography are numerous. By using blockchain, it is possible to create secure and tamper-proof records of transactions, secure data sharing, and identity management.

In summary, while cryptography is a complex field, there are several strategies and solutions that can be employed to address key issues and ensure cybersecurity. Multi-factor authentication, quantum-resistant encryption algorithms, cryptographic key management, and blockchain technology are just some of the solutions available to navigate the complexities of cryptography [1].

5 Cryptography Tool: Kerberos

The Massachusetts Institute of Technology (MIT) created Kerberos as a network authentication system to address network security issues. When a client and a server are linked through an unsecured network, the Kerberos protocol allows them to use strong cryptography to verify their identities to each other. After authenticating with each other, the client and server can encrypt their correspondence to ensure privacy and data integrity [14]. For public key authentication, authorization data transfer, and delegation, Windows OS presently uses Kerberos version five and extensions. A security support provider (SSP) permits the Kerberos authentication client to be implemented, suggesting that the Kerberos authentication client is accessible via the Security Support Provider Interface (SSPI). Kerberos uses a Key Distribution Center (KDC) that is connected with the domain controller's other Windows Server security services. Because a security account database is required, the KDC makes use of the domain's Active Directory Domain Services database [14].

The Kerberos authentication procedure may be broken down into six parts, as seen in Fig. 1. To begin, a client seeking to access a service submits a request to the authentication server for an authentication ticket (i.e., ticket-granting ticket [TGT]). The authentication server then validates the client's access to the resources and returns an encrypted TGT and a session key. Third, the client transmits the TGT, together with the requested resource, to the Ticket-Granting Service (TGS), which is generally the KDC. The TGS then supplies the client with a service-specific valid session key. The client then hands over the key to the service. In exchange, the service allows the now-authorized client access [14].

As an analogy for Kerberos, Halback presents the example of a carnival. When you arrive at a carnival, you must enter through the front gate, where you will be

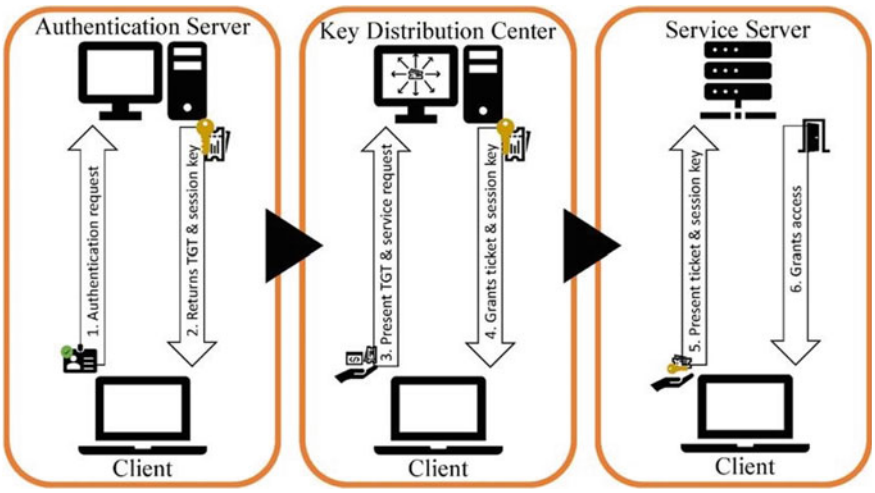


Fig. 1 Authentication process of Kerberos [14]

given a wrist band. This is analogous to turning on your computer and entering your credentials to get access to your workplace network. During this procedure, your computer asks the domain controller for the TGS (a wrist band) to enter the realm (carnival). To ride the rides, you must first enter the funfair and walk to a ticket counter to purchase tickets. The rides represent access to a file sharing server storing your work files, and the tickets represent the service-specific valid session key and ticket. You wait in line for a certain ride after receiving your tickets (ticket and session key) from file server. The ride attendant validates that you are tall enough and of legal riding age, collects your ticket, and allows you to board the ride. This is analogous to the client submitting the session key and ticket to the file server, which grants access to the files if the client has permission. Your ticket would only enable you to access the ride (server) once. You'd have to return to the ticket booth and repeat the process if you wanted to ride again (access the server again) or ride another ride (access a different server) [14].

6 Conclusion

In conclusion, cryptography plays an increasingly important role in our lives. Its impact is felt in various aspects of our daily routines, from securing our financial transactions to protecting sensitive communications between individuals, government agencies, and businesses.

As technology continues to advance, new cryptographic techniques are emerging to address the challenges that come with securing data in an increasingly interconnected world. However, with new cryptographic technologies come new problems and vulnerabilities that must be addressed. Keeping up with the latest trends, best practices, and solutions is critical to maintaining the security of our digital lives.

In summary, cryptography will continue to be a crucial tool for securing our digital world. It is incumbent upon all of us to stay informed and educated about the latest trends and techniques in the field to ensure that our data remains secure and protected.

References

1. Kessler GC. An overview of cryptography
2. May P, Ehrlich HC, Steinke T (2006) ZIB structure prediction pipeline: composing a complex biological workflow through web services. In: Nagel WE, Walter WV, Lehner W (eds) Euro-par 2006, vol 4128. LNCS. Springer, Heidelberg, pp 1148–1158
3. Kh-Madhloom J (2022) Dynamic cryptography integrated secured decentralized applications with blockchain programming. Wasit J Comput Math Sci 1(2):21–33
4. Guo H, Yu X (2022) A survey on blockchain technology and its security. Blockchain Res Appl 3(2):100067. ISSN 2096-7209. https://doi.org/10.1016/j.bcra.2022.100067
5. Zhang P et al (2021) A blockchain system based on quantum-resistant digital signature. Secur Commun Netw, pp 1–13

6. Al Badawi A, Bates J, Bergamaschi F, Cousins DB, Erabelli S, Genise N, Halevi S, Hunt H, Kim A, Lee Y, Liu Z, Micciancio D, Quah I, Polyakov Y, Saraswathy RV, Rohloff K, Saylor J, Suponitsky D, Triplett M, Vaikuntanathan V, Zucca V (2022) OpenFHE: open-source fully homomorphic encryption library. In: Proceedings of the 10th workshop on encrypted computing & applied homomorphic cryptography (WAHC'22). Association for Computing Machinery, New York, NY, pp 53–63. https://doi.org/10.1145/3560827.3563379
7. Samardzic N, Feldmann A, Krastev A, Devadas S, Dreslinski R, Peikert C, Sanchez D (2021) F1: a fast and programmable accelerator for fully homomorphic encryption. In: MICRO-54: 54th annual IEEE/ACM international symposium on microarchitecture (MICRO '21). Virtual Event, Greece. ACM, New York, NY, USA, 15 Pages. https://doi.org/10.1145/3466752.3480070
8. Joseph D, Misoczki R, Manzano M et al (2022) Transitioning organizations to post-quantum cryptography. Nature 605:237–243. https://doi.org/10.1038/s41586-022-04623-2
9. Paquin C, Stebila D, Tamvada G (2020) Benchmarking Post-quantum Cryptography in TLS. In: Ding JP, Tillich JP (eds), Post-quantum cryptography, PQ crypto 2020, Lecture notes in computer science, vol 12100. Springer, Cham. https://doi.org/10.1007/978-3-030-44223-1_5
10. Ahn J, Kwon H-Y, Ahn B, Park K, Kim T, Lee M-K, Kim J, Chung J (2022) Toward quantum secured distributed energy resources: adoption of post-quantum cryptography (PQC) and quantum key distribution (QKD). Energies 15(3):714. https://doi.org/10.3390/en15030714
11. Abdalzaher MS, Fouda MM, Emran A, Fadlullah ZM, Ibrahem MI (2023) A Survey on key management and authentication approaches in smart metering systems. Energies 16(5):2355. https://doi.org/10.3390/en16052355
12. Abdalzaher MS, Fouda MM, Emran A, Fadlullah ZM, Ibrahem MI (2023) A Survey on key management and authentication approaches in smart metering systems. Energies 16:2355. https://doi.org/10.3390/en16052355
13. Sharafi M, Fotouhi-Ghazvini F, Shirali M, Ghassemian M (2019) A low power cryptography solution based on chaos theory in wireless sensor nodes. IEEE Access 7:8737–8753. https://doi.org/10.1109/ACCESS.2018.2886384
14. Demers D, Lee H (2022) Kerberoasting: case studies of an attack on a cryptographic authentication technology. Int J Cybersecur Intell Cybercrime 5(2):25–39. https://vc.bridgew.edu/ijcic/vol5/iss2/3

Food-Block: IoT–Blockchain-Based Framework for Agriculture Food Supply Chain

Martin Parmar and Parth Shah

Abstract India has the second most massive agricultural food production in the world. However, due to the shortfall of traditional agriculture food supply chain management (SCM), India suffers significant food loss that affects the economy and limits farmer income. Recently, blockchain has emerged as a potential substitution of traditional food supply chain. Blockchain is a fully digitized technology that ensures end-to-end integrity with complete transparency of data. Due to the advancement of technology, traditional farming integrates with blockchain, Internet of Things (IoT), artificial intelligence (AI), and many other to bring significant innovation in the agricultural and food supply process. The technologies accumulate processes and analyze the information with minimum human intervention and with less error-prone data. This paper carries existing research on agriculture food supply and identified the potential of blockchain to address the issues. The paper proposed a blockchain-based Food-Block framework with technical specifications for future system development using IoT–blockchain integration.

Keywords Food supply chain · Blockchain · IoT–blockchain integration · Distributed ledger · Smart contract

1 Introduction

Advances in IoT technology have been paced in the previous decade, and presently, there is a new expansion of technology such as blockchain and artificial intelligence (AI). Earlier blockchain is known only for cryptocurrencies such as bitcoin, which is the first application of blockchain. A blockchain is an advanced exchange record kept up by a system of different registering machines that do not depend on a confided

M. Parmar (✉) · P. Shah
Chandubhai S. Patel Institute of Technology (CSPIT), Charotar University of Science and Technology (CHARUSAT), Anand, Gujarat, India
e-mail: martinparmar.ce@charusat.ac.in

P. Shah
e-mail: parthshah.ce@charusat.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_11

outsider [1]. Distinct trade data records block regulated through express programming stages that license the data to be transmitted, arranged, set aside, and addressed in a coherent structure. The blockchain's unique arrangement is such that every block has a header in its block along with a stamp of the period, trade data, and the link to the previous block. A hash is being made for every block and becomes implied in the heading of the next block. Figure 1 presents brief information of each parameter of block-header.

All the blocks maintain a single long chain. The blocks are added in a chronological fashion after another block shown in Fig. 2. The block is validated by the miners after finding the correct nonce given in the block-header. To find the correct nonce, miners must invest their computing power to generate more hashrate to validate the block among the rest of the miners. On successfully validating the block, the miner broadcasts the correct nonce to other miners into the network. The final consent will be given after verifying the given nonce by the miners as per the proof-of-work (PoW) consensus mechanism. Finally, the block is added into blockchain immutably to all distributed ledgers.

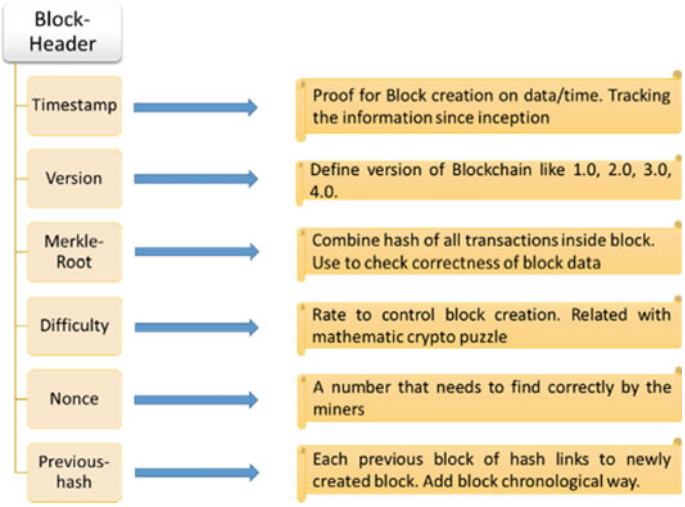


Fig. 1 Block-header

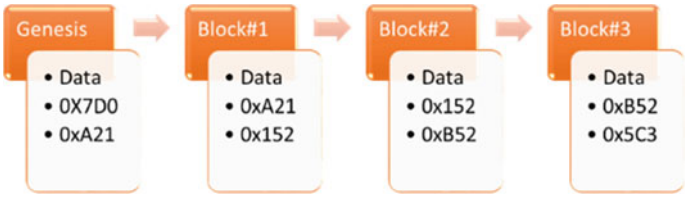


Fig. 2 Block of chain

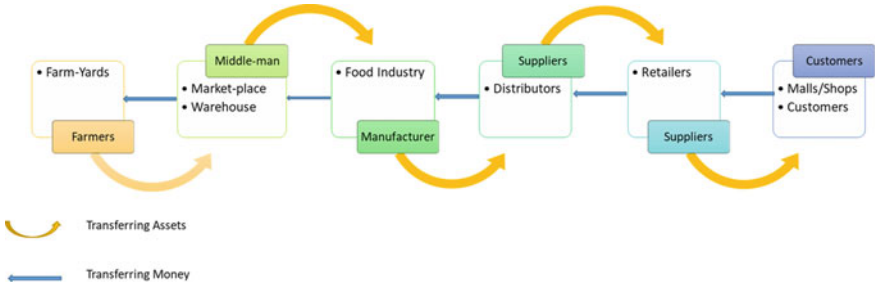


Fig. 3 Movements of food and money in a food supply chain

Blockchain's main feature is that it keeps a transparent view among the network of nodes, and it works on the agreement of the nodes [2]. The consensus algorithm forms a critical mechanism in creating the new blocks in the blockchain network. Consensus algorithms such as proof-of-work (PoW), proof-of-stake (PoS), and proof-of-authority (PoA) are being used to verify nodes to be added [3].

The blockchain-based supply chain promises the reliability of traceability and authentication and removes the intermediary person [4]. The applications of blockchain in supply chains are extended with the assistance of the Internet of Things (IoT) and machines providing operational data subsequently [5].

1.1 Food Supply Chain

As shown in Fig. 3, the food goes through multiple parties such as farmers, intermediaries, manufacturers, suppliers, and retailer, and finally, it comes in the hand of the customer. Thus, the movement of food from the farmer to the customer and money has a domino-like movement. The money goes in the opposite direction of the food from the customer to the farmer, the retailer, distributor, and processor [6].

1.2 Food Safety

The safety of the food lies within the caring of food, keeping it away from contamination, and ensuring that the food saves enough supplements in it to have a good impact of food on the person eating it [6]. Menacing food and water show that the food may infect through the germs or contain germs that can cause illnesses or contamination of food.

1.3 Food Security

According to the Food and Agriculture Organization (FAO), there is food security when all people consistently have physical, social, and financial access to adequate, sheltered, and nutritional nourishment that meets their dietary needs and nourishment inclinations for a functioning and solid life. Achieving this goal has proven to be incredibly challenging due to the compassionate situations brought on by ecological disasters, terrible political and ethnic conflicts, etc.

1.4 Food Integrity

Food sharing within the supply chain is a component of food integrity. Every stakeholder (farmer, distributor, processor, and retailer) should provide the most precise information possible on the products' place of origin. This issue is extremely stressful in China, where incredibly rapid development has created real problems with straight imposition. By using blockchain techniques, food associations can reduce food fraud by promptly identifying and connecting outbursts to their unique sources [6].

2 Blockchain for Food Supply Chain

The food supply chain is a combination of more than one stakeholder. The different stakeholders in a food supply chain can be farmers, distributors, shipping companies, retailers, food processors, and finally, the consumer. Figure 4 shows how each stakeholder interacts with blockchain by storing and retrieves information securely [7]. Each stakeholder can verify and monitor the product into supply chain network using blockchain.

Following are essential stakeholders that involve the different phases of supply chain ecosystem.

- Production: Here, the production refers to the food production from the food producers, i.e., the farmers [8]. It represents the agricultural activities and all types of fertilizers, pesticides, and other materials used to produce the food.
- Processing of food: The food processing phase concerns the transformation of the raw food to product, whether it can be partial or complete. The batch code is being processed and can be stored on the block [8].
- Distribution: When the product is ready after packing, it is ready for distribution to the sellers. In addition, there can be possibilities depending on the product that can be stored or the approximate time it will take to reach the seller.
- Retailer: The grocer or the retailers are the buyers from distributors, which provide the food to the end users (customers).

Fig. 4 Food supply chain system with blockchain

- Consumption: The consumer is the last key of this chain, who will buy the food product/item, and can trace the details of the product from the farm to the fork.

Table 1 shows the above-listed stakeholders with their key role and issues using conventional supply chain process. The table also highlights potential solutions from different research papers.

Table 1 highlights following potential of blockchain.

- To address the issue of information traceability since the inception to final delivery of the food [13].
- To provide fine open transparency to all the involving stakeholders to inspect and monitor the food during its entire life cycle.
- Decentralized nature of blockchain brings all the participants on common platforms and enforces the trust management using distributed consensus mechanisms.
- To store information in persistent storage which cannot be modified by anyone**.**

3 Food-Block Framework

Figure 5 presents the proposed framework based on the blockchain for agriculture food supply chain ecosystem. The framework consists of the following main operations to execute the entire supply chain system.

- Registration of all stakeholders such as farmers, manufacturer, distributors, suppliers, and end customers into blockchain network using valid crypto wallet.
- A crypto wallet generates public and private keys. All the stakeholders have their own Account-ID as public key to make any transaction. The private key is kept secure which is used to sign the transaction.
- **Farmers** store their food details into blockchain network after validating and verifying.

Table 1 Food supply chain stakeholders [9–12]

Stakeholder	Role	Issues	Solution by blockchain
Food producer	The creators are farmers, agricultural businesses, and producers of any final food product	How to choose the right fertilizers and treatments, plant their crops, choose the right seeds or other raw materials, decide when to harvest, etc.	Keep track of and record facts that cannot be changed, such as the use of pesticides, fertilizers, and the water cycle Transparency and place
Supplier	Make it possible for items to be transported from fields to storage facilities and processing plants Establish links among merchants, distributors, and other middlemen in the food processing industry	How to move raw materials in an environment that can support food production while moving raw materials quickly to their destination How to guarantee producer, distributor, and retailer transparency	Check the quality of the product while it is in transit by avoiding losses and regularly checking variables like temperature and humidity
Food processor	Large-scale food producers sell their products to private or public businesses, who buy them in their natural state	How to perform manual tasks such as packaging goods for sale The best way to inspect the input product's quality is crucial for this stakeholder	Exhibit a high level of trust and confidence by providing clear and trustworthy details about processed food products, such as the source of the raw ingredients, the input quality, the process parameters, the genuineness, storage conditions, etc. Charge more for its products due to increased visibility and trust. Visibility will also aid in eradicating waste and inefficiencies
Retailer	Retailers are businesses that contact with customers directly, such as retail chains, online merchants, and brick-and-mortar stores	How to examine the products' provenance, quality, and shelf life. If any harm is done to the customers, they may also be subject to legal action	Maintain end-to-end visibility of the products during their full path from the farm to the shelf Address product-related problems, improve chain-wide responsibility, and boost customer happiness
Consumer	The final users of food goods are consumer	How to provide value for money in light of their top priority of receiving the best products possible at all costs	Obtaining clear and dependable access to the cycle of food products allows for a trusted environment and improved brand experience

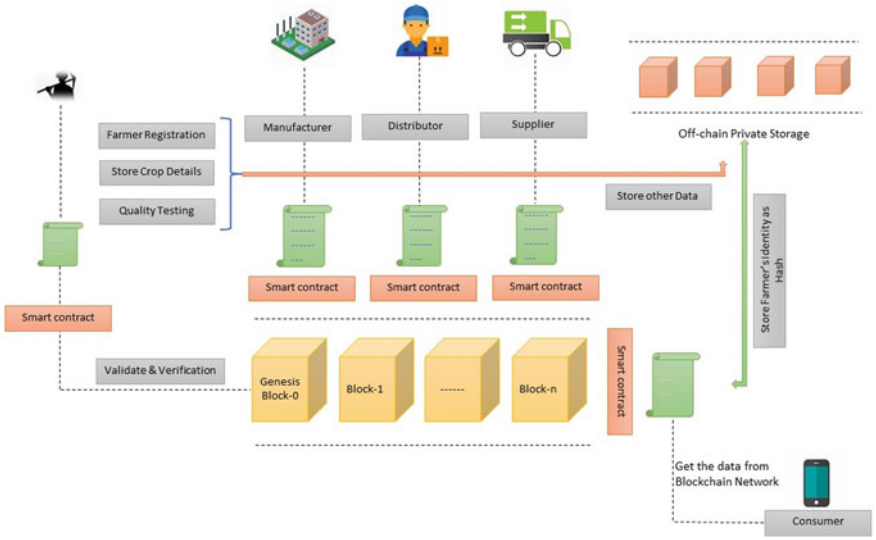


Fig. 5 Food-Block blockchain framework

- **Manufacturer** or food processing industry can see all the details about farmers with their products and process to purchase.
- Manufacturer forms end product and make an entry into blockchain network. All the products will be managed into different lots.
- **Distributors** and **supplier** can see different manufacturers with products and process to buy the products. The distributor and supplier can see the origin of products.
- Manufacturer, distributors and suppliers can see the number of stocks held by any stakeholders into the supply chain system.
- **Retailers** and **customers** can see the entire life cycle of the final product from the blockchain network.

3.1 IoT–Blockchain Integration

The model presented in Fig. 6 handles data generation, analysis, and storing information after quality check into the blockchain network [14]. The information is verified and inspected at each and every stage such as farm fields, transportation, and finally at retailer shop.

The data are captured using IoT devices on the fields. IoT devices bring the atomization using a sensor to monitor soil conditions, crop details, water level, temperature, and past control [15]. These result in better crop analysis and increase productivity of agriculture food. Data analytics and food quality checking are performed using advanced technology artificial intelligence and machine learning [16].

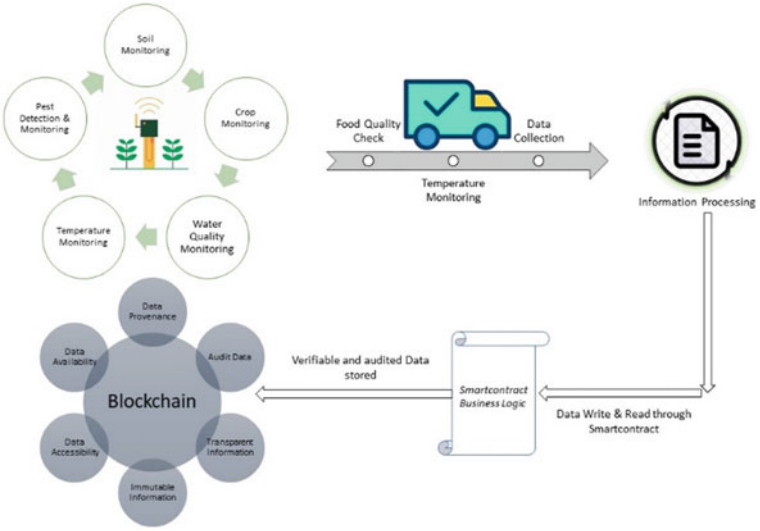


Fig. 6 IoT–blockchain integration

All the information after quality testing is stored and processed through smart contract into blockchain network. Stakeholders such as manufacturer or food processors may check the required information from blockchain. A food inspection is also performed during transportation, especially for those foods that must maintain certain temperature levels using sensor devices [17]. All the data are processed and stored deterministic through a smart contract that enforces rules and regulation of handling data [18]. Blockchain provides a complete transparent view to monitor the condition of food during the entire lifespan of the product.

3.2 Decentralized Application Flow

Figure 7 shows technical flow of application integration with blockchain network. There are main five components of the system.

1. Web browser with wallet integration.
2. Software interface such as UI/UX to interact with blockchain in order to store and retrieve information.
3. Providers are kind of blockchain services offered by third party such as INFURA, QuickNode, and Alchemy to set up blockchain network.
4. Smart contract development and testing.
5. Storing smart contract to Ethereum virtual machine or other platforms such as Polygon and Solana.

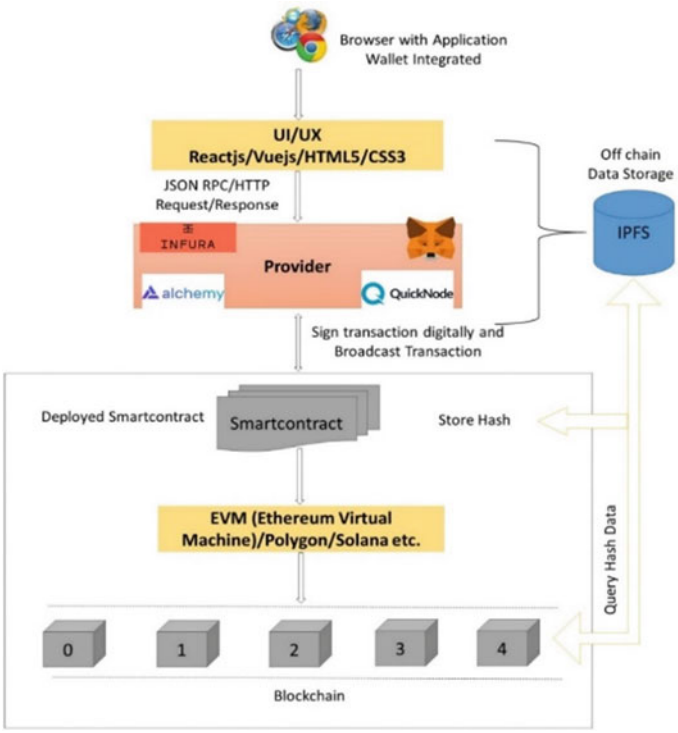


Fig. 7 Application integration with blockchain

Off-chain data storage IPFS is used to store the volume of information in files that returns unique hash. The unique hash will be stored and mapped with a blockchain-distributed ledger. This will minimize storage usage and increase the performance.

3.3 Blockchain Platforms and Electric Load Consumption

Figure 8 builds on the data generated from CCRI Crypto Sustainability API Indices.

The chart provides context for annualized electricity load and presumptively reflects consistent electricity use at the time. The horizontal axis represents different blockchain platforms available to develop decentralized application. The vertical axis represents electric load in megawatt as well as electric consumption in megawatt per hour.

All the data are based on proof-of-stack (PoS) consensus mechanisms. In case of proof-of-work (PoW), electric load and consumption increase drastically as PoW consumes more electric power for mining.

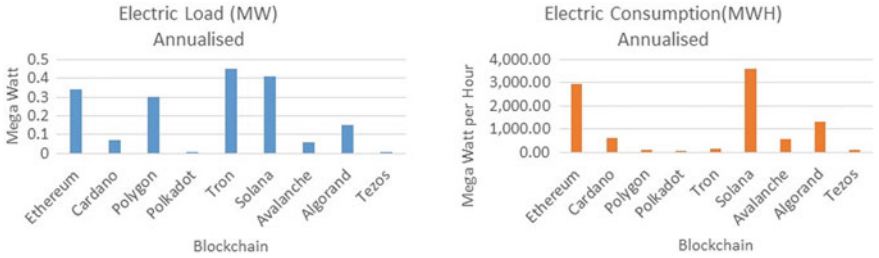


Fig. 8 Proof-of-stack electric load consumption data

4 Conclusion

Blockchain is one of the revolutionary technologies that brings transparency, end-to-end information tracking, and information tracing. Several studies have shown the potential of blockchain for the use cases such as supply chain and agriculture along with IoT integration; however, but none of the work has practical implementation of the exact model or framework to complete technical flow with analysis. The framework Food-Block is tested with five nodes, but it does not have any limit to extend the nodes in the future as the network starts to grow. The stakeholders or the end users can access information stored in blockchain using web or mobile applications. For the better performance, we are not keeping all information in blockchain, but we maintain additional information on the local storage server. In the future work, we will test the entire blockchain network using different blockchain-based platforms and integrate the artificial intelligence (AI) for better food quality analysis.

References

1. Krichen M, Ammi M, Mihoub A, Almutiq M (2022) Blockchain for modern applications: a survey. Sensors 22(14):5274. https://doi.org/10.3390/s22145274
2. Khan MA, Hossain ME, Shahaab A, Khan I (2022) Shrimp Chain: a blockchain-based transparent and traceable framework to enhance the export potentiality of Bangladeshi shrimp. Smart Agric Technol 2:100041. https://doi.org/10.1016/j.atech.2022.100041
3. Guo H, Yu X (2022) A survey on blockchain technology and its security. Blockchain Res Appl 3(2):100067. https://doi.org/10.1016/j.bcra.2022.100067
4. Hang L, Ullah I, Kim D-H (2020) A secure fish farm platform based on blockchain for agriculture data integrity. Comput Electron Agric 170:105251. https://doi.org/10.1016/j.compag.2020.105251
5. Dai H-N, Zheng Z, Zhang Y (2019) Blockchain for internet of things: a survey. IEEE Internet Things J 6(5):8076–8094. https://doi.org/10.1109/JIOT.2019.2920987
6. Kamilaris A, Fonts A, Prenafeta-Boldú FX (2019) The rise of blockchain technology in agriculture and food supply chains. Trends Food Sci Technol 91:640–652. https://doi.org/10.1016/j.tifs.2019.07.034

7. Uddin MA, Stranieri A, Gondal I, Balasubramanian V (2021) A survey on the adoption of blockchain in IoT: challenges and solutions. Blockchain Res Appl 2(2):100006. https://doi.org/10.1016/j.bcra.2021.100006
8. Ronaghi MH (2021) A blockchain maturity model in agricultural supply chain. Inf Process Agric 8(3):398–408. https://doi.org/10.1016/j.inpa.2020.10.004
9. Parmar M, Shah DP (2020) Uplifting blockchain technology for data provenance in supply chain. Int J Adv Sci Technol 29(6):18
10. Sendros A, Drosatos G, Efraimidis PS, Tsirliganis NC (2022) Blockchain applications in agriculture: a scoping review. Appl Sci 12(16):8061. https://doi.org/10.3390/app12168061
11. Demestichas K, Peppes N, Alexakis T, Adamopoulou E (2020) Blockchain in agriculture traceability systems: a review. Appl Sci 10(12):4113. https://doi.org/10.3390/app10124113
12. Xu J, Guo S, Xie D, Yan Y (2020) Blockchain: a new safeguard for agri-foods. Artif Intell Agric 4:153–161. https://doi.org/10.1016/j.aiia.2020.08.002
13. Lin W et al (2020) Blockchain technology in current agricultural systems: from techniques to applications. IEEE Access 8:143920–143937. https://doi.org/10.1109/ACCESS.2020.3014522
14. Torky M, Hassanein AE (2020) Integrating blockchain and the internet of things in precision agriculture: analysis, opportunities, and challenges. Comput Electron Agric 178:105476. https://doi.org/10.1016/j.compag.2020.105476
15. Rodríguez JP, Montoya-Munoz AI, Rodriguez-Pabon C, Hoyos J, Corrales JC (2021) IoT-agro: a smart farming system to Colombian coffee farms. Comput Electron Agric 190:106442. https://doi.org/10.1016/j.compag.2021.106442
16. Jamil F, Ibrahim M, Ullah I, Kim S, Kahng HK, Kim D-H (2022) Optimal smart contract for autonomous greenhouse environment based on IoT blockchain network in agriculture. Comput Electron Agric 192:106573. https://doi.org/10.1016/j.compag.2021.106573
17. Noby DA, Khattab A (2019) A survey of blockchain applications in IoT systems. In: 2019 14th international conference on computer engineering and systems (ICCES), Cairo, Egypt, pp 83–87. https://doi.org/10.1109/ICCES48960.2019.9068170
18. Panarello A, Tapas N, Merlino G, Longo F, Puliafito A (2018) Blockchain and IoT integration: a systematic survey. Sensors 18(8):2575. https://doi.org/10.3390/s18082575

Speech Emotion Recognition Using Machine Learning

Rohini R. Mergu, Rupali J. Shelke, Yash Bagade, Prajjwal Walchale, and Himshruti Yemul

Abstract This paper describes improved research on speech emotion recognition (SER) systems. The definition, classification of the state of emotions and the expressions of emotions are introduced theoretically. In this research article, a SER system based on the CNN classifier and MFCC feature extraction are developed. Mel Frequency Cepstral Coefficients (MFCC) are excerpted from audio signals that are accustomed to train various classifiers. All seven emotions were categorized using a convolutional neural network (CNN). Surrey Audio Visually Expressed Emotion (SAVEE), Ryerson Affective Speech and Song Audiovisual Database (RAVDESS), Toronto Affective Speech Set (TESS), Crowdsourced Affective Multimodal Actor Dataset (CREMA-D) databases were used as experimental datasets. This study shows all four datasets using the CNN classifier. With 1D-CNN, the overall emotion recognition accuracy is 43%, the gender recognition accuracy is 81%, and the gender-neutral recognition accuracy of emotions is 48%. Using 2D-CNN, the overall accuracy rate for emotion recognition is 67.58%, the accuracy rate for gender recognition is 98%, and the accuracy rate for non-gender recognition of emotions is 65%.

Keywords SER · CNN · 1D · 2D · Recognition · Accuracy · Gender · Emotions

1 Introduction

Emotions play a vital role in our interpersonal relationships. Analysing and understanding our emotions helps us make logical and rational decisions. By providing feedback and communicating your feelings, you can align and understand the feelings of others. Human language provides a wealth of information about an individual's mental state and emotions. This has expanded into a new field of study known as "automatic emotion recognition." Its main goal is to know and acquire the conveyed emotions. There are enough applications to detect human emotions when interacting with robotics, voice monitoring, synchronous and virtual learning,

R. R. Mergu (✉) · R. J. Shelke · Y. Bagade · P. Walchale · H. Yemul
Walchand Institute of Technology, Solapur, Maharashtra, India
e-mail: rohinimergu@gmail.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_12

clinical research, business applications, computer games, entertainment, call centres, cardboard systems, banking, and more. In e-learning programming and classroom, focus is on improving the quality of lessons by providing information about students' emotional states. For example, teachers can use SER to identify different strategies to make their lessons enjoyable.

A successful SER system needs to address three key issues.

(1) A good emotional audio dataset
(2) Extraction of effective features
(3) Reliable classifier design.

2 Related Work

Researchers have developed various methods to detect emotions in speech signals. Many studies have proposed traditional speech recognition methods [1, 2]. First, this research has been done in terms of feature extraction, which are separated by temporal and spectral features. Time-domain features such as velocity, zero-crossing rate, amplitude, signal energy, maximum energy, and minimum amplitude can be accessed from the temporal feature domain. Some studies use temporal pitch information of speech signals to perform emotion recognition with temporal features [3]. Another study considered speech rate and Mel frequency coefficients (MFCCs) [4]. Many studies, those mentioned above, focus on the spectral features alternate to the temporal features. The MFCC is one of the spectral characteristics extricated from the audio signal to categorize emotions as happy, anger, and sadness [5]. A few researchers have put forth an emotion detection that assorts emotions with a spectrograph as a characteristic [6]. Recent studies are carried out with transfer-learning and skill using VGG-19 patterns [7] trained using ImageNet [8], according to spectral features like Spectrogram [9].

A 2D-CNN speech emotion recognition is shown using MFCC, one of the spectral features to perform speech emotion recognition on deep learning models incorporating spectral features. We use 2D-CNN with RADVESS, CREMA-D, SAVEE, TESS [10] datasets to efficiently classify sentiment. Convolution operations in one direction is performed by1D-CNN only; by contrast, the convection core of 2D-CNN operates in two dimensions. It can also be used to analyse time series data [11]. A CNN architecture consists of three main building blocks as first a convolutional layer, second a pooling layer, and finally a fully connected layer [12].

3 The Proposed Method

The entire system is shown in Fig. 1. First, select four datasets. These datasets are SAVEE, TESS, CREMA-D, RADVESS. The target audio signal data is then provided to the feature extraction block. The feature extraction block produces MFCC

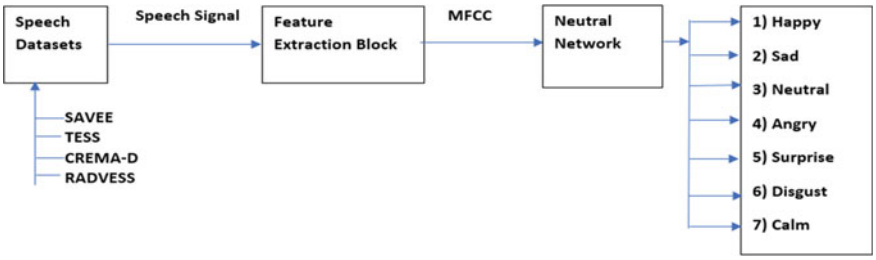


Fig.1 Block diagram

features with components. The model classifies emotions (happiness, anger, calm, fear, sadness, surprise, disgusting, neutral) based on the generated MFCC function as data enters the neural network model.

3.1 MFCC

MFCC as spectral function for speech emotion recognition system is used. MFCCs are generated using cepstral analysis on Logmel spectrograms. Cepstral analysis is an analysis technique that uses a scaled log spectrum with an inverse Fourier transform. This analysis assists to identify a clear sound of the speaker at fundamental and specific frequency. A spectrogram through short-time Fourier transform is been converted from the raw audio data and a Mel filter bank is enforced to render the spectral image intelligible to the human inner cochlea. This is due the human aural set-up is non-linear, subtlety to low-frequencies and coarseness to high-frequencies. Then MFCC through discrete-cosine transform is obtained using cepstral scale analysis. The spectrum coverage is defined using MFCC. This helps to recognize the sound and pitch structure of the vocal track, therefore it noted as a SER feature. We used LIBROSA python library to extract MFCC. Figs. 2 and 3 shows some sample Mel spectrograms as Mel-spectrogram of male happy voice and female happy voice.

3.2 CNN

Convolutional neural networks solve the problem of mislayed spatial and local information due to reduction in dimensionality, which is a common problem of fully connected networks to process data with spatial images. It usually comprises of a convolution layer that performs convolution operations on the input features. The 1D-CNN only performs convolutions in one direction, whereas the 2D-CNN convolution kernel performs them in two dimensions. It can also be used to analyse time series data. Table 1 shows SER accuracy for 1D and 2D-CNN for 7 emotions.

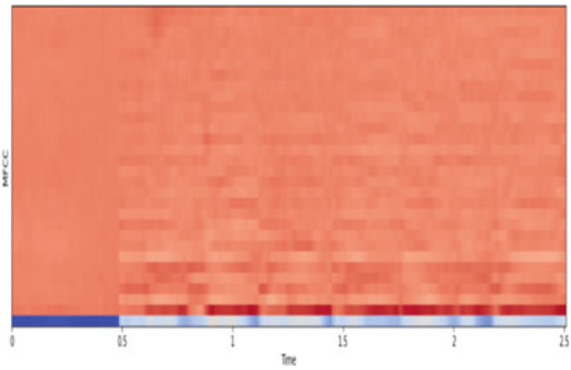


Fig. 2 Mel-spectrogram of male

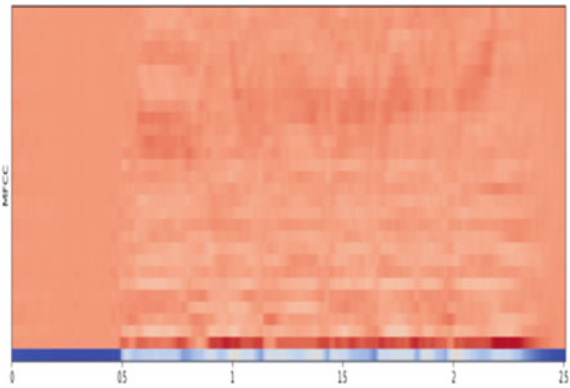


Fig. 3 Mel-spectrogram of female

Table 1 SER accuracy using 1D-CNN and 2D-CNN for 7 emotions

Model	Accuracy (%)		Accuracy
1D-CNN	43	**2D-CNN**	**67.58**
(a) Gender recognition	81	(a) Gender recognition	98
(b) Emotion recognition without gender	48	(b) Emotion recognition without gender	65

Our 2D MFCC CNN uses the function as an input, represented by a 2D matrix. It consists 4 convolutional blocks, encompasses a convolutional layer, a combination layer. The input size of the model is ($N \times 30 \times 216 \times 1$), here N is used to represent the number of sample input data, 30 represents feature measurements count, 216 represents number of frames/sample and 1 is the channel depth/sample. Some studies have shown that up to 4 convolutional layers, neural network performance increases with the number of layers; however, performance drops dramatically beyond 4 layers [13]. So, we chose to use 5 convolutional classes. Each convolutional table has a kernel that extracts local resources. The size of all kernels is (4×10), and the backup

function of each class uses the RelU function. The number of pieces is reduced by the amount of spring water (2 × 2) at the end of level. To prevent overfitting, the normal operation and discarding of each block was performed; the abandonment rate was 0.2, where 20% of the neurons were not isolated from training for each type of procedure.

4 Result

In this experiment, we used four datasets, namely SAVEE [14], TESS [15], CREMA-D [16], and RADVESS [17]. Most of these samples last 2 to 4 s. So, we decided that the relevant information would initially be available within 2.5–3 s, so we only sampled the data within 2.5 s. In this way, messages over 2.5 s are interrupted, and messages under 2.5 s are filled with zeros. The sampling information is given by: sampling rate = 44.4 kHz, sampling type = kaiser_fast. A new model is introduced and experimentally analysed it using the Keras [18] and also Tensorflow frameworks are used for analysis [19]. GOOGLE COLAB works in the GPU runtime environment. The LIBROSA [20] library is used to analyse audio data and extract MFCC. Category cross entropy is used as a loss function to train the model. ADAM with learning rate 0.001 is used as the optimizer. During the learning process, training is performed exactly as a metric. Batch sizes are considered 571, 20 epochs.

The number of training data is 9121 and 3040 samples for validation. Divided into seven emotions. Each is divided into 14 gender-separated classes, and after training for 50 epochs, 2D-CNN representation achieves 67.58% accuracy on the test set. Initially training set and test sets were separate. Therefore, the model shows high accuracy on new data is used. Figure 3 shows the accuracy plots and confusion matrices for all experimental models. The blue coloured and orange coloured lines show the training and test set accuracy. For each axis, the horizontal and vertical refer to the CNN model, especially the 2D-CNN model, which showed the highest accuracy on the test set against 1D-CNN model.

The two network models 1D-CNN and 2D-CNN are tested as shown in Figs. 4 and 8. We achieve an accuracy of 43% for 1D-CNN and 67.58% for 2D-CNN as shown in Figs. 12 and 13. The Emotional Speech Recognition System also demonstrated gender accuracy with 1D-CNN and 2D-CNN achieving 81% and 98% accuracy, respectively, shown in Figs. 6 and 10. Shows below the confusion matrix with the accuracy of each label correctly predicted by the model. Figures 5, 7, 9 and 11 shows accuracy confusion matrix. The comparison of our work with existing literature is shown in Table 2.

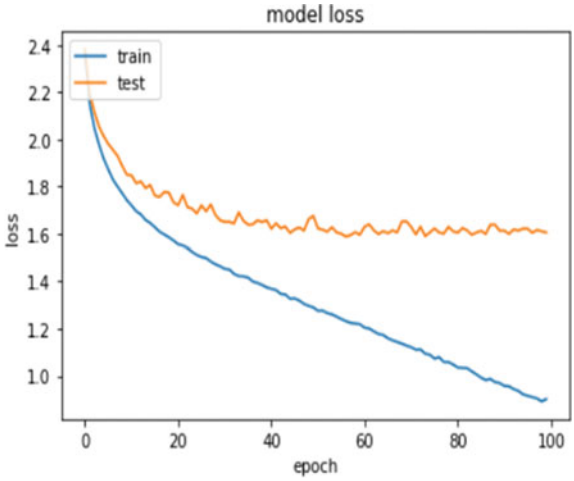


Fig. 4 1D-CNN train and test graph

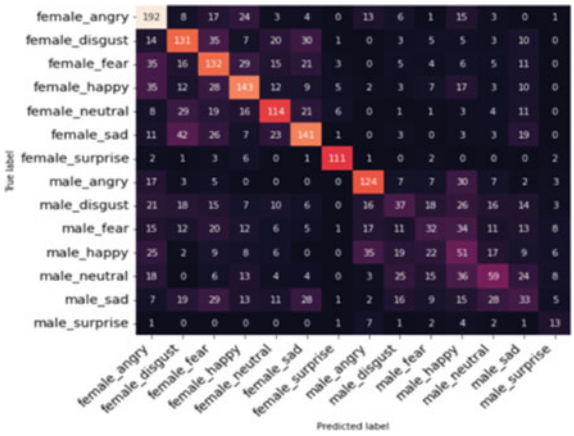


Fig. 5 1D-CNN accuracy confusion matrix

5 Conclusion

The proposed system examines the accuracy and the performance distribution of Speech Emotion Recognition using methods of the 1D-CNN and 2D-CNN. For the SAVEE, RADVESS, TESS, CREMA-D datasets, against 1D-CNN method, the 2D-CNN method with the MFCC showed high accuracy in emotion recognition for all types of emotions with an average of 67.58% accuracy, showing the best performance.

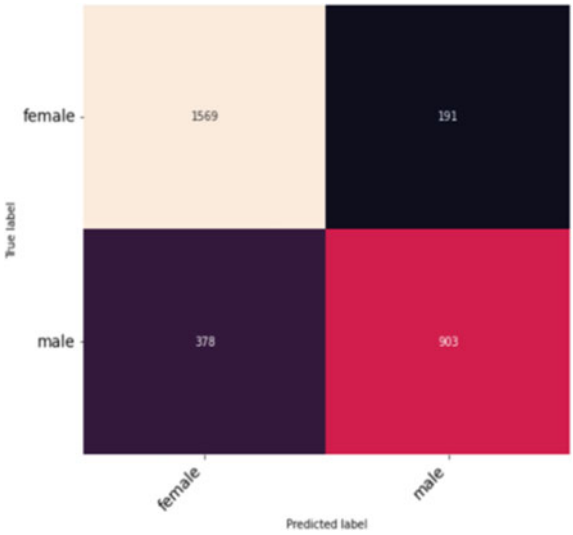


Fig. 6 1D-CNN gender accuracy

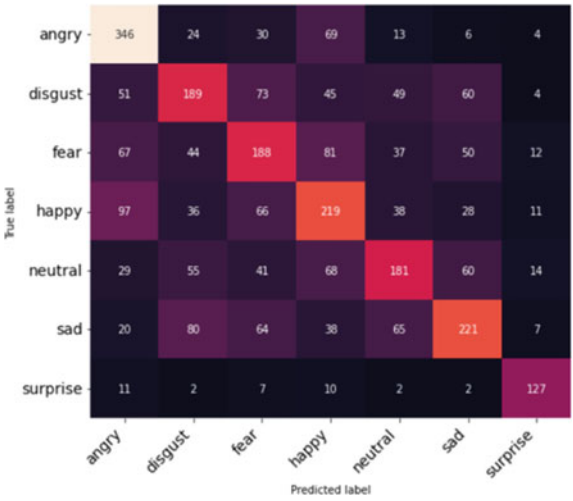


Fig. 7 1D-CNN emotion accuracy

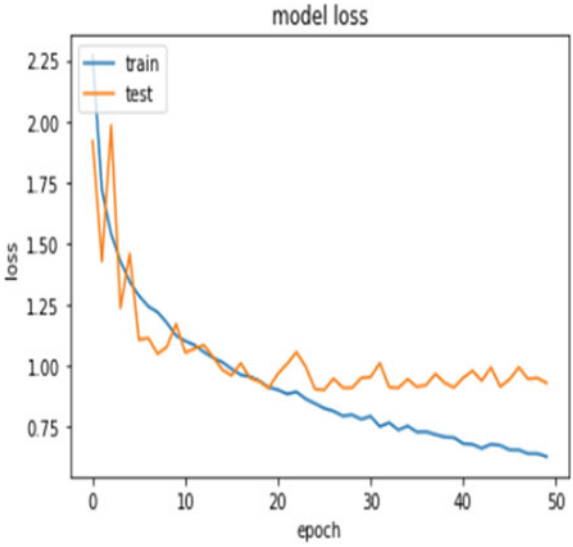


Fig. 8 2D-CNN train and test graph

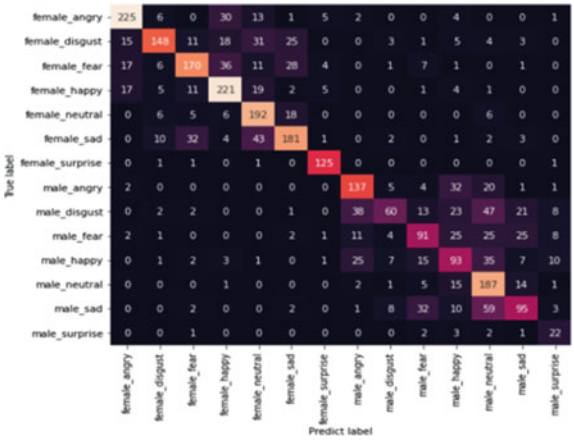


Fig. 9 2D-CNN accuracy confusion matrix

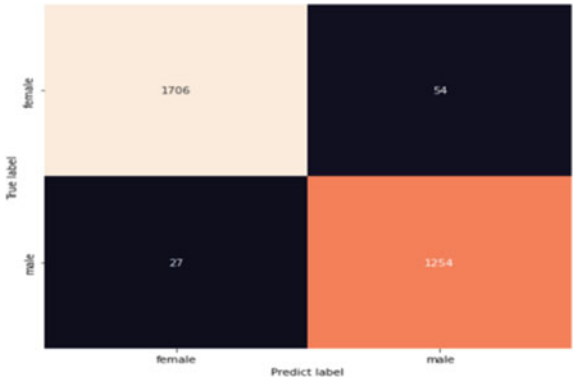


Fig. 10 2D-CNN gender accuracy

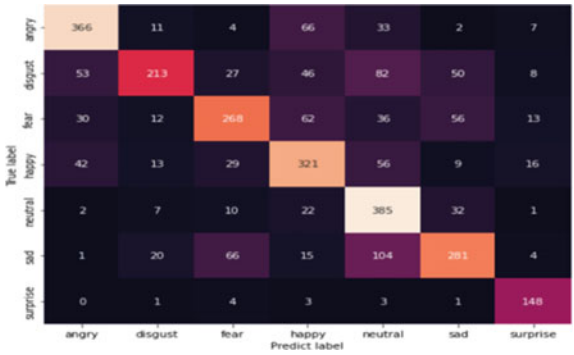


Fig. 11 2D-CNN emotion accuracy

	precision	recall	f1-score	support
female_angry	0.48	0.67	0.56	287
female_disgust	0.45	0.50	0.47	264
female_fear	0.38	0.47	0.42	282
female_happy	0.50	0.50	0.50	286
female_neutral	0.51	0.49	0.50	233
female_sad	0.52	0.51	0.51	279
female_surprise	0.85	0.86	0.86	129
male_angry	0.56	0.60	0.58	205
male_disgust	0.27	0.18	0.22	207
male_fear	0.26	0.16	0.20	197
male_happy	0.21	0.24	0.22	209
male_neutral	0.37	0.27	0.31	215
male_sad	0.21	0.15	0.18	216
male_surprise	0.27	0.41	0.32	32
accuracy			0.43	3041
macro avg	0.42	0.43	0.42	3041
weighted avg	0.42	0.43	0.42	3041

Fig. 12 1D-CNN result

Fig. 13 2D-CNN result

	precision	recall	f1-score	support
female_angry	0.84	0.84	0.84	287
female_disgust	0.74	0.71	0.72	264
female_fear	0.76	0.62	0.68	282
female_happy	0.78	0.69	0.73	286
female_neutral	0.76	0.77	0.77	233
female_sad	0.60	0.78	0.68	279
female_surprise	0.91	0.97	0.94	129
male_angry	0.70	0.71	0.70	202
male_disgust	0.57	0.61	0.59	215
male_fear	0.47	0.39	0.43	195
male_happy	0.51	0.47	0.49	200
male_neutral	0.65	0.58	0.61	226
male_sad	0.47	0.57	0.51	212
male_surprise	0.53	0.84	0.65	31
accuracy			0.67	3041
macro avg	0.66	0.68	0.67	3041
weighted avg	0.68	0.67	0.67	3041

Table 2 Comparison of our work with existing literature

Reference	Emotions	Accuracy percentage	Year	Technique/model
Seunghyun Yoon	4	68.8–71.8	2018	2D-CNN
Aharon Satt	4	66	2017	2D-CNN
Gaurav Sahu	6	56.6	2019	2D-CNN
Andrew Maas	–	61.2	2015	1D-CNN
Our work	**8**	67.58	**2022**	**2D-CNN**

References

1. New TL, Foo SW, De Silva LC (2003) Speech emotion recognition using hidden Markov models. Speech Commun 41(4): 603–623
2. El Ayadi M, Kamel MS, Karray F (2011) Survey on speech emotion recognition: features, classification schemes, and databases. Pattern Recognit 44(3):572–587
3. Byun S, Lee S (2016) Emotion recognition using tone and tempo based on voice for IoT. Trans Korean Inst Electr Eng 65(1):116–121. https://doi.org/10.5370/kiee.2016.65.1.116.
4. Hong I, Ko Y, Kim Y, Shin H (2019) A study on the emotional feature composed of the mel-frequency cepstral coefficient and the speech speed. J Comput Sci Eng 13(4):131–140. https://doi.org/10.5626/JCSE.2019.13.4.131
5. Likitha MS, Gupta SRR, Hasitha K, Raju AU (2017) Speech based human emotion recognition using MFCC. In: 2017 International conference on wireless communications, signal processing and networking (WiSPNET), pp 2257–2260. https://doi.org/10.1109/WiSPNET.2017.8300161
6. Park S, Kim D, Kwon S, Park N (2018) Speech emotion recognition based on CNN using spectrogram. In: Information and control symposium, pp 240–241
7. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556

8. Deng J, Dong W, Socher R, Li L, Li K, Fei-Fei L (2009) Image-Net: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition (CVPR), pp 248–255. https://doi.org/10.1109/CVPR.2009.5206848
9. Lee J, Yoon U, Jo G (2020) CNN-based speech emotion recognition model applying transfer learning and attention mechanism. J KIISE 47(7):665–673. https://doi.org/10.5626/JOK.2020.47.7.665
10. Livingstone SR, Russo FA (2018) The Ryerson audio-visual database of emotional speech and song (RAVDESS): a dynamic, multimodal set of facial and vocal expressions in North American English. PLoS One 13(5):e0196391. https://doi.org/10.1371/journal.pone.0196391
11. Tang W, Long G, Liu L, Zhou T, Jiang J, Blumenstein M (2020) Rethinking 1D-CNN for time series classification: a stronger baseline. arXiv: 2002.10061
12. Mao Q, Dong M, Huang Z, Zhan Y (2014) Learning salient features for speech emotion recognition using convolutional neural networks. IEEE Trans Multimed 16:2203–2213. https://doi.org/10.1109/TMM.2014.2360798
13. Huang L, Dong J, Zhou D, Zhang Q (2020) Speech emotion recognition based on three-channel feature fusion of CNN and BiLSTM. In: 2020 the 4th international conference on innovation in artificial intelligence (ICIAI), pp 52–58. https://doi.org/10.1145/3390557.3394317
14. https://www.kaggle.com
15. https://www.kaggle.com/datasets/barelydedicated/savee-database
16. Yoon S, Byun S, Jung K (2018) Multimodal speech emotion recognition using audio and text. In: 2018 IEEE spoken language technology workshop (SLT). https://doi.org/10.1109/SLT.2018.8639583
17. https://www.kaggle.com/datasets/ejlok1/toronto-emotional-speech-set-tess
18. https://www.kaggle.com/datasets/ejlok1/cremad
19. https://www.kaggle.com/datasets/uwrfkaggler/ravdess-emotional-speech-audio
20. https://keras.io/getting_started/

WhatsApp Chat Analyser

Puja Chavan, Priyanka Ramteke, Prashil Ramteke, Suyog Patil, Prashant Raut, and Pranav Shetkar

Abstract Among all communication channels, WhatsApp has been the most popular and effective. Numerous individual and group conversations make up this. As a result, they might contain some unstated facts. In this initiative, the data from those talks is thoroughly analysed. Regardless of the subject, this tool offers a thorough study of the information that WhatsApp provides. Matplotlib, seaborn, Pandas, word cloud, URL extract, emoji, collection, NLTK libraries, among others were used in its construction. Graphs can be efficiently plotted using them, as well as data frames. It is possible to apply this library to the greatest data set because it is used to build data frames and plot various graphs. Inspecting and modelling data is a technique that aims to find some valuable information before drawing certain conclusions. We have both sentiment and chat analysis in WhatsApp Chat Analyser. Also group and individual analysis can be done. Above that we have used 12h format according to the latest update of WhatsApp.

Keywords WhatsApp chat · Sentiment · Matplotlib · Seaborn · Pandas · Word cloud · URL extract · Emoji · NLTK

P. Chavan · P. Ramteke (✉) · P. Ramteke · S. Patil · P. Raut · P. Shetkar
Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India
e-mail: ramteke.priyanka21@vit.edu

P. Chavan
e-mail: puja.cholke@vit.edu

P. Ramteke
e-mail: prashil.ramteke21@vit.edu

S. Patil
e-mail: suyog.patil21@vit.edu

P. Raut
e-mail: mahadeo.prashant21@vit.edu

P. Shetkar
e-mail: prakash.pranav21@vit.edu

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_13

1 Introduction

With 2 billion users, across 180 countries, WhatsApp has established itself as the most widely used social networking site. Over 55 billion messages are sent each day via WhatsApp [1]. Because it enables significant communication, emotional expression and data sharing, WhatsApp is essential to people's lives. As a result, it is frequently asserted that WhatsApp conversations are frequently used to record and display an individual's moods and thought processes. The process of separating and extracting subjective information from language, such as angle and opinions, is known as sentiment analysis.

Prior to recently, the primary focus of sentiment analysis was opinion polarity or categorizing text as positive, negative or neutral. WhatsApp may play a significant role in people's lives because it enables the sharing of emotions through in-depth conversations in addition to facts. In Chennai, India, a survey on students was carried out to find out how popular WhatsApp is with young people. According to the findings, participants spend sixteen hours online and use WhatsApp for a certain amount of time each day [2]. As a result, WhatsApp chats are frequently used to convey an individual's emotions, as well as to record their thoughts and feelings neutral.

2 Literature Review

According to the research done by the authors of the 'Survey Analysis on the usage and Impact of WhatsApp Messenger', writers analysis of WhatsApp users, ranging in age from 18 to 50, shown that around 79% of participants use WhatsApp at least 15 to 60 min a day [3]. An additional study by the researchers behind 'Impact of WhatsApp on Youth: A Sociological study' 100 random WhatsApp users in the age between 18 and 30 in the city of Agra [4]. According to their research, 63% of consumers utilized WhatsApp frequently around 50 times a day and 21% users used it regularly, about 20 times per day, while the remaining 16% used it 100 times per day or more frequently. This study's findings suggest that WhatsApp has a significant influence on how young people communicate [5].

Data mining is the process of evaluating data from a broad perspective and distilling it into valuable information. It is the process of reviewing data from several angles and condensing it into meaningful information using operations like clustering, classification, associative relative technique and other similar ones. The characteristics that could contribute to this addiction are identified through the categorization procedure [6].

User-to-user communication, broadcast messaging and group conversations are just a few of the different types of communication that WhatsApp offers its users. Users can communicate by exchanging simple text messages, multimedia items (including audio, video and image files), contact cards and geographical data [7].

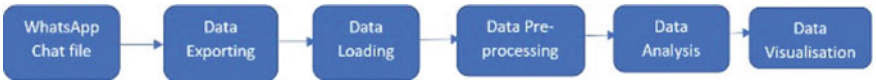


Fig. 1 Block diagram of proposed system

Whatsapp Chat Analyser [7, 8] describes the web-based service WhatsApp chat Analyser, which can analyse WhatsApp chat communication by receiving chat records. In order to retrieve a text-based version of the conversation histories, messages are transmitted via the internet with end-to-end encryption, and chat histories are stored in an encrypted database on the mobile device; thus, this is currently the only way to access chat data. Timestamps, anonymous user names and message kinds are collected from chat histories.

3 Methodology

3.1 Data Analysis

Cleaning, processing, examining and modelling data is a procedure that aims to find some helpful information before pointing to certain conclusions. Figure 1 shows block diagram of proposed system. Data analysis is the act of taking raw data and turning it into information that users can use to make decisions. A simple statistical analysis of WhatsApp communication is provided by this project. These are the findings of the analysis.

- To determine the total number of messages, words, files and links shared during WhatsApp session.
- To identify the group members who are the most active.
- To determine which group used the greatest number of emojis.
- To determine the week's busiest day.
- To identify the most frequently occurring and frequently used terms in a conversation between two people or a group.
- To determine the daily talk frequency.
- To identify the users who are most supportive, most neutral and most critical.

3.2 Tools Used

- Conducted all data analysis using the Python programming language.
- Jupyter notebook for iterative code execution.
- The use of PyCharm as an IDE.

- Constructed an app framework using Streamlit. The quickest way to create and deploy web applications is with Streamlit.

3.3 Implementation

- Created a new project in PyCharm.
- And then created 3 files in that project and named them as following bullet point pre-processor. py: The first step is data pre-processing. That is because it leads to data sets, that are cleaner, coherent and much more manageable, a must for any business trying to get valuable insights from the collected data [9].
- The input we take is a text file consisting of the chats (group chat or personal), so we need to do some pre-processing and get a data frame out of it. Here, the data is structured and segregated into different fields for the date, time, user's name and message. Also from date, year, month, day, hour and minute are extracted.

A string of characters known as a 'Regular Expression', or RegEx, stands for a search pattern. RegEx can be used to determine whether a string includes a given search pattern [10]. WhatsApp Chat data has date and time of a chat. It contains information that the model can use, but because date and time are available in so many different formats, working with such data can be challenging. RegEx must be used in this situation to deal with the various formats.

- app.py: This is the main file where we create the Streamlit app.
- helper.py: We made some functions in this file that are used in the main app.py file.

4 Result and Discussion

In order to train our machine learning classification model on the acquired data, it is read and processed. Following the evaluation, the model is serialized. Figure 2 shows user interface of WhatsApp Chat Analyser Web App which was framed using seaborn library.

4.1 Analysis Performed

1. Highest Statistics: Fig. 3 includes total messages, total words, shared material and shared links.
2. Activity Map: In Fig. 4, the busiest and least busy days are displayed.
3. Daily Timeline: In Fig. 6, it shows frequency of chat in a day.

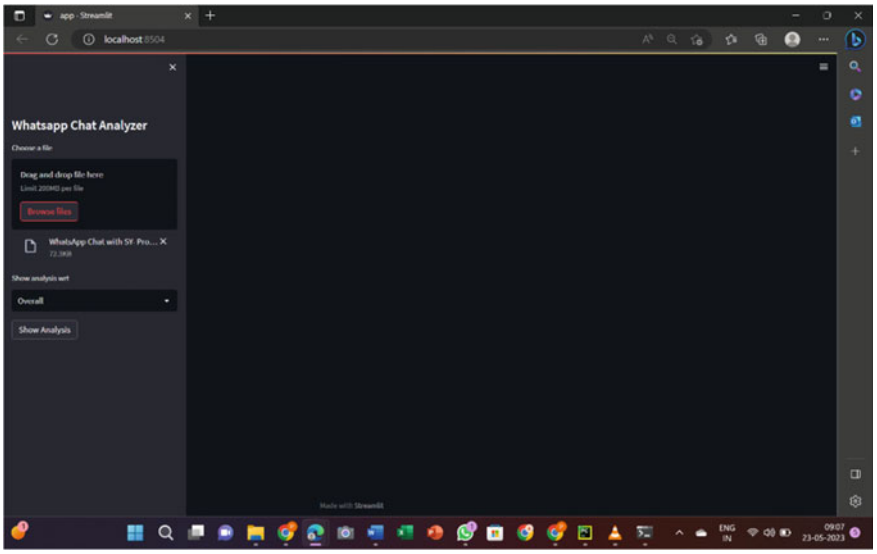


Fig. 2 User interface of app

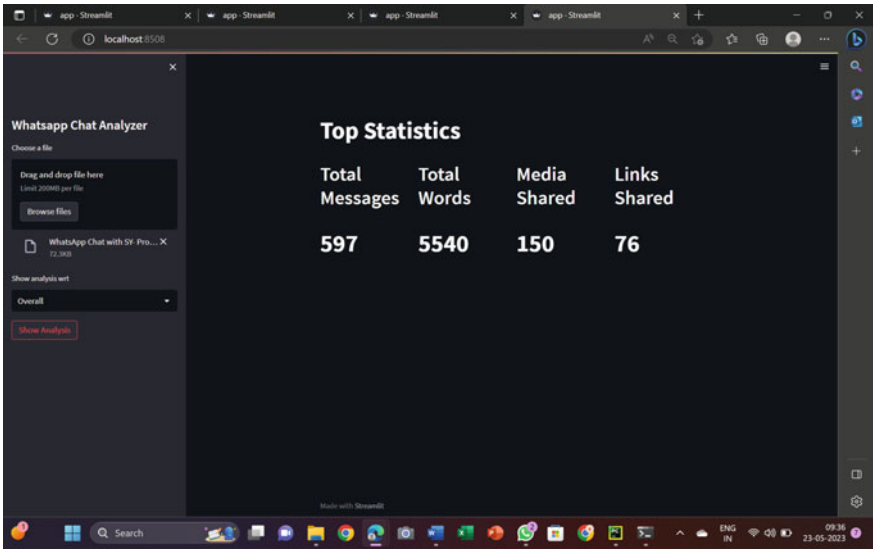
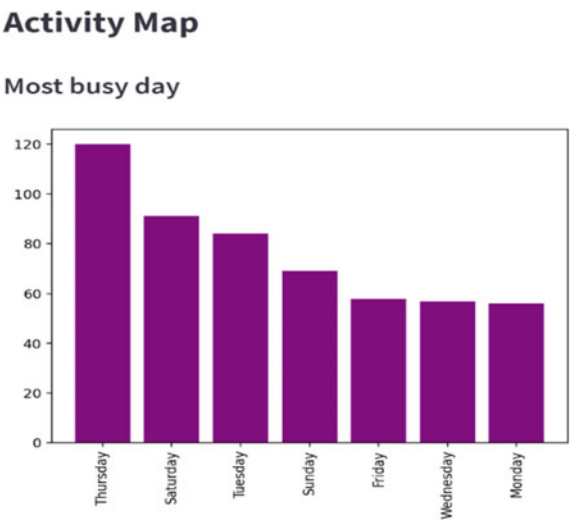


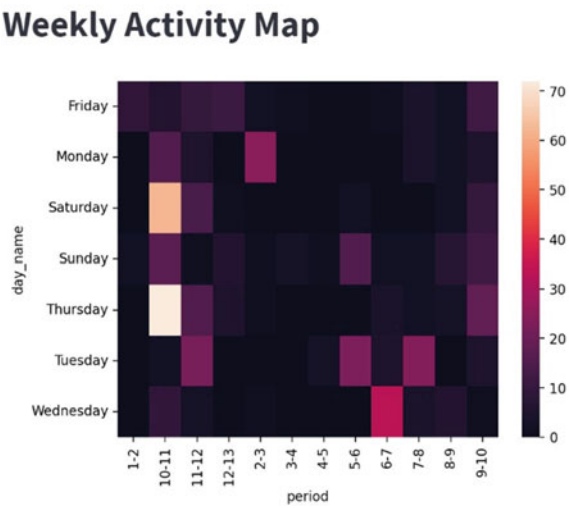
Fig. 3 Top statistics

Fig. 4 Activity map



4. Weekly Activity Map: In Fig. 5, it shows heat map where time range of chats and what time which day the chatting happens the most, least and everything in between.
5. Word Cloud: In Fig. 8, it shows most commonly and frequently used word. Here, frequency of chat is directly proportional to font size.
6. Most Busy Users: In Fig. 7, it shows most active people. Figure 9 shows probability of most common word used.
7. Emoji Analysis: In Fig. 10, it shows most commonly and frequently used emojis**.** The number of emojis used in an online communication is negatively connected

Fig. 5 Weekly activity



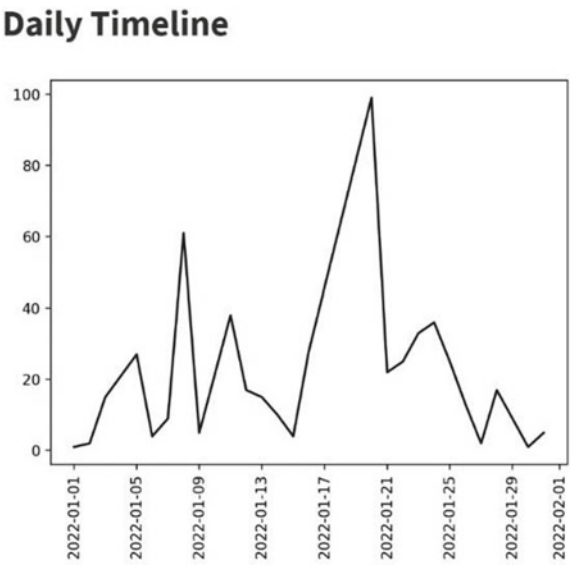
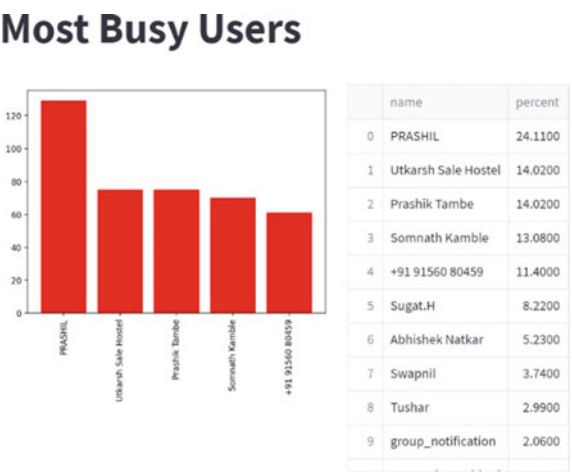


Fig. 6 Daily timeline



	name	percent
0	PRASHIL	24.1100
1	Utkarsh Sale Hostel	14.0200
2	Prashik Tambe	14.0200
3	Somnath Kamble	13.0800
4	+91 91560 80459	11.4000
5	Sugat.H	8.2200
6	Abhishek Natkar	5.2300
7	Swapnil	3.7400
8	Tushar	2.9900
9	group_notification	2.0600

Fig. 7 Mostly busy user

to how professional the participants are [11]. Emojis are frequently utilized as a substitute for words when feelings cannot be expressed with simple words, and people feel more relaxed employing them instead.

8. WhatsApp Chat Sentiment Analysis: In Fig. 11, it shows most positive users, most neutral users and most negative users. A sentiment is an attitude, belief or conclusion brought on by a sensation. Sentiment analysis, commonly referred to as opinion mining, examines how individuals feel about particular things [12]. According to research, introverts prefer texting rather than phone calls as a form of communication, and thus, they text more frequently [13].

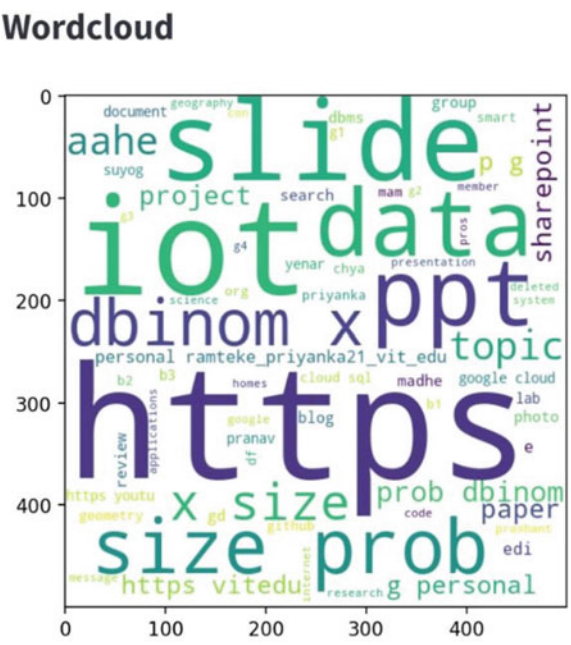


Fig. 8 Word cloud

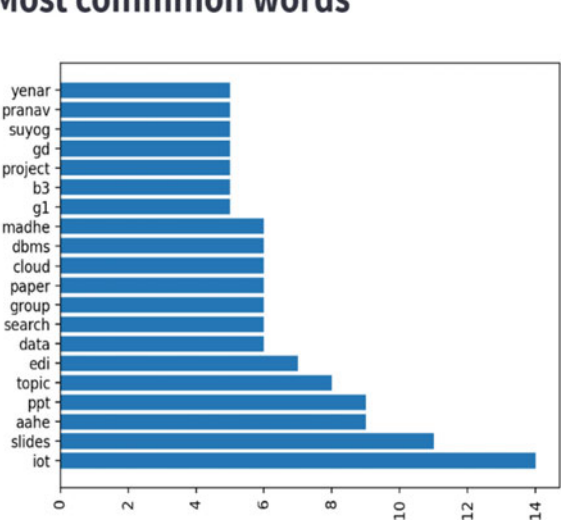


Fig. 9 Most common words

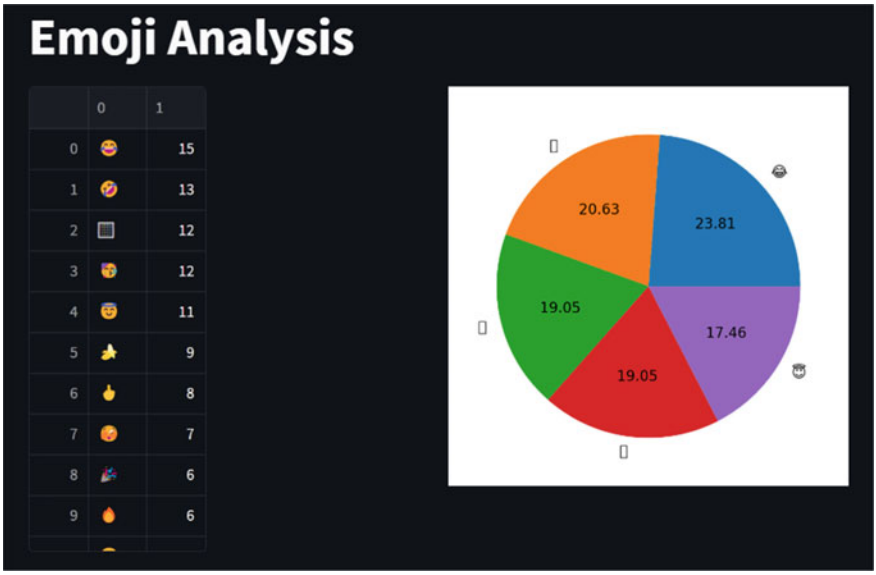


Fig. 10 Emoji analysis

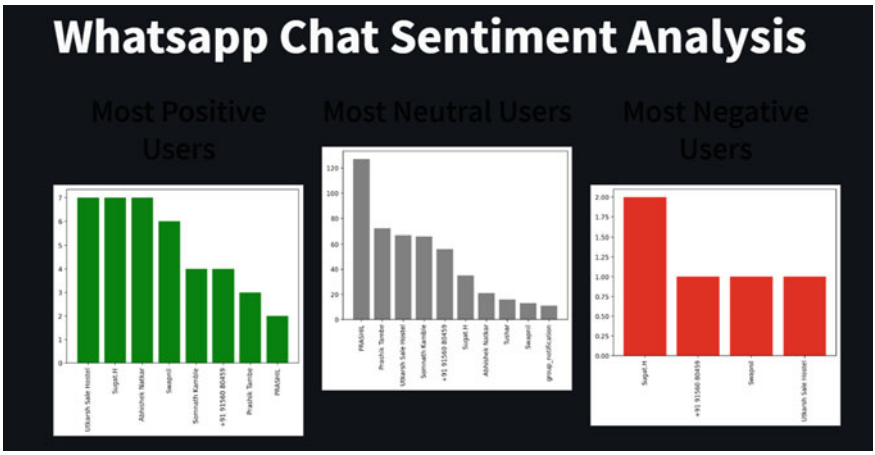


Fig. 11 Sentiment analysis

Proposed system shows overall analysis of group chat and also analyse the chat of any particular user for that click on drop down menu, it will show all group members. You can select any member as per your choice.

5 Conclusion

In this paper, we have discussed WhatsApp application and its libraries, to analyse WhatsApp group chat and visually display the top 10 and top 20 chat group members. A plot was provided for both individual and group chat. When the job was completed, the anticipated outcomes were attained, and the analysis was able to demonstrate the level of engagement (daily and weekly), time spent according to 12hour format, sentiment analysis, most busy user in the group, emoji analysis and most common word.

References

1. Andjelic J (2019) Retrieved from https://fortunly.com/statistics/whatsapp- statistics#gref
2. Ka Jebakumar J (2014) Whatsapp: A trend setter in mobile communication among chennai youth. IOSR J Human Soc Sci. 19:01-06. https://doi.org/10.9790/0837-19970106
3. Kumar N, Sharma S (2017) Survey analysis on the usage and impact of Whatsapp Messenger. Glob J Enterp Inf Syst. 852. https://doi.org/10.18311/gjeis/2016/15741
4. Bhatt MA, Arshad DM Impact of Whatsapp on youth: a sociological study. IRA-Int J Manag Soc Sci
5. Dahiya S, Mohta A, Jain A (2020) Text classification based behavioural analysis of Whatsapp chats. In: Proceedings of the fifth international conference on communication and electronics systems (ICCES 2020) IEEE conference record # 48766; IEEE Xplore ISBN: 978-1-7281–5371-1
6. Radha D, Jayaparvathy R, Yamini D (2016) Analysis on social media addiction using data mining technique, Int J Comput Appl. 139(7):23–26 (0975–8887)
7. Anglano C (2014) Forensic analysis of whatsapp messenger on android smartphones. Digital Invest J. 11(3):201–213
8. Ravishankara K, Dhanush V, Srajan IS (2020) Whatsapp chat analyzer. Int J Eng Res Technol (IJERT). 9(05) ISSN: 2278-0181
9. Mohta A, Jain A, Saluja A, Dahiya S (2020) Pre-processing and emoji classification of WhatsApp chats for sentiment analysis. In: 2020 fourth international conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I- SMAC) https://doi.org/10.1109/i-smac49090.2020.9243443
10. Regex Retrieved from https://regexr.com/
11. Hadi H, Bakir NS (2019) The use of emoticons among university students: a pragmatic study. Zanco J Hum Sci 23. https://doi.org/10.21271/zjhs.23.1.19
12. Fang X, Zhan J (2015) Sentiment analysis using product review data. J Big Data 2(1):1–14
13. Zhou Z, Xu K, Zhao J (2018) Extroverts tweet differently from introverts in Weibo. EPJ Data Sci 7:18. https://doi.org/10.1140/epjds/s13688-018-0146-8

Real-Time Suppression of Non-stationary Noise for Web-Based Calling Applications

Radha Senthilkumar, T. V. Raghavasimhan, R. S. Tejeshwini, C. Kavipriya, and P. Jayanthi

Abstract Background noise affects the quality of audio for the far-end user. This noise must be suppressed without degrading the speaker's voice and must be heard by the far-end user. Non-stationary noise from the background affects the quality of call audio for the far-end user as it evades detection and suppression by traditional algorithms. Utilizing deep learning-based techniques increases the computation time and resources, which can be partially mitigated by the use of cloud computing, however at the cost of users' privacy. This work focuses on improving portability by realizing real-time noise suppression in the sender-side browser. It utilizes client-side JavaScript frameworks and libraries on the sender's end to process and deliver noise-suppressed audio for further transmission to the receiver's end through WebRTC for use in audio-video conferencing applications. The proposed work tackles the challenges of performing the noise suppression in the client-side with a very high PESQ of 4.64 and STOI of 0.99 with a real-time factor 2.06.

Keywords Noise suppression · Latency · Portability · Conferencing · DEMUCS · PESQ · STOI · Non-stationary noise

1 Introduction

The effects of background noise during phone calls has been identified and studied for a significant amount of time. Noise suppression aims to reduce the effect of background noise and make the signal of interest more prominent. This noise could be stationary and non-stationary noise. Stationary noise refers to noise that does not change with time, such as a fan spinning in the caller's background. The other variety involves random, sudden noises that affect the perceptibility of the audio. Most existing techniques are focused only on the stationary noise and the techniques used for non-stationary noise consume more resources for computation. Traditionally, statistical behavior of the noise was used to estimate its noise distribution, which

R. Senthilkumar · T. V. Raghavasimhan · R. S. Tejeshwini · C. Kavipriya · P. Jayanthi (✉)
Department of Information Technology, Anna University, Chennai 600044, India
e-mail: pjayanthi@mitindia.edu

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_14

could then be exploited for the removal of similar components from the actual noisy audio. This produces quick results owing to the computational simplicity of the methods involved, namely Wiener filtering and spectral subtraction. The efficacy of these techniques against stationary noise is well-known; however, they falter against non-stationary noise. The major advantage of these techniques is still the low latency of the process. Mobile phones utilize traditional digital signal processing algorithms to reduce noise using hardware and software techniques which impose constraints on the hardware of the device. Phones utilize a dual-microphone system, which is applicable only for specific devices that can exploit a "candy-bar" or a "touch-bar" form factor. This makes it unsuitable for smart devices like TVs, Amazon Echo, etc. Services like Google Meet send the user's audio data to a cloud server where it is processed to suppress noise using machine learning algorithms on customized, proprietary hardware designed and developed for this specific purpose. The cost and scale of operation cannot be replicated by other service providers. There are also valid concerns about the privacy of the call data that is being transferred to a server [6]. Furthermore, the underlying protocol for web-based conferencing can work in a peer-to-peer fashion and proceeding in that direction can avoid the privacy concerns that could arise. With the rise of smart devices and the availability of a browser on almost every smart device, using the browser as an environment for providing noise suppression at the client-side holds a lot of promise in reducing latency and improving privacy and appears to be the most versatile. This work identifies a three-pronged goal toward noise suppression: reduced latency, improved quality and higher portability across various platforms. This work focuses on improving portability by realizing real-time noise suppression in the sender-side browser. It utilizes client-side JavaScript frameworks and libraries on the sender's end to process and deliver noise-suppressed audio for further transmission to the receiver's end through WebRTC. However, this has not been the direction of work in this domain, and therefore, the frameworks and libraries required for this task are not available for use in the client-side. The required libraries were hence built in order to facilitate the realization of this goal.

2 Literature Survey

This section further describes the methods developed in the past to reduce background noise in audio. These techniques are suitable only for a particular set of tasks—some of them apply to mobile phones alone, others are for use with specialized hardware and some other techniques do not work in a real-time setting (Fig. 1).

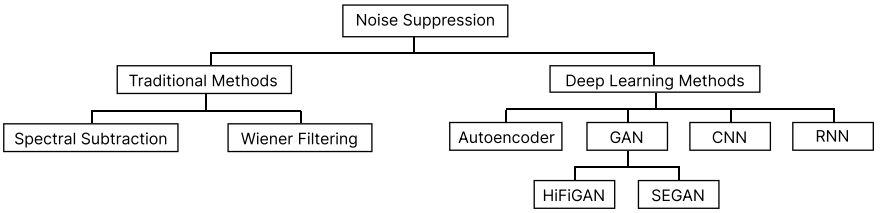


Fig. 1 Traditional and deep learning-based noise suppression methods

2.1 Traditional Methods

Traditional methods rely on statistical modeling of the noise and to estimate the noise representation from the given audio. The primary assumption in these methods is that noise is stationary or varying slowly with respect to time and does not undergo rapid changes suddenly. Spectral Subtraction [1] utilizes noise information from the parts of the audio where the speech signal is absent and models the noise of the audio based on these parts. It uses the short-term spectral amplitude and neglects phase information in developing the noise model, which enforces the constraint that noise is temporally stationary. There is, however, a lot of remnant noise after the suppression. This limitation can be tackled [11] by using the algorithm iteratively on the output or by altering the value being subtracted. Spectral mapping is used [9] for mapping corresponding noisy and clean signals to estimate the noise and to further remove it using spectral subtraction. Wiener filter, used largely in mobile phones, is an adaptive technique that guarantees noise reduction in stationary environments. The amount of noise reduced by the wiener filter is shown to be directly proportional to the degradation of the speech signal [2]. Hence, after a point, this technique leads to more signal loss. To conclude, the study in this section describes the promise shown by the traditional techniques in the domain of stationary noise suppression, but as their fundamental assumption is that noise is stationary, they are ineffective when sudden noise peaks are observed in the audio.

2.2 Deep Learning Methods

Initial attempts at using deep neural networks (multilayer perceptrons) for speech enhancement did not show much promise. The use of a recurrent neural network (RNN) [10] for the noise suppression task was further improved by Mozilla's RNNoise. RNNs have proven to be more effective in identifying and learning patterns across the temporal domain. However, they are not suitable for real-time applications as they are computationally more intensive. The use of Deep Autoencoders [5] for mapping from noisy to clean speech was proposed as an improvement. These were surpassed by CNN-based approaches [8], which again suffered from complexity

unsuitable for real-time systems. DEMUCS architecture [3] was an attempt to make noise suppression work on lighter hardware. It employs an encoder-decoder architecture with an LSTM that extracts features and performs noise suppression over the latent representation produced by the encoder, rather than the actual signal. Speech Enhancement GAN (SEGAN) [12] relies on the effectiveness of Autoencoders as well as GANs to denoise an audio signal but it takes more time for training and it is not stable. High Fidelity GAN (HiFi-GAN) [13] was a highly successful GAN approach for speech enhancement that provides studio level speech quality but is unsuitable for any real-time system due to the complexity of the model. Deep learning-based methods can benefit from a large noise corpus from diverse environments for training and hence accurately model noise despite its dynamic behavior with respect to time. Hence, in accordance with the above study, we choose the Deep Extractor for Music Sources (DEMUCS) architecture, which provides a good trade-off between the portability and quality.

3 Proposed Methodology

The proposed work unfolds with two modules: the audio-video chat application and the noise suppression module. The first module is responsible for creating and connecting the call in a secure manner and also for acquiring the media streams captured by the device from the browser, to be passed on to the next module. The second module reduces the noise from the audio stream separated from the media stream. The overall scheme of these is provided in Fig. 2.

3.1 Audio-Video Chat Application

Audio-Video Chat Application module is built with Javascript, using the WebRTC API, which implements the protocol of the same name. The WebRTC API provides

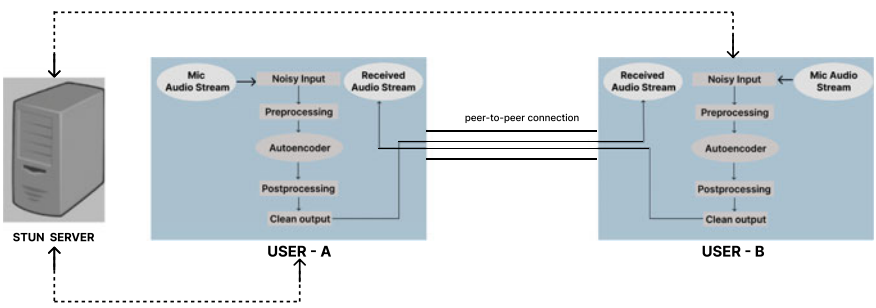


Fig. 2 Architecture diagram

a secure and abstract way of setting up peer-to-peer communication through WebSockets with a signaling server that only is in action till the two users are connected. The server does not by itself receive any of the data sent through the sockets. The communication is hence private.

Algorithm 1 describes the method used which is based on the WebRTC protocol. The Web Real-Time Communications (WebRTC) protocol is used for communication over web using audio, video, and text streaming between connected users. It consists of various components such as MediaStream, RTCPeerConnection, RTCDataChannel, and Signaling Server. A user initiates a call by creating a MediaStream and an RTCPeerConnection object. The peers establish a direct connection using NAT traversal techniques like Session Traversal Utilities for NAT (STUN) and Traversal Using Relays around NAT (TURN), which enable the traversal of NAT/firewall devices to reach the peer's IP address. The architecture of the system is depicted in Fig. 3.

Algorithm 1 AudioWebRTC

```
Acquire MediaStream from browser
Separate Audio and Video streams
Pass Audio and Video to WebDEMUCStreamer
Display user video as localstream
Generate sessionID
Send request to STUN server
if sessionID of user A == sessionID of user B then
    Set up PeerConnection between A and B
    while hangup is false do
        Send MediaStream from user A to user B
    end while
end if
```

The application allows users to connect to start a call with video and audio. It produces a unique session ID stored in a Firebase Firestore DB, which is a NoSQL database. Firebase is preferred for the implementation since the transmission of data is handled using web sockets and it is processed/stored in the cloud. The STUN

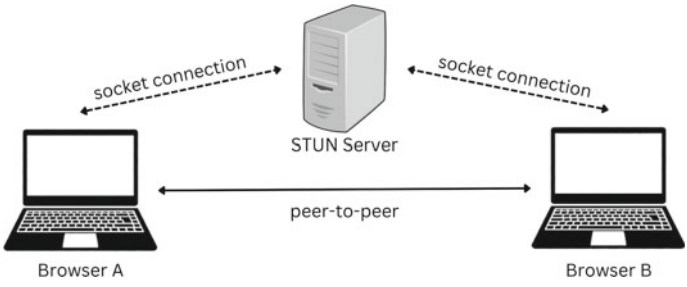


Fig. 3 The architecture of audio-video chat application that utilizes WebRTC

server used for signaling needs to store the unique IDs of each call session and the IDs of the individual participants so that they can be signaled. The second user can connect to the call by entering this ID, shared by the user who initiated the call. The WebRTC API provides methods to initiate peer-to-peer connection through a signaling server that acts as an intermediary.

3.2 Noise Suppression Module

Noise suppression module is built using the Web Audio API and Tensorflow.js. Web Audio API provides methods to read and manipulate audio streams and process the audio in a separate thread. Tensorflow.js is an API for building machine learning models and performing inference and can be used both on the client-side and server-side. It is, however, still under development. Hence, unlike the complete Tensorflow framework, model architectures are not directly available and have to be rebuilt. Tensorflow.js offers support for hardware acceleration via WebGL and is beneficial as it exploits the hardware of the system properly. Its low-level APIs allow developers to build custom models from scratch using standard building blocks like layers, optimizers, and loss functions. The DEMUCS [3] architecture works in the waveform domain to analyze the audio data. It uses an encoder-decoder-based architecture that produces higher portability. The encoder uses a stack of convolutional layers to extract information and produce a compressed representation of audio. The decoder reconstructs the original audio signal. The pretrained DEMUCS model is loaded using Tensorflow.js and the required preprocessing is performed by converting the audio to tensors.

Figure 4 shows the preprocessing step Short-term Fourier Transform (STFT) applied before passing its magnitude to the DEMUCS model. The model extracts the representation of the noise in the given audio and uses this to denoise the audio. The postprocessing step of Inverse STFT converts the audio back to a time-series. The required preprocessing modules were unavailable for client-side JavaScript in the pre-existent libraries and hence had to be redeveloped. Many of the existing audio processing libraries utilize numpy arrays, and all operations are defined to operate on these arrays. Client-side JavaScript does not offer support for numpy arrays, and hence an alternative was required. The tensor representation can be manipulated using Tensorflow.js library functions. This was used as a replacement for the numpy array object. The functions were adapted to suit the behavior of tensor objects. The advantage that came with this migration is the WebGL support offered by tensors. The computations on the tensor object can support hardware acceleration on systems with a GPU, but these computations are still optimal for running on a CPU.

Algorithm 2 takes MediaStream object from the browser as input. This has the provision to treat video and audio as separate streams. The audio stream extracted cannot be operated on by JavaScript methods or tensor operations. The format is converted to usable format by typecasting it into an AudioBuffer, which is a wrapper over the ArrayBuffer object in JavaScript. This array cannot be manipulated by tensor

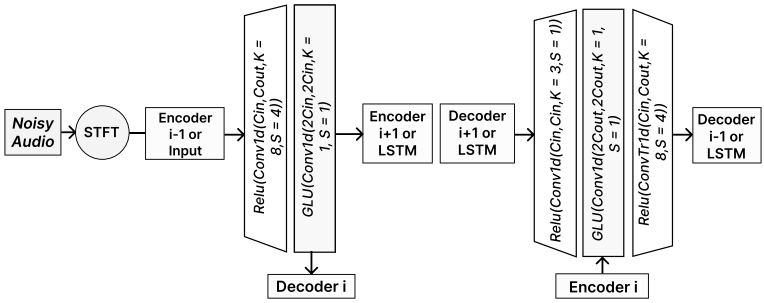


Fig. 4 DEMUCS architecture

methods, and hence must be converted into a tensor. In order to replicate the STFT method available in standard libraries in other languages, the tensor must be center-padded before the transform is applied. This must be reversed during postprocessing to extract the original audio. The data is then fed to the model chunk-by-chunk for inference in real-time. The resultant tensor is further processed before streaming it through WebRTC. The above computations require operating over complex-valued tensors, which have a greater overhead in operations like slicing and indexing. There is limited support for slicing operation over complex-valued tensors in client-side JavaScript. Since the focus remains on improving portability, all users irrespective of their hardware must be able to utilize the applications. Hence, the methods have to be engineered such that the operations are performed over real-valued tensors representing the real and imaginary parts separately and the resultant outputs are combined back only when required. This technique is used in both STFT and ISTFT functions in this module to avoid the overhead of slicing operation overhead.

Algorithm 2 WebDEMUCStreamer

Acquire $MediaStream$ from browser
Separate Audio and Video streams
Convert Audio stream to $AudioBuffer$
Convert $AudioBuffer$ to tensor of $frameLength$ and apply padding
Preprocess the tensor
    Apply $STFT$ on the tensor
    Split the result of STFT into $Magnitude$ and $Phase$
    Scale the $Magnitude$
    Reshape the $Magnitude$
Perform model inference
Postprocess the tensor
    Perform inverse scaling
    Perform $ISTFT$ by combining magnitude and phase only for required window
Convert tensor to $AudioBuffer$
Connect to the $MediaStream$

4 Outcomes and Results

4.1 Dataset

The clean and noisy data are taken from the Microsoft Scalable Noisy Speech Dataset of the Microsoft DNS Challenge 2022 [4]. The noise files are taken from the DEMAND corpus.

4.2 Metrics and Evaluation

Perceptual Evaluation of Speech Quality (PESQ) is a standard test recommended by ITU-T P.862 for speech quality assessment which predicts subjective opinion scores of the denoised audio [7]. Short-Time Objective Intelligibility (STOI) is an intelligibility measure which is highly correlated with the intelligibility of degraded speech signals. STOI is typically measured using a scale from 0 to 1, where a higher score indicates higher intelligibility.

Table 1 shows the performance metrics of DEMUCS-based noise suppression module, RNN model, Spectral Subtraction and Wiener Filter methods, which were discussed in Sect. 2. The DEMUCS model outperforms all other methods with the highest values of PESQ and STOI. Table 2 provides performance details for the real-time implementation of DEMUCS for noise suppression on an Intel i5 8th gen CPU at 1.6 GHz. The Real-Time Factor (RTF) of 2.06 indicates good real-time behavior of the system with very low lag due to model inference.

Figure 5a shows the spectrogram for a clean audio file. There is a significant difference from Fig. 5b which shows the same for the corresponding noisy audio. The noisy audio was generated by adding various noises from a noise corpus onto the clean audio and hence the correlation is reliably possible. Figure 5c–e shows the spectrograms for the noisy audio after being denoised by a RNN model, Spectral Subtraction and Wiener Filter, respectively. The output indicates that while there is an improvement (indicated by darker reds), there is also a huge difference from the clean audio. However the audio after it is denoised by the DEMUCS model (Fig. 5f) shows greater similarity to the clean audio and is a massive improvement over other techniques.

Table 1 Performance comparison of noise suppression techniques discussed under Sect. 2 along with DEMUCS

Technique	PESQ (out of 5)	STOI (out of 1)
Spectral subtraction	1.461	0.349
Wiener filter	4.408	0.986
RNN	1.385	0.617
DEMUCS	4.643	0.999

Table 2 Real-time evaluation of DEMUCS model inference

Parameter	Value
Model size	127.9 MB
Initial lag	41.3 ms
Stride	16.0 ms
Time per frame	32.9 ms
RTF	2.06
Total lag	74.2 ms

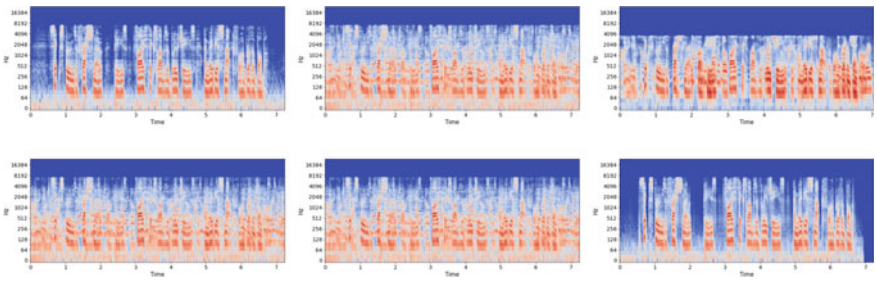


Fig. 5 Spectrograms: **a** clean audio, **b** noisy audio, **c** RNN denoised audio, **d** spectral subtraction denoised audio, **e** Wiener filter denoised audio, **f** DEMUCS denoised audio

5 Conclusions and Future Work

In this article, we endorse the use of deep learning techniques in a client-side environment with the support of web browsers to improve the quality of web-based calls. A noise suppression module for Audio-Video conferencing using WebRTC protocol has been proposed and implemented using client-side JavaScript. The libraries unavailable in the language were reimplemented using JavaScript and Tensorflow.js. The proposed noise suppression module, DEMUCS outperforms other deep learning methods as well as conventional methods with a PESQ score of 4.64 (out of 5) and STOI score of 0.99 (out of 1) (Table 2). The server does not by itself receive any of the data through the sockets. The communication is hence private. The main goal however was that of increasing portability, which was achieved by using client-side modules that can run within the browser environment. The computations required are adapted to perform well within the constraints of the browser and thus, making such an application possible. Improved portability can be further achieved by handling cases for WebGL and WebASM frameworks by optimizing individually for these platforms. This will guarantee major performance improvements for implementing such an application in large-scale video conference systems. The quality can be improved by using a more sophisticated generative model such as a GAN. While the focus here has been on call noise suppression, the deployment of deep learning models on the web that run on the client-side is a technique that has to be explored

more. This also pushes further for bridging gaps between native applications and web applications, which will largely benefit smart devices that are not meant to be dedicated computers.

References

1. Boll S (1979) Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans Acoust Speech Signal Process 27(2):113–120. https://doi.org/10.1109/TASSP.1979.1163209
2. Chen J, Benesty J, Huang Y, Doclo S (2006) New insights into the noise reduction Wiener filter. IEEE Trans Audio Speech Lang Process 14(4):1218–1234. https://doi.org/10.1109/TSA.2005.860851
3. Defossez A, Synnaeve G, Adi Y (2020) Real time speech enhancement in the waveform domain. arXiv preprint arXiv:2006.12847
4. Dubey H, Gopal V, Cutler R, Aazami A, Matusevych S, Braun S, Eskimez SE, Thakker M, Yoshioka T, Gamper H et al (2022) ICASSP 2022 deep noise suppression challenge. In: ICASSP 2022–2022 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 9271–9275. https://doi.org/10.48550/ARXIV.2202.13288
5. Feng X, Zhang Y, Glass J (2014) Speech feature denoising and dereverberation via deep autoencoders for noisy reverberant speech recognition. In: 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 1759–1763
6. Grant M, Kennedy J, Zhu J, Tan J, Markovski S, Popa C (2021) Videoconferencing privacy and security concerns in the pandemic
7. Hu Y, Loizou PC (2007) Evaluation of objective quality measures for speech enhancement. IEEE Trans Audio Speech Lang Process 16(1):229–238
8. Isik U, Giri R, Phansalkar N, Valin JM, Helwani K, Krishnaswamy A (2020) PocoNet: better speech enhancement with frequency-positional embeddings, semi-supervised conversational data, and biased loss. arXiv preprint arXiv:2008.04470
9. Juang BH, Rabiner LR (1987) Signal restoration by spectral mapping. In: ICASSP'87. IEEE international conference on acoustics, speech, and signal processing, vol 12, pp 2368–2371
10. Maas A, Le QV, O'Neil TM, Vinyals O, Nguyen P, Ng AY (2012) Recurrent neural networks for noise reduction in robust ASR. In: 13th annual conference of the international speech communication association, INTERSPEECH
11. Navneet Upadhyay AK (2015) Speech enhancement using spectral subtraction-type algorithms: a comparison and simulation study. Procedia Comput Sci 54:574–584. https://doi.org/10.1016/j.procs.2015.06.066
12. Pascual S, Bonafonte A, Serra J (2017) SEGAN: speech enhancement generative adversarial network. arXiv preprint arXiv:1703.09452
13. Su J, Jin Z, Finkelstein A (2021) HiFi-GAN-2: studio-quality speech enhancement via generative adversarial networks conditioned on acoustic features. In: 2021 IEEE workshop on applications of signal processing to audio and acoustics (WASPAA). IEEE, pp 166–170

Event Management Using Blockchain

Ajay Talele, Rushal Patil, Saloni Nimgaonkar, Parth Shethji, and Mayur Patil

Abstract The rapid advancement of blockchain-based technology, online booking of tickets has come as a great relief to most enthusiasts. Event management is a good fit for blockchain technology because of its attributes like transparency, security, and immutability. This project develops a system which provides a real time booking and transfer of tickets, events for customers and in turn examines the potential of blockchain technology in event management. People can book tickets. The controls are decentralized or controlled by algorithm. They can book, buy, and retrieve information and transfer tickets.

Keywords Blockchain · Immutable · Solidity · Tickets · Transactions

1 Introduction

The process of planning and executing an event is known as event management. This process includes pre-event planning all the way to post-event planning. In today's changing world, we must learn how to manage in-person events and virtual programs. Managing a virtual event involves the same processes as managing an in-person event, with the added challenge of doubling the excitement of your content. While traveling, making connections, and getting free food are advantages of in-person gatherings, the

A. Talele · R. Patil (✉) · S. Nimgaonkar · P. Shethji · M. Patil
Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India
e-mail: rushal.patil@vit.edu

A. Talele
e-mail: ajay.talele@vit.edu

S. Nimgaonkar
e-mail: saloni.nimgaonkar21@vit.edu

P. Shethji
e-mail: parth.shethji21@vit.edu

M. Patil
e-mail: mayur.patil21@vit.edu

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_15

success of a virtual event mostly depends on the quality of its content. It is important to sure that the presenters are ready to deliver their topic online while hosting a virtual event, and that the content is engaging and condensed.

Blockchain is a decentralized irreproducible digital ledger of economic activity that may be used to record nearly any exchange of value, not only money. Without a centralized administrator, electronic cash can be managed via blockchain. Based on the cryptocurrency based on the cryptocurrency bitcoin, uses a distributed database. This means that no central authority like a government or bank supervises the transaction (nodes). The blockchain is made up of a chain of data, including transaction data that has a timestamp that cannot be altered after it has been recorded.

The emergence of blockchain technology has introduced new possibilities for improving event management procedures. Smart contracts, which are self-executing contracts with the contents of the agreement explicitly put into lines of code, are one of the essential components of blockchain technology. A variety of event management tasks, including ticketing, payment, and settlement, may be automated with the help of smart contracts, making sure that everyone engaged in the event is aware of their duties and responsibilities. The decentralized nature of blockchain technology is another benefit for event management. This makes it harder for malicious parties to infiltrate the system since data is kept throughout a network of nodes rather than in a single location. In order to better understand how blockchain technology could be used to increase efficiency.

The issues that arise during event management comprise of fraudulent ticketing, ineffective payment methods, and insufficient transparency in event preparation. A blockchain-based approach was set out to address these problems, and it included secure ticketing, smart contracts for event organization and execution, and decentralized payment mechanisms. To evaluate the practicality of these solutions, a blockchain-based event management platform prototype was created and put to the test. Event planners can design, administer, and advertise events using the platform, while attendees can safely buy tickets, follow the event's development, and offer feedback.

Blockchain technology is a distributed digital ledger that enables safe and open transactions without the need for a centralized middleman. It is a decentralized system that employs cryptographic techniques to preserve the confidentiality and integrity of the data kept on it.

Due to its potential to offer supply chain transparency, security, and immutability, blockchain technology has drawn attention as a prospective solution. The study focuses on three primary areas: IOT-enabled energy management, smart logistics and transportation, and smart business models. It looks at how blockchain technology and Industry 4.0 may advance supply chains toward sustainability. Blockchain is a decentralized ledger that is used to conduct transactions and securely trade virtual currency. It is a distributed database that maintains an ever-expanding, impenetrable data structure, with blocks of varying widths holding batches of distinct transactions. Bitcoin is a permissionless peer-to-peer network that allows users to connect to the network and submit new transactions to validate and create new blocks.

Blockchain can be utilized to enhance several procedures, including event planning, payment systems, and tickets. The creation of tamper-proof and immutable records using blockchain technology may aid in the prevention of fraud, guarantee the transparency of the event planning and execution process, and offer a safe and effective payment system. Event planners can give guests experiences that are more frictionless, safe, and pleasant by utilizing the power of blockchain.

2 Literature Review

Derick et al. [1] This paper investigates the use of web services in an online reservation system through a case study of a hotel reservation system. The study finds that web services are critical to the functioning of the system, using SOAP-based web services to interact with external systems such as payment gateways and inventory management systems. The authors note the importance of web services in ensuring the security and reliability of the online reservation system. Overall, the paper provides valuable insights into the use of web services in online reservation systems and highlights their importance in ensuring smooth functioning.

Yaga et al. [2] This paper gives a general understanding of blockchain technology, a decentralized digital ledger that securely and openly records transactions. The main characteristics of blockchain technology are covered in the article, including its distributed structure, immutability, and consensus process. The writers also look at how blockchain technology is being used in numerous fields, including supply chain management, healthcare, and finance. The article offers a thorough knowledge of blockchain technology and how its capacity for safe and open transaction processing can revolutionize a number of sectors.

Hasan et al. [3] The paper discusses the importance of documentation in an online booking system. The authors highlight the challenges of developing and maintaining such a system and emphasize the need for clear and concise documentation to ensure its smooth functioning. The paper presents a framework for documenting an online booking system, which includes the system's design, implementation, testing, and maintenance. The authors also emphasize the importance of updating the documentation regularly to keep up with changes in the system. Overall, the paper provides a valuable guide for documenting an online booking system and highlights the importance of clear and concise documentation in ensuring its successful development and maintenance.

Sahoo et al. [4] This paper explores the potential of blockchain technology for sustainable supply chain management. The authors discuss the challenges faced by supply chain management in ensuring sustainability, such as ethical sourcing, transparency, and accountability. The paper examines the features of blockchain technology that can address these challenges, such as its decentralized and immutable ledger, smart contracts, and supply chain traceability. The authors also discuss several

case studies that demonstrate the application of blockchain technology in sustainable supply chain management. Overall, the paper provides a valuable overview of the potential of blockchain technology in promoting sustainability in supply chain management.

Singh et al. [5] This paper discusses the potential of blockchain technology as the future of security. The authors examine the security challenges faced by traditional systems, such as centralized control, data breaches, and hacking. The paper explores the features of blockchain technology that can address these challenges, such as its decentralized nature, cryptographic algorithms, and consensus mechanism. The authors also discuss several case studies that demonstrate the application of blockchain technology in enhancing security in various industries, such as finance, healthcare, and identity management. Overall, the paper provides a comprehensive overview of the potential of blockchain technology as the future of security.

3 Methodology

For event planners, we have created a management system that makes purchasing tickets and managing them very simple. We utilized an online compiler to do this since it makes coding and compilation simple. Remix was the compiler that we chose. Traditional web development methods, such as HTML, CSS, and JavaScript, are used to create the frontend. Also, we used the programming language solidity to create smart contract. Using the remix IDE, we created a smart contract first. Three methods were created–one for setting up an event, one for purchasing tickets, and one for ticket transfers. We also mapped events to see how any ongoing occurrences are being handled.

We must first compile the smart contract made using the built-in compiler in order to access the system. Following that, we must select our account and run the code. For educational reasons, we may utilize Goerli test net or Remix's built-in fake account feature. We can create our account using MetaMask, but first we must mine for ethers. Once the account is provided and the smart contract is deployed, we can then start using our system. Now, when we access the deployed contracts tab in Remix ide, we can see our contract and an option for various other tasks. The first step is to establish an event, where we are required to provide the event's name, date, cost, and the number of tickets available. Following that, there is another button to see our event but for that you must provide the event ID before you can purchase or transfer tickets. Another button is dedicated to indicating the number of remaining tickets. To the bottom right of the screen there is terminal where we can see the result of our transaction. If it is successful, the transaction cost is automatically subtracted otherwise the terminal generates an error.

4 Results

At the end of the project, we created a smart contract that allows us to host an event and then buy or transfer tickets to anyone we like. The smart contract produced using Remix IDE. This is the website we created using web development technologies (Figs. 1, 2, 3, and 4).

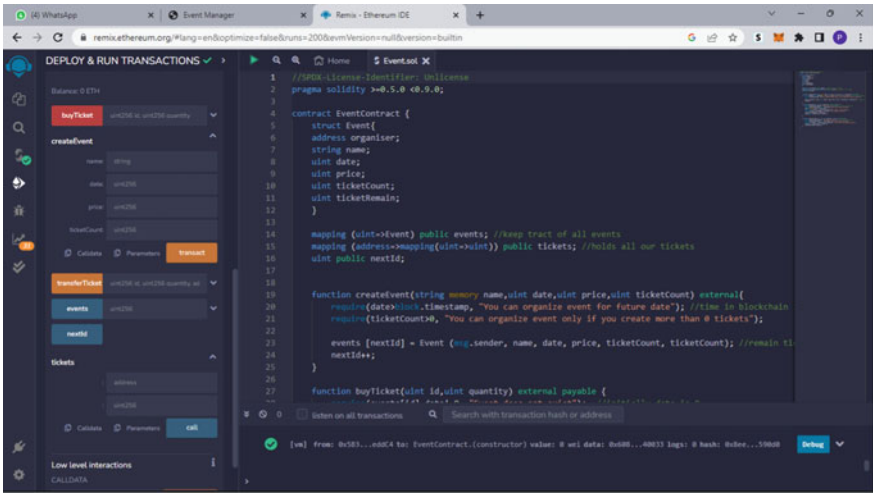


Fig. 1 Remix IDE code implementation

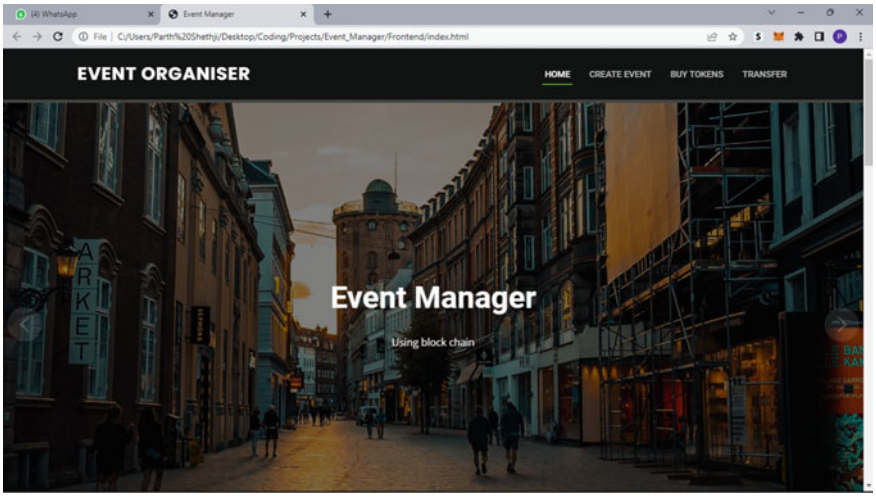


Fig. 2 Website home page

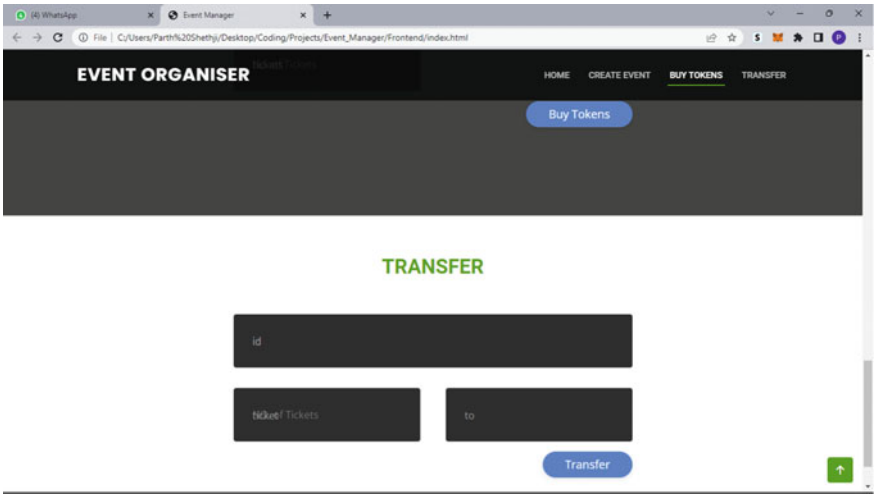


Fig. 3 Ticket transfer page

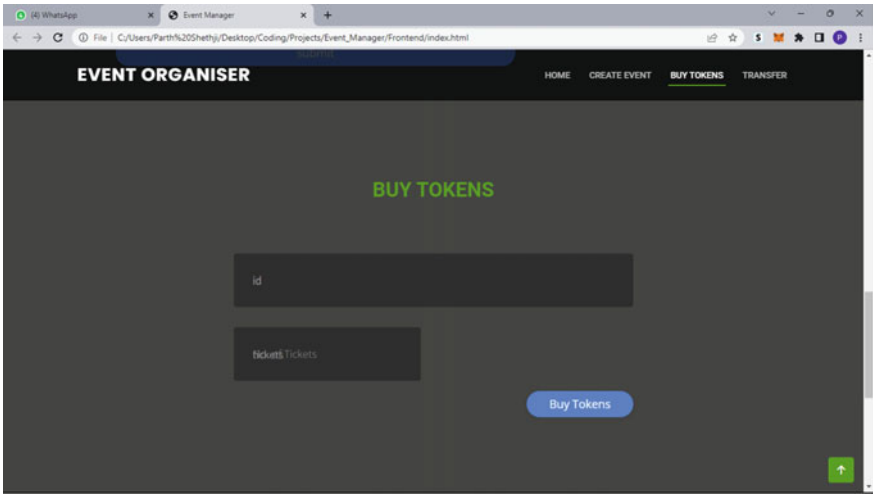


Fig. 4 Ticker booking page

5 Future Scope

The backend of the website will be our primary emphasis now that the frontend and smart contract have been built. In addition to developing the smart contract, we also used the built-in Remix IDE to deploy the code. Now we need to link the smart contract to the website's backend. Also, we may make our own account on MetaMask and use Ethereum instead of the fake accounts. The Ethereum API must

be integrated with the website's backend code in order to connect smart contracts to it. Either web3.js or ethers.js, two well-liked JavaScript libraries for communicating with the Ethereum network, can be used to do this. On our own local blockchain, we also want to host the smart contract using Ganache or Hardhat. We also plan to develop a decentralized identity verification system using blockchain technology that enables event participants to demonstrate their identity without disclosing any personal information. This can guard guests' privacy and aid in preventing fraud. Blockchain technology may be utilized to provide a decentralized platform for event financing where anybody can make donations and earn incentives in return. This can make it simpler for event planners to generate money and expose their events to more people. With the use of blockchain technology, we can establish a safe and open system for gathering participant input and using the results of the analysis to enhance subsequent events.

6 Conclusion

In this project, we presented some considerations for the implementation of the smart contract using blockchain technology. Often, those who purchase tickets attempt to resell them on the illicit market by raising the price. There is no middleman in this case because we are employing blockchain technology, and all transactions are governed by written smart contracts. To ensure the security of a blockchain system or network, a detailed risk analysis is performed. Cybersecurity frameworks, security testing processes, and secure coding techniques are used to protect a blockchain system against online fraud, breaches, and other. This project demonstrates how blockchain technology has the ability to revolutionize event management and improve the entire experience for event planners, attendees, and other stakeholders. Moreover, it talks about potential uses of blockchain technology in event management as well as future research prospects.

References

1. Derick Wasonga JO (2019) An investigation into web services: A case of an online reservation system. In: IEEE-SEM. 10(6), ISSN 2320–9151
2. Yaga D, Peter M, Roby N, Scarfone K (2019) Blockchain technology overview. arXiv preprint arXiv:1906.11078. In: Maleh Y, Shojafar M, Alazab M, Romdhani I (eds) Blockchain for cyber security and privacy, Edi-1, Taylor and Francis group, LLC, 6000 broken sound parkway NW, Suite 300, Boca Raton, FL 33487–2742
3. Hasan S (2014) Documentation of online booking system. Online hotel booking system: Asia Pacific institute of innovation technology, Malaysia

4. Sahoo S, Kumar S, Sivarajah U, Lim WM, Westland JC, Kumar A (2022) Blockchain for sustainable supply chain management: Trends and ways forward. Electr Comm Res. 1–56
5. Singh S, Singh N (2016) Blockchain: future of financial and cyber security. In: 2016 2nd international conference on contemporary computing and informatics (IC3I), pp 463–467 IEEE

Impact of Artificial Intelligence on Human Resource Management Practices in Private Organizations of India

Divya Hiran, Hemant Kothari, and Shivoham Singh

Abstract The term "artificial intelligence" (AI) refers to a technology that gives computers the ability to learn from and make decisions or recommendations based on data that has been gathered in the past. When it comes to the management of human resources, artificial intelligence can be utilized in a wide variety of contexts for the sake of streamlining procedures and increasing operational effectiveness. In the current research work the main objectives were to measure the impact of AI-based applications on organizational performance of private companies in India and does AI-based application helps in effective decision making for the top-level management of private companies in India. The survey method is being applied in this quantitative research with data collection instrument, the questionnaire. The non-probability-based convenience sampling method is used for sample selection and sample size of about 120 respondents as senior officials, executives, managers, IT expert, etc. being considered from different companies of India. Accordingly, the outcomes of the research work confirm that there is statistically significant impact of artificial intelligence-based human resource operations/practices on the organizational performance of private companies in India as the null hypothesis $H_0 1$ being rejected. It was also found that there is significant impact of artificial intelligence-based human resource practices on the organizational decision-making process of private companies in India. Based on the results it is recommended that the private companies in India should invest in implementation of AI-based applications for HRM practices so that it helps them in better decision making, improves organizational overall performance and eases the HR operational activities.

Keywords AI · Human resource management · Recruitment

D. Hiran
Department of Home Science, Govt. Meera Girl's College, Udaipur, India
e-mail: divyahiran123@gmail.com

H. Kothari · S. Singh (✉)
Pacific Academy of Higher Education and Research University, Udaipur, India

S. Singh
Symbiosis Institute of Business Management, Hyderabad (SIBM-H), Hyderabad, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_16

1 Introduction

Recruiting, selection and acquisition of talent, reporting, payroll and compensation are just some of the HR tasks that could benefit greatly from the use of AI technologies. Today's cutting-edge artificial intelligence (AI) tools have a profound effect on how businesses are run. Executives in the HR field are confident that incorporating AI into HR administration tasks would boost efficiency and enhance the working conditions for all employees. This will allow for better personnel management by providing more resources, more time and more precise data. Nilsson (2005) argues that, with human-level artificial intelligence, machines should be able to perform most of the tasks that human intelligence requires.

Increasing amounts of human resources data are being generated in the cloud by human and machine learning efforts, and these analyses can be improved with the use of artificial intelligence (AI). Accordingly the key to a company's long-term success lies in its ability to strategically blend people, process and technology to provide game-changing value at optimum cost. As a result of AI, many HR administrative tasks can be automated for efficient and trustworthy transaction processing and service provision. Insight into automation intelligent through the technology-agnostic chabot is the primary subject of the study by Nilsson (2005), which also covers conversational AI capabilities for HR transactions. With AI taking over mundane jobs that recruiters have traditionally handled, the HR sector is entering a "new age," or what some have dubbed "the Fourth Industrial Revolution."

Today due to the use of technology the efficiency and performance of any organization depend on the use of AI in their business environment [1]. Similarly the use of AI plays important role in human resource management domain. Whether it is making choices about complex problems or providing personalized assistance, AI has shown to be an effective tool [2]. Most companies, according to Jain [3], use cutting-edge technology for various HR functions, including recruitment, performance evaluation and the use of cloud-based HR services. Jarrahi [4] suggests that AI could aid in decision making, ambiguity management and especially split opinions inside an organization.

2 Related Work

According to O'Connor's [5] paper, "Artificial Intelligence (AI) in Human Resource Management" AI has already had and will uphold to have a significant impact on HRM in the coming years. Human resources experts also ought to be better prepared for potential difficulties. Therefore, HR professionals who care about their futures should do what they can to keep up with the latest developments in the area and build a solid groundwork in HR expertise upon which to build.

Wisetsri et al. [6] discussed about the use of artificial intelligence (AI) into human resources operations that would improve enterprises owing to the fact that these

programmes can observe, anticipate and diagnose to help HR departments make better decisions. Assessing AI in human resources management (HRM) is the goal of the observation. The research work by the author includes the importance of AI in HRM, conceptual model for artificial intelligence and various implementation aspects related to AI-enabled human resource practices. The study will help the policymakers in better understanding the benefits of AI in HRM activities.

According to Tiwari et al. [7] the technology has accelerated all industries in a competitive world. AI helps the sector expand quicker and function more efficiently. Finance, HR, marketing, production, etc. use this technology. The AI system has improved performance and daily operations. Due to the dynamic and competitive workplace, managers are under pressure and realizing the necessity for artificial intelligence. Quantitative research and regression analysis were employed by authors. AI is used in HR processes from talent recruiting to performance evaluation. The research examine how AI affects HR functions and HR department functions and understand innovativeness and HR operations. The survey included IT company HR personnel. The investigation showed that AI influences simplicity of use and innovativeness. The study explains artificial intelligence, which is driving Industry 4.0.

The main goal of the study given by Tsymbaliuk et al. [8] is to evaluate how artificial intelligence is used in the HR practices in Ukrainian firms and the potential spread of AI in HR. The survey's findings partially supported the second working hypothesis, according to which the use of AI in HR has both good and negative effects, and they supported the first working hypothesis, according to which the expansion of AI in HR is a natural process. The capacity to carry out various HR operations and procedures utilizing AI technology, which needs adequate training and expertise, has become one of the most crucial abilities of HR professionals as a result of the widespread adoption of AI in HR today.

Gaidhani [9] discussed that most people believe that soon human resource managers will be replaced by artificial intelligence. The study aims to address the fact that artificial intelligence actually speeds up and simplifies workers' tasks by automating and enhancing repetitive, difficult tasks. Through the use of current examples, the study also demonstrates how artificial intelligence is reshaping and supporting numerous human resource practices. Finally, it lists a number of issues and effects related to implementation of artificial intelligence technique in real-time business environment.

Sanyaolu and Atsaboghena [10] HRM must adapt to new business areas to maximize organizational growth. The study examines the importance of artificial intelligence (AI) in human resources departments for recruiting, selection, on boarding, retention, compensation management and employee management. AI-enhanced HRM is changing how firms recruit, manage and engage employees. Artificial intelligence lets robots make better judgements based on past facts and behaviour than humans. This trend has replaced all manual labour with machines, prompting HR professionals to become more strategic. AI's strengths and drawbacks in HRM units were examined. The study discusses how AI may help firms improve their HR functions. The Human Resource Department in the organization is undergoing a

digital revolution and employing numerous approaches to simplify resources through big data analysis. Amla and Malhotra [11] discuss artificial intelligence and cloud computing. Most organizations use artificial intelligence or digital technology in human resource management, such as a machine learning, chatbot and robot process automation, to help recruiting, screening, on boarding and interviewing, among other things.

Jia et al. [12] create an efficient and corresponding system of intelligent decision support for HR planning and strategy by integrating these disciplines with knowledge and data mining [13, 14]. The recruitment process may be handled more efficiently and with higher quality with AI solutions.

3 Methodology

3.1 Population and Sample Selection

The main research intend is to identify the impact of the AI-based HRM system on the organizational decision making and ease of HR operations such as human resource planning (HRP), hiring, recruitment, selection, training, on boarding, payroll system, employee relations, compensation and benefits, performance appraisals, productivity, performance management and off boarding. The study population includes all private organizations in India. The sample includes the respondents as managers and employees working in selected companies where AI-enabled application is being used. In order to collect primary data a well-structured questionnaire was being used. After data is collected, it is synthesized using Social Science Statistical tools like SPSS. To examine how an AI-based HRM system influences decision making and productivity in Indian businesses, a multiple regression analysis is conducted.

3.2 Survey Instrument—Questionnaire

Respondents being considered in the research work include HR managers, IT managers, unit mangers, IT executives, communicators, supervisor of service, top management officials, digitization managers, IT architects as well as data engineers and other senior managers. The questionnaire has around 35 questions in total, 5 of which are related to each of the 7 variables of the study, which are organizational performance, employee recruitment, employee training and development, decision making, employee performance evaluation, employee compensation and ease of HR operations. In addition, the questionnaire includes questions regarding the ease of HR operations.

Table 1 Result Cronbach's alpha

Cronbach's alpha	No. of items
0.785	35

Sample Design includes:

Sampling technique: non-probability-based convenience sampling.

Sample size: 120 respondents from various private organizations related to Financial, EduTech., Pharmaceutical and others sectors in India were being considered for the study.

Demographic Details of Respondents:

Gender-wise classification: About 60% respondents were male, and 40% were female presently working in private organizations in India being considered proportionately. Mainly the respondents belong to the age group 25–35 years with having experience > 5 years.

Data collection: Both primary and secondary sources of data collection were being used. The primary data was collected with the help of well-structured questionnaire. The instrument questionnaire was being tested for reliability and validity using Cronbach's alpha test.

According to the outcomes shown in Table 1, it was found that the Cronbach's alpha value is 0.785; as this score is greater than the value of 0.70 it can be concluded that the instrument designed or the questionnaire is reliable and has good internal consistency between the questions confirming that the questions have shared covariance and will help in finding the impact of AI on HR operations. The secondary source of data collection is being done through various online resources like government sites, company sites, online available annual reports, etc. also from newspapers and AI and HRM-based journals.

3.3 Objectives

1. To measure the impact of AI-based applications on organizational performance of private companies in India.
2. AI-based application helps in effective decision making for the top-level management of private companies in India.

3.4 Hypotheses (S)

1. $H_0$1 (Null): There is no statistically significant impact of artificial intelligence-based human resource practices on the organizational performance of private companies in India.

H_a1 (Alternative): There is statistically significant impact of artificial intelligence-based human resource practices on the organizational performance of private companies in India.

2. $H_0$2 (Null): There is no significant impact of artificial intelligence-based human resource practices on the organizational decision-making process of private companies in India.

H_a2 (Alternative): There is significant impact of artificial intelligence-based human resource practices on the organizational decision-making process of private companies in India.

3. $H_0$3(Null): AI-based application does not ease the human resource management operations in private companies of India.

H_a3(Null): AI-based application does ease the human resource management operations in private companies of India.

3.5 Regression Models

$$\text{Model-1}: \ \text{OP}_\text{i} = \beta 0 + \beta 1\text{AIETD}i + \beta 2\text{AIER}i + \beta 3\text{AIPEPE}i + \beta 4\ \text{AIEC}i + \varepsilon \text{m}1i$$

whereas

$\beta 0$ = Constant or risk.

$\beta 1$, $\beta 2$, $\beta 3$,4 = Coefficients.

OPi = Organizational performance of "i" private firm in India.

AIETDi = AI-based employee training and development of "i" private firm in India.

AIERi = AI-based employee recruitment of "i" private firm in India.

AIPEPEi = AI-based employee performance evaluation of "i" private firm in India.

AIECi = AI-based employee compensation of "i" private firm in India.

εm1i = Indicates the error term of firm "i."

$$\text{Model-2}: \ \text{DM}_i = \gamma 0 + \gamma 1\ \text{AIETD}i + \gamma 2\ \text{AIER}i + \gamma 3\ \text{AIPEPE}i + \gamma 4\ \text{AIEC}i + \varepsilon \text{m}2i$$

whereas

$\gamma 0$ = Constant or risk.

$\gamma 1$, $\gamma 2$, $\gamma 3$,4 = Coefficients.

DMi = Decision making of "i" private firm in India.

AIETDi = AI-based employee training and development of "i" private firm in India.

AIERi = AI-based employee recruitment of "i" private firm in India.

AIPEPEi = AI-based employee performance evaluation of "i" private firm in India.

AIECi = AI-based employee compensation of "i" private firm in India.

εm2i = Indicates the error term of firm "i."

$$\text{Model-3}: \ \text{EHRO}_i = \partial 0 + \partial 1 \ \text{AIETD}i + \partial 2 \ \text{AIER}i + \partial 3 \ \text{AIPEPE}i + \partial 4 \ \text{AIEC}i + \varepsilon \text{m3}i$$

whereas

∂_0 = Constant or risk.

∂_1, ∂_2, ∂_3, ∂_4 = Coefficients.

EHROi = AI-based ease of HR operations of "i" private firm in India.

AIETDi = AI-based employee training and development of "i" private firm in India.

AIERi = AI-based employee recruitment of "i" private firm in India.

AIPEPEi = AI-based employee performance evaluation of "i" private firm in India.

AIECi = AI-based employee compensation of "i" private firm in India.

εm3i = Indicates the error term of firm "i."

Multiple regression analysis is being used to find the association between dependent and independent variables. The proposed model can be classified into three sub-models Model-1, Model-2 and Model-3 as shown below in Fig. 1.

4 Results

Based on the descriptive statistics values as shown in Table 2 it is clear that majorly the mean score of the dataset is between the values 3.3 and 4.01 which indicates that the dataset is having closer values. Although the maximum and minimum values lie between 5 and 1, respectively. Contrarily, the absence of an outlier is also reflected in the standard deviation numbers, which show that the data for all variables is not spread out. As a result, the data is normally distributed, and the majority of respondents gave the same answers across the board.

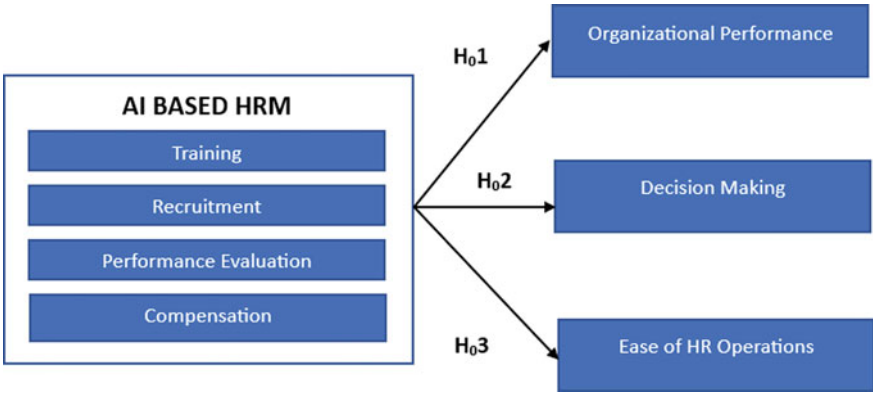


Fig. 1 Research framework

Table 2 Descriptive analysis

Descriptive statistics					
	Obs	Minimum	Maximum	Mean	Std. deviation
AIER	120	1.00	5.00	3.5000	1.27021
AIETD	120	1.00	5.00	3.3167	1.14483
AIPEPE	120	1.00	5.00	4.0167	1.13006
AIEC	120	1.00	5.00	3.7417	1.35036
OP	120	1.00	5.00	3.6167	1.37311
DM	120	1.00	5.00	3.8833	1.08607
EHRO	120	1.00	5.00	3.3333	1.37403
Observation (listwise)	120				

4.1 AI-Based Application and Organizational Performance

$H_0 1$ (Null): There is no statistically significant impact of artificial intelligence-based human resource practices on the organizational performance of private companies in India (Tables 3, 4 and 5).

After finding the reliability and internal consistency of the instrument now the impact of artificial intelligence (AI)-based human resource practices on the organizational performance of private companies in India is being analysed through statistical technique multiple regression analysis.

$$\text{Model-1}: \ \text{OP}i = \beta 0 + \beta 1\text{AIETD}i + \beta 2\text{AIER}i + \beta 3\text{AIPEPE}i + \beta 4\ \text{AIEC}i + \varepsilon m1i$$

$$\text{Model-1}: \ \text{OP}i = -0.499 + 0.330\text{AIETD}i + 0.205\text{AIER}i + 0.312\text{AIPEPE}i + 0.305\text{AIEC}i + 0.133$$

Table 3 Variables details: Model-1

Variables details as entered/removed[a]			
Model	Entered variables	Removed variables	Technique/method
1	Employee recruitment (AIER) Employee training and development (AIETD) Employee performance evaluation (AIPEPE) Employee compensation (AIEC)[b]		Enter

[a]Dependent variable: (OP: organizational performance)
[b]All requested variables entered

Table 4 Model-1 outcomes

Model summary				
Model	R value	Value of R-square	Value of adjusted R-square	Standard error of the estimate
1	0.953[a]	0.908	0.905	0.26627

[a] Predictors: (Constant), employee recruitment (AIER), employee training and development (AIETD), employee performance evaluation (AIPEPE), employee compensation (AIEC)

Table 5 Model-1: ANOVA results

ANOVA[a]						
Model		Value of sum of squares	(df) Degree of freedom	Mean square value	*F*-value	Sig
1	Regression	80.638	4	20.159	284.331	0.000[b]
	Residual	8.154	115	0.071		
	Total	88.792	119			

[a]Dependent variable: (OP: organizational performance)
[b]Predictors: (Constant), employee recruitment (AIER), employee training and development (AIETD), employee performance evaluation (AIPEPE), employee compensation (AIEC)

The results of the multiple regression analysis conducted for this study's Model-1, which represents the relationship between artificial intelligence-based HRM and organizational success, are displayed in Table 6. The statistical results for Model-1 show that the model represents good fitness, demonstrating that the independent variables utilized in the current study are adequate to offer useful conclusions. Therefore, the independent or explanatory variables included in the study are sufficient to produce significant results, as indicated by the *R*-square value (0.908) of Model-1, which states that they account for 90.8% of the data. In ANOVA statistics, the *F*-value and *P*-value are included. The multiple regression Model-1 in the current investigation reveals a superior fit to the dataset with the studied variables, according to the *F*-test value of 284.331. Also, the results of the *F*-test are supported by the *P*-value (0.000) of the *F*-test, which is significant at the 0.05 level or 95% confidence interval.

Table 6 Model-1: coefficients details

Coefficients[a]						
Model		Coefficients values un-standardized		Coefficients values standardized	*t*	Sig
		B	Std. error	Beta		
1	(Constant)	− 0.499	0.133		− 3.750	0.000
	Employee recruitment (AIER)	0.205	0.023	0.301	8.798	0.000
	Employee training and development (AIETD)	0.330	0.024	0.437	13.884	0.000
	Employee performance evaluation (AIPEPE)	0.312	0.025	0.408	12.242	0.000
	Employee compensation (AIEC)	0.305	0.020	0.477	15.203	0.000

[a]Dependent variable: OP: organizational performance

So, it can be concluded that the hypothesis $H_0 1$ is being rejected which specifies that there is statistically significant impact of artificial intelligence-based human resource practices on the organizational performance of private companies in India.

4.2 AI-Based Application and Decision Making

$H_0 2$ (Null): There is no significant impact of artificial intelligence-based human resource practices on the organizational decision-making process of private companies in India (Tables 7, 8, 9 and 10).

$$\text{Model-2}: \ \text{DM}_i = \gamma 0 + \gamma 1\ \text{AIETD}i + \gamma 2\ \text{AIER}i + \gamma 3\ \text{AIPEPE}i + \gamma 4\ \text{AIEC}i + \varepsilon m2i$$

$$\text{Model-2}: \ \text{DM}_i = -0.694 + 0.34\ \text{AIETD}i + 0.130\text{AIER}i + 0.357\text{AIPEPE}i + 0.358\text{AIEC}i + 0.248$$

Multiple regression analysis was being used to find that does AI-based applications help in better decision making in reference with India companies. The regression analysis: Model-2 sub model results in the form of *R*, *R*-Square, ANOVA and Coefficients are shown below in the tables. In this analysis artificial intelligence-based HR practices (Employee recruitment, Employee training and development, Employee performance Evaluation and Employee compensation) are being considered as independent variables and decision making (DM) being considered as dependent variable. The model summary Model-2 table shows the *R* and *R*-Square values of the model.

Table 7 Variables details: Model-2

Variables details as entered/removed[a]			
Model	Entered variables	Removed variables	Technique/method
2	Employee recruitment (AIER) Employee training and development (AIETD) Employee performance evaluation (AIPEPE) Employee compensation (AIEC)[b]		Enter

[a]Dependent variable: (DM: decision making)
[b]All requested variables entered

Table 8 Model-2 outcomes

Model summary				
Model	*R* value	Value of *R*-square	Value of adjusted *R*-square	Standard error of the estimate
2	0.863[a]	0.744	0.735	0.49620

[a]Predictors: (Constant), employee recruitment (AIER), employee training and development (AIETD), employee performance evaluation (AIPEPE), employee compensation (AIEC)

Table 9 Model-2: ANOVA results

ANOVA[a]						
Model		Value of sum of squares	(df) degree of freedom	Mean square value	*F*-value	Sig
2	Regression	82.352	4	20.588	83.619	0.000[b]
	Residual	28.315	115	0.246		
	Total	110.667	119			

[a]Dependent variable: (DM: decision making)
[b]Predictors: (Constant), employee recruitment (AIER), employee training and development (AIETD), employee performance evaluation (AIPEPE), employee compensation (AIEC)

R-square value of the model, which indicates how well the chosen research model fits the research data and how much variability in the dependent variable is accounted for by the independent variable. Accordingly, the *R* value found to be 0.863 is showing good correlation between AI-based HR practices and decision making. Therefore, the independent factors of the study are sufficient to produce significant results, as indicated by the *R*-Square value (0.744) of Model-2, which states that these variables explain 74.4% of the data. In ANOVA statistics, the *F*-value and *P*-value are included. The multiple regression Model-2 in the current study reveals superior fit to the dataset with the studied variables, according to the *F*-test value of 83.619. Also, the results of the *F*-test are supported by the *P*-value (0.000) of the *F*-test, which is significant at the 0.05 level or 95% confidence interval.

Table 10 Model-2: coefficients details

Coefficients[a]						
Model		Coefficients values un-standardized		Coefficients values standardized	*t*	Sig
		B	Std. error	Beta		
2	(Constant)	− 0.694	0.248		− 2.801	0.006
	Employee recruitment (AIER)	0.130	0.043	0.171	2.997	0.003
	Employee training and development (AIETD)	0.341	0.044	0.405	7.701	0.000
	Employee performance evaluation (AIPEPE)	0.357	0.047	0.418	7.518	0.000
	Employee compensation (AIEC)	0.358	0.037	0.502	9.572	0.000

[a]Dependent variable: DM: decision making

Finally, it can be concluded that AI-enabled HRM practices helps in better decision making for the top management. India companies are required to use AI-based applications for decision making in HRM practices: recruitment, training and development, employee performance appraisals or evaluation and employee compensation.

4.3 AI and Ease of Human Resource Management Operations

$H_0 3$ (Null): AI-based application does not ease the human resource management operations in private companies of India (Tables 11, 12, 13 and 14).

$$\text{Model-3}: \ \text{EHRO}_i = \partial 0 + \partial 1\ \text{AIETD}i + \partial 2\ \text{AIER}i + \partial 3\ \text{AIPEPE}i + \partial 4\ \text{AIEC}i + \varepsilon m3i$$

$$\text{Model-3}: \ \text{EHRO}_i = -0.184 + 0.289\text{AIETD}i + 0.033\text{AIER}i + 0.402\ \text{AIPEPE}i + 0.266\ \text{AIEC}i + 0.187$$

The multiple regression Model-3 based on the hypothesis $H_0 3$ finds the relationship between artificial intelligence-based HRM (Employee recruitment, Employee training and development, Employee performance evaluation and Employee compensation) and ease of HR Operations. In Model-3 the artificial intelligence-based HRM aspects are being considered as independent variables and ease in HRM operation being considered as dependent variable. Having *R* value found to be 0.898 is shown

Table 11 Variables details: Model-3

Variables details as entered/removed[a]			
Model	Entered variables	Removed variables	Technique/method
3	Employee recruitment (AIER) Employee training and development (AIETD) Employee performance evaluation (AIPEPE) Employee compensation (AIEC)[b]		Enter

[a]Dependent variable: (EHRO: ease of HR operations)
[b]All requested variables entered

Table 12 Model-3 summary

Model summary				
Model	*R* value	Value of *R*-square	Value of adjusted *R*-square	Standard error of the estimate
3	0.898[a]	0.807	0.800	0.37448

[a]Predictors: (Constant), employee recruitment (AIER), employee training and development (AIETD), employee performance evaluation (AIPEPE), employee compensation (AIEC)

Table 13 Model-3: ANOVA results

ANOVA[a]						
Model		Value of sum of squares	(df) degree of freedom	Mean square value	*F*-value	Sig
3	Regression	67.464	4	16.866	120.268	0.000[b]
	Residual	16.127	115	0.140		
	Total	83.592	119			

[a]Dependent variable: (EHRO: ease of HR operations)
[b]Predictors: (constant), employee recruitment (AIER), employee training and development (AIETD), employee performance evaluation (AIPEPE), employee compensation (AIEC)

in the table model summary. Model-3 implies good correlation between AI-based applications and ease in HRM operations. Similarly, Model-3 statistical results indicate that the model represents the best fit which confirms that the independent or explanatory variables being considered in the study are adequate to present valuable outcomes. The corresponding results convey that as "Model-3" *R*-square value of 0.807 explains that about 80.7% of the data values of the independent variables are adequate to obtain or generate the significant results. Accordingly the ANOVA analysis calculates both *F*-value and *P*-value values. *F*-test value of 120.268 indicates that the multiple regression "Model-3."

Table 14 Model-3: coefficients details

Coefficients[a]						
Model		Coefficients values un-standardized		Coefficients values standardized	*t*	Sig
		B	Std. error	Beta		
3	(Constant)	− 0.184	0.187		− 0.984	0.327
	Employee recruitment (AIER)	0.193	0.033	0.292	5.883	0.000
	Employee training and development (AIETD)	0.289	0.033	0.395	8.643	0.000
	Employee performance evaluation (AIPEPE)	0.298	0.036	0.402	8.324	0.000
	Employee compensation (AIEC)	0.266	0.028	0.428	9.411	0.000

[a]Dependent variable: EHRO: ease of HR operations

Finally, it can be concluded that AI-based application does ease the human resource management operations. Indian companies are required to use more AI-based applications support related to HRM operation or practices such as recruitment, training and development, employee relations, performance evaluation and compensation to increase the organizational effectiveness.

5 Conclusions

There is no doubt that the adoption of AI-based HR activities for individuals has a higher impact on improving the organization's efficiency. Although AI programmes do not have the same emotional or cognitive capabilities as people, they can nonetheless analyse, forecast and diagnose data, making them an effective tool for any type of company. Mainly three hypotheses (H1, H2 and H3) were being analysed in this research work. The outcomes of the hypothesis testing ($H1_0$) indicate that Model-1 represents good fitness, demonstrating that the independent variables utilized in the concern research are sufficient to offer useful conclusions. Therefore, the *R*-square score (0.908) of Model-1 indicates that the study's independent variables explain 90.8% of the data and are adequate to provide meaningful results. Furthermore, the *F*-test *P*-value (0.000) is significant at the 0.05 level or 95% confidence interval, validating the *F*-test results. The hypothesis $H_0 1$ is being rejected which specifies that there is statistically significant impact of artificial intelligence-based human resource practices on the organizational performance of private companies in India.

The outcomes of hypothesis $H_0 2$ also conclude that as P-value < 0.05 which indicates that the null hypothesis is being rejected there is significant impact of artificial intelligence-based human resource practices on the organizational decision-making process of private companies in India. Model-3's R-square score (0.807) indicates that the independent variables explain 80.7% of the data and are adequate to provide meaningful findings. ANOVA statistics include F-value and P-value. The F-test value 120.268 suggests that the multiple regression Model-3 in this investigation fits the dataset better with the study variables. The F-test P-value (0.000) is significant at 0.05 level or 95% confidence interval, validating its results. Based on the above discussion it can be concluded that the null hypothesis $H_0 3$ is being rejected and which indicates that AI-based application does ease the human resource management operations in private companies of India.

References

1. Buzko I, Dyachenko Y, Petrova M, Nenkov N, Tuleninova D, Koeva K (2016) Artificial intelligence technologies in human resource development. Comput Model New Technol 20(2):26–29
2. De Geofroy Z, Evans MM (2017) Are emotionally intelligent employees less likely to hide their knowledge? Knowl Process Manag 24(2):81–95
3. Jain DS (2018) Human resource management and artificial intelligence. Int J Manag Soc Sci Res 7(3):56–59
4. Jarrahi MH (2018) Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making. Bus Horiz 61(4):577–586
5. O'Connor SW (2020) Artificial intelligence in human resource management. North Eastern University, New York
6. Wisetsri W, Vijai C, Chueinwittaya K, Jirayus P (2022) Artificial intelligence in human resources management: an overview, pp 2688–2693
7. Tiwari P, Pandey R, Garg V, Singhal A (2021) Application of artificial intelligence in human resource management practices, pp 159–163. https://doi.org/10.1109/Confluence51648.2021.9377160
8. Tsymbaliuk S, Vasylyk A, Bilyk O (2022) Artificial intelligence in HR: practices and prospects of the spread in Ukraine. Eurasia Proc Sci Technol Eng Math 17:152–160
9. Gaidhani S (2020) The role of artificial intelligence in automation of HR practices
10. Sanyaolu E, Atsaboghena R (2022) Role of artificial intelligence in human resource management: overview of its benefits and challenges. https://doi.org/10.13140/RG.2.2.22297.29283
11. Amla M, Malhotra PM (2017) Digital transformation in HR. Int J Interdiscip Multidiscip Stud 4(3):536–544

12. Jia Q, Guo Y, Li R, Li Y, Chen Y (2018) A conceptual artificial intelligence application framework in human resource management. In: ICEB 2018 proceedings, p 91
13. Hislop D, Coombs C, Taneva S, Barnard S (2017) Impact of artificial intelligence, robotics and automation on work
14. Rajesh DS, Kandaswamy MU, Rakesh MA (2018) The impact of artificial intelligence in talent acquisition lifecycle of organizations. Int J Eng Develop Res 6(2):709–717
15. Jain S (2017) Is artificial intelligence the next big thing in hr? In: International conference on innovative research in science, technology and management

Artificial Intelligence in Finance and Accounting: Opportunities and Challenges

Pushpkant Shakdwipee, Khushbu Agarwal, Hemlata Kunwar, and Shivoham Singh

Abstract Artificial intelligence (AI) is revolutionizing the finance and accounting industry by offering numerous opportunities for increased efficiency, improved decision-making, and enhanced customer experience. AI can automate tedious and time-consuming tasks, provide accurate insights and predictions, and help identify patterns and trends in large datasets. However, the adoption of AI in finance and accounting also presents several challenges, including issues related to data quality, bias, lack of transparency, privacy, regulatory compliance, ethics, and expertise. The integration with legacy systems, reliance on third-party vendors, cost, scalability, and workforce impact are also significant challenges that must be addressed. To fully leverage the benefits of AI in finance and accounting, businesses must address these challenges and implement AI solutions responsibly and ethically. By doing so, they can gain a competitive advantage, improve operational efficiency, and deliver better value to customers. However, several factors, including trust in AI, regulatory environment, availability of data, and cost, could impact the adoption of AI in finance and accounting. By addressing these challenges and factors, businesses can unlock the full potential of AI and gain a competitive advantage in the industry.

Keywords Artificial intelligence · Finance · Accounting · Opportunities · Challenges · Big data · Automation

P. Shakdwipee · K. Agarwal · H. Kunwar · S. Singh (✉)
PAHER University, Udaipur, Rajasthan, India
e-mail: shivohamsingh@gmail.com

S. Singh
Symbiosis Institute of Business Management, Hyderabad (SIBM-H), Hyderabad, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_17

1 Introduction

Artificial intelligence (AI) has emerged as a transformative technology that has the potential to revolutionize various industries. The finance and accounting industry is no exception, and AI has already started to transform how financial services are delivered, managed, and regulated.

In the finance and accounting industry, AI technology can be used for a range of tasks such as automating repetitive tasks, predicting market trends, detecting fraudulent activities, and making investment decisions. With AI-powered applications, financial institutions can automate tasks such as risk assessment, credit scoring, and underwriting, thereby improving operational efficiency and reducing costs.

One of the significant benefits of AI technology in finance and accounting is that it can analyze large volumes of data in real-time, which can help financial institutions to make informed decisions. AI can help detect fraudulent activities such as money laundering and insider trading by analyzing transaction patterns and identifying anomalies. It can also assist in identifying investment opportunities by analyzing market trends, news, and social media sentiment.

However, the adoption of AI technology in the finance and accounting industry also presents some challenges. One of the biggest challenges is data privacy and security, which is critical in the finance industry. As AI algorithms rely on large amounts of data to make decisions, there is a risk that sensitive financial information may be exposed to cyberattacks or breaches.

Another challenge is the lack of transparency in AI algorithms. AI algorithms are often considered black boxes, which means that it is challenging to understand how they arrive at certain decisions. This lack of transparency can make it difficult for financial institutions to comply with regulatory requirements and explain their decision-making processes to stakeholders.

AI technology can also help financial institutions to improve customer experience and engagement by personalizing services and offering personalized financial advice. For example, chatbots powered by AI can provide customers with quick and personalized responses to their queries and concerns, thereby improving customer satisfaction.

Furthermore, AI can assist in compliance and regulatory reporting, which is a critical aspect of the finance and accounting industry. Financial institutions are required to comply with a wide range of regulations, and AI-powered applications can help automate compliance tasks and reduce the risk of non-compliance.

Another area where AI can make a significant impact in the finance and accounting industry is in asset management. AI algorithms can analyze vast amounts of data to identify investment opportunities and make predictions about market trends. This can help asset managers to make better investment decisions and improve portfolio performance.

However, the adoption of AI technology in the finance and accounting industry requires significant investment in both technology and talent. Financial institutions need to invest in the development and deployment of AI-powered applications, as

well as in training their employees to work with these applications. This can be a significant challenge, especially for smaller financial institutions that may not have the resources to invest in AI.

AI technology presents significant opportunities for the finance and accounting industry. However, it also presents challenges that need to be addressed to fully realize its potential. Financial institutions need to balance the benefits of AI with the challenges associated with its adoption to ensure that they can leverage this technology to improve their operations, enhance customer experience, and achieve better business outcomes.

2 Review of Literature

2.1 Global Perspective

Here is a brief review of literature on the global perspective of artificial intelligence in finance and accounting, along with relevant citations and references:

According to a report by the World Economic Forum, AI has the potential to transform the financial services industry by improving efficiency, reducing costs, and enhancing customer experience. The report highlights the importance of balancing innovation with regulatory compliance and ethical considerations [20].

A study by PwC found that AI could add $15.7 trillion to the global GDP by 2030, with the finance and insurance industry being one of the most significant beneficiaries. The report also highlights the need for businesses to invest in AI technology and talent to remain competitive [18].

A survey conducted by Deloitte found that 76% of financial institutions believe that AI will significantly transform the industry in the next three years. The report highlights the importance of addressing challenges related to data quality, transparency, and governance to fully realize the potential of AI [7].

A study by McKinsey and Company found that AI has the potential to generate significant value in the finance and accounting industry by automating manual processes, improving decision-making, and enhancing customer experience. The report emphasizes the need for financial institutions to invest in building the right infrastructure, talent, and partnerships to leverage the benefits of AI [13].

A research paper by BNM Institute of Technology explores the use of machine learning and AI in financial fraud detection. The study highlights the potential of AI in detecting fraudulent activities in real-time by analyzing transaction data and identifying anomalies [10].

A research paper by the University of Sussex examines the ethical and legal implications of AI in financial services. The study highlights the need for transparency and accountability in AI algorithms and emphasizes the importance of ensuring that AI is used to serve the best interests of customers and society [2].

A research paper by the University of Bristol explores the impact of AI on the audit profession. The study highlights the potential of AI in improving audit quality, reducing costs, and enhancing the efficiency of the audit process. The paper also identifies challenges related to the interpretation of results and the need for human intervention in complex cases [17].

A research paper by the University of Edinburgh examines the use of AI in credit risk assessment. The study finds that AI algorithms can improve the accuracy of credit risk models by analyzing non-traditional data sources and detecting patterns that traditional models may miss. The paper also highlights the need for transparency and explains ability in AI models to ensure they are fair and unbiased [3].

A research paper by the University of Oxford explores the use of AI in regulatory compliance. The study finds that AI algorithms can assist in identifying compliance risks, automating compliance tasks, and improving the efficiency of compliance processes. The paper also highlights challenges related to data privacy, explain ability, and accountability [1].

A research paper by the University of Cambridge examines the use of AI in financial forecasting. The study finds that AI algorithms can improve the accuracy of financial forecasts by analyzing large datasets and identifying patterns and trends. The paper also highlights the importance of ensuring that AI models are transparent and explainable to stakeholders [11].

2.2 Indian Perspective

Here is a brief review of literature on the Indian perspective of artificial intelligence in finance and accounting, along with relevant citations and references:

A report by the National Association of Software and Services Companies (NASSCOM) highlights the potential of AI in transforming the Indian financial services industry. The report identifies opportunities for AI in areas such as customer experience, risk management, fraud detection, and compliance [16].

A research paper by the Indian Institute of Technology (IIT) Delhi examines the use of AI in credit risk assessment. The study finds that AI algorithms can improve the accuracy of credit risk models by analyzing non-traditional data sources and detecting patterns that traditional models may miss. The paper also highlights challenges related to data privacy, explain ability, and regulatory compliance [12].

A research paper by the Indian Institute of Management Bangalore (IIMB) explores the use of AI in financial forecasting. The study finds that AI algorithms can improve the accuracy of financial forecasts by analyzing large datasets and identifying patterns and trends. The paper also highlights the importance of ensuring that AI models are transparent and explainable to stakeholders [19].

A research paper by the Indian Institute of Technology (IIT) Roorkee examines the use of AI in fraud detection in the Indian banking sector. The study finds that AI algorithms can assist in detecting fraudulent activities by analyzing transaction data and identifying anomalies. The paper also highlights the need for data privacy and ethical considerations in the use of AI for fraud detection [5].

A research paper by the Institute for Development and Research in Banking Technology (IDRBT) explores the use of blockchain and AI in the Indian banking sector. The study finds that the combination of blockchain and AI can improve the efficiency, transparency, and security of financial transactions. The paper also highlights challenges related to regulatory compliance and standardization [14].

A research paper by the Indian Institute of Technology (IIT) Delhi examines the use of machine learning in fraud detection in the Indian insurance industry. The study finds that machine learning algorithms can improve the accuracy of fraud detection by analyzing historical claims data and identifying patterns. The paper also highlights challenges related to data quality and interpretability [14].

A report by Deloitte India highlights the potential of AI in transforming the Indian banking sector. The report identifies opportunities for AI in areas such as customer service, risk management, and compliance. The report also highlights the importance of building a robust data infrastructure and addressing ethical and regulatory concerns [8].

A research paper by the Indian School of Business (ISB) explores the use of AI in financial reporting. The study finds that AI algorithms can improve the speed and accuracy of financial reporting by automating tasks such as data extraction and analysis. The paper also highlights challenges related to data quality and the need for human oversight [4].

A report by the Confederation of Indian Industry (CII) examines the potential of AI in the Indian accounting sector. The report identifies opportunities for AI in areas such as audit, tax, and financial reporting. The report also highlights the importance of developing a skilled workforce and addressing ethical and regulatory concerns [4].

3 Opportunities of Artificial Intelligence in Finance and Accounting

Artificial intelligence (AI) offers a wide range of opportunities in finance and accounting. Here are some of the key opportunities:

3.1 Improved Efficiency

AI can automate repetitive and time-consuming tasks such as data entry, account reconciliation, and financial reporting. This can free up staff to focus on more complex tasks and enable faster decision-making.

3.2 Enhanced Accuracy

AI can analyze large datasets and identify patterns and trends that may be missed by human analysts. This can improve the accuracy of financial forecasting, risk management, and fraud detection.

3.3 Better Customer Service

AI-powered chatbots and virtual assistants can provide 24/7 customer support and personalized recommendations based on customer data.

3.4 Cost Savings

By automating tasks and reducing errors, AI can help to reduce costs associated with manual processing and compliance.

3.5 Increased Compliance

AI can help to identify and prevent fraudulent activities, ensure regulatory compliance, and monitor transactions for suspicious behavior.

3.6 Improved Risk Management

AI can analyze data in real-time to identify potential risks and alert stakeholders to potential issues before they escalate.

3.7 New Revenue Streams

AI can help to identify new business opportunities and revenue streams by analyzing customer data and market trends.

3.8 Advanced Data Analytics

AI can analyze large datasets and identify patterns that can help businesses to make informed decisions. With the help of machine learning algorithms, businesses can identify market trends and patterns that can help them to predict customer behavior and make strategic decisions.

3.9 Predictive Maintenance

In the accounting and finance industry, predictive maintenance can help businesses detect problems before they occur. For example, predictive maintenance can help financial institutions monitor their hardware, software, and network components for signs of wear and tear, allowing them to replace components before they fail.

3.10 Fraud Detection

AI algorithms can identify patterns of fraudulent activities and alert businesses before they become a major issue. Machine learning algorithms can analyze financial transactions and flag any suspicious activities that deviate from normal spending patterns, allowing financial institutions to take appropriate action.

3.11 Investment Recommendations

AI algorithms can help financial advisors make informed investment recommendations based on a customer's investment history and risk tolerance. Machine learning algorithms can analyze past investments and market trends to help advisors make more informed recommendations.

3.12 Intelligent Automation

In the finance and accounting industry, AI-powered intelligent automation can help businesses automate their processes, reduce errors, and increase efficiency. For example, intelligent automation can help businesses automate account reconciliation, invoice processing, and other manual tasks.

3.13 Personalization

AI can help businesses personalize their products and services based on customer data. For example, banks can use AI algorithms to analyze a customer's spending habits and offer customized financial products that match their needs.

3.14 Real-Time Decision-Making

AI can process vast amounts of data in real-time, allowing businesses to make informed decisions quickly. For instance, AI algorithms can help traders to identify opportunities in the market and execute trades in real-time.

3.15 Improved Credit Risk Assessment

AI can analyze a wide range of data sources to help businesses assess the creditworthiness of customers. By analyzing data such as credit history, employment history, and social media activity, AI algorithms can provide more accurate credit risk assessments.

3.16 Portfolio Management

AI algorithms can analyze market trends and customer data to help businesses manage their investment portfolios more effectively. By providing real-time insights into market trends, AI can help businesses make informed decisions and minimize risk.

3.17 Regulatory Compliance

AI can help businesses stay compliant with regulatory requirements by analyzing data in real-time and identifying potential compliance issues. For instance, AI algorithms can monitor transactions for suspicious activities and alert businesses to potential compliance issues.

3.18 Cybersecurity

AI can help businesses detect and prevent cyberthreats by analyzing network data and identifying potential vulnerabilities. By analyzing data in real-time, AI algorithms can identify potential threats before they cause significant damage.

Overall, AI offers a wide range of opportunities in finance and accounting, from personalization and real-time decision-making to improved credit risk assessment and regulatory compliance. However, businesses must address challenges such as data quality, privacy, and ethical considerations to ensure that AI is used responsibly and effectively.

4 Challenges of Artificial Intelligence in Finance and Accounting

While AI offers a multitude of opportunities in finance and accounting, it also presents a number of challenges. Here are some of the key challenges:

4.1 Data Quality

AI algorithms rely on high-quality data to provide accurate insights and recommendations. However, data quality can be a major challenge in the finance and accounting industry, as data may be scattered across different systems and sources and may be incomplete or inaccurate.

4.2 Bias

AI algorithms can be biased if they are trained on biased data. In the finance and accounting industry, bias can occur if data are skewed toward certain demographics or if data are incomplete.

4.3 Lack of Transparency

Some AI algorithms are opaque, meaning that it is difficult to understand how they arrive at their recommendations. This lack of transparency can make it difficult for stakeholders to understand how decisions are being made.

4.4 Privacy Concerns

AI algorithms may require access to sensitive data such as financial records, which can raise privacy concerns. It is important to ensure that data are protected and that privacy regulations are adhered to.

4.5 Regulatory Compliance

The use of AI in finance and accounting must comply with various regulatory requirements, such as data protection regulations and anti-money laundering laws. Ensuring compliance can be challenging, especially as regulations continue to evolve.

4.6 Ethical Considerations

The use of AI in finance and accounting raises ethical considerations, such as whether AI should be used to make decisions that affect people's livelihoods. It is important to ensure that AI is used ethically and responsibly.

4.7 Lack of Expertise

Developing and implementing AI solutions require expertise in areas such as data science, machine learning, and software development. However, there is currently a shortage of professionals with these skills, making it difficult for businesses to develop and implement AI solutions.

4.8 Integration with Legacy Systems

Many businesses in the finance and accounting industry still rely on legacy systems that may not be compatible with new AI solutions. Integrating AI with legacy systems can be challenging and may require significant investment in infrastructure.

4.9 Reliance on Third-Party Vendors

Some businesses may rely on third-party vendors to develop and implement AI solutions. This can create additional challenges, such as data security and intellectual property rights.

4.10 Cost

Implementing AI solutions can be expensive, particularly for smaller businesses with limited budgets. The cost of implementing AI can include the development of custom algorithms, the cost of hardware and software, and ongoing maintenance and updates.

4.11 Scalability

As businesses grow, they may need to scale their AI solutions to accommodate larger volumes of data and users. Ensuring that AI solutions can scale effectively can be challenging and may require additional investment in infrastructure and resources.

4.12 Workforce Impact

The use of AI in finance and accounting may have an impact on the workforce, particularly in areas such as data entry and analysis. Businesses must consider the impact of AI on their employees and may need to provide training and support to help employees adapt to new technologies.

While the opportunities of AI in finance and accounting are numerous, there are also a number of challenges that must be addressed to ensure that AI is used effectively and responsibly. By addressing these challenges, businesses can gain the benefits of AI while mitigating the associated risks.

5 Conclusion

Artificial intelligence (AI) has the potential to transform the finance and accounting industry, offering numerous opportunities for increased efficiency, improved decision-making, and enhanced customer experience. AI can automate tedious and time-consuming tasks, provide accurate insights and predictions, and help identify patterns and trends in large datasets.

However, the adoption of AI in finance and accounting also presents several challenges. These include issues related to data quality, bias, lack of transparency, privacy, regulatory compliance, ethics, and expertise. Additionally, the integration with legacy systems, reliance on third-party vendors, cost, scalability, and workforce impact are also significant challenges that must be addressed.

To fully leverage the benefits of AI in finance and accounting, it is important for businesses to address these challenges and implement AI solutions responsibly and ethically. By doing so, they can gain a competitive advantage, improve operational efficiency, and deliver better value to customers.

One of the factors is the level of trust in AI. While AI has the potential to provide accurate insights and predictions, there is also a risk of errors and bias. To build trust in AI, businesses must be transparent about how AI algorithms are developed and how they arrive at their recommendations.

Another factor is the regulatory environment. As AI becomes more prevalent in finance and accounting, regulatory bodies are likely to introduce new regulations and guidelines to ensure that AI is used responsibly and ethically. Businesses must keep up-to-date with regulatory developments and ensure that their AI solutions comply with the relevant regulations.

The availability of data is also an important factor. AI relies on large volumes of high-quality data to provide accurate insights and recommendations. However, the availability of data can be limited, particularly in emerging markets or in industries where data are not readily available.

Finally, the cost of implementing AI solutions could be a significant factor, particularly for smaller businesses with limited budgets. While AI can provide significant benefits in terms of efficiency and cost savings, the upfront investment required to implement AI solutions could be a barrier for some businesses.

Overall, while the opportunities of AI in finance and accounting are significant, businesses must be aware of the challenges and factors that could impact the adoption of AI. By addressing these challenges and factors, businesses can unlock the full potential of AI and gain a competitive advantage in the industry.

References

1. Arner DW, Barberis JN, Buckley RP (2019) The evolution of FinTech: A new post-crisis paradigm? Georgetown J Int Law 50(4):1149–1215
2. Barral M, Klass L, Floridi L (2020) The ethics and law of artificial intelligence in financial services. J Bus Ethics 166(4):723–737
3. Bień K, Bognár T (2021) Exploring credit risk assessment with machine learning: A review of methods and fair lending considerations. J Bank Finance 129:104–118
4. Chakrabarti R, Chawla K, Subramanian R (2020) AI in financial reporting: Enabling the shift from compliance to insights. Indian school of business
5. Chaudhary A, Bhatnagar V (2019) Fraud detection in Indian banking sector: Using machine learning and analytics. J Financial Crime 26(4):1114–1127
6. CII (2019) Artificial intelligence in accounting: A CII perspective
7. Deloitte (2020) Financial services industry outlooks: Insights from the global financial services industry 2020 survey
8. Deloitte (2020) Future of banking in India: Artificial intelligence transforming the future of banking
9. IDRBT (2020) Blockchain and artificial intelligence: Opportunities and challenges for Indian banking
10. Jaswal A, Choudhary N, Agarwal P, Sinha N (2019) Fraud detection in financial systems using machine learning. Int J Comput Sci Inf Secur 17(12):12–19
11. Krauss C, Do XA (2020) Artificial intelligence and financial forecasting: A review. Int J Forecast 36(1):223–247
12. Majumdar A, Mehta N, Rao B (2019) Credit risk assessment using machine learning: A review. Decision 46(4):313–328
13. McKinsey & Company (2019) The future of risk management in the digital era
14. Mishra RK, Shetty R (2020) Machine learning in fraud detection in Indian insurance industry. Int J Comput Appl 175(1):22–30
15. Mishra RK, Sahu S, Biswas S (2020) Blockchain and artificial intelligence in Indian banking sector: Opportunities and challenges. J Public Aff 20(2):2042
16. NASSCOM (2020) Artificial intelligence and its impact on the Indian financial services industry
17. Papoutsaki E, Iatridis GE, Shailer G (2021) Audit in the age of artificial intelligence: The future of auditing. J Bus Res 130:644–652
18. PwC (2018) Sizing the prize: What's the real value of AI for your business and how can you capitalize?
19. Sharma P, Kodali R (2020) Artificial intelligence for financial forecasting in India: Issues and challenges. Int J Emerg Mark 15(5):902–917
20. World Economic Forum. (2020). The future of financial services: How disruptive innovations are reshaping the way financial services are structured, provisioned, and consumed.

Available Bandwidth Estimation in MANET Using FPECM-MFL-GRRSU for Adaptive Video Streaming

Sanjay Agal

Abstract There is a huge challenge to stream high-quality video from home media sources to TV sets over a wireless network. The data will suffer delay as well as packet loss if the bandwidth of the video exceeds the available bandwidth of a network. Consequently, for ensuring the quality of service (QoS) of real-time multimedia communication, the measurement of available bandwidth is of great importance. Hence, a DL-based available bandwidth estimation (BWE) approach utilizing fuzzy probability entropy C-means (FPECM)-Mamdani fuzzy logic (MFL) and gated recurrent radial secant unit (GRRSU) for adaptive video streaming services in mobile ad-hoc networks (MANET) is proposed in this work. The proposed work begins by initializing the nodes. Grounded on the distance, the initialized nodes are clustered utilizing the FPECM algorithm. After that, by using the MFL technique, routing is done to find the routes and the optimal routes are chosen by utilizing the Ushiki self-awareness wild horse optimizer (USA-WHO). Test packet transmission is executed and the features of the data packet as well as the corresponding path are extracted. For dimensionality reduction, these features are given to the linear information discriminant analysis (LIDA). Consequently, to train the network for BWE, the diminished features are provided to the GRRSU. The proposed method is compared and found to be more efficient than the existing methods in the experimental evaluation.

Keywords Bandwidth estimation (BWE) · Mobile ad-hoc networks (MANETs) · Linear information discriminant analysis (LIDA) · Gated recurrent radial secant unit (GRRSU) · Fuzzy probability entropy C-means (FPECM)

1 Introduction

Independent mobile nodes, which are randomly deployed as well as can join or leave the network on the move, are comprised in MANETs. In MANET, to coordinate the communication or to carry data betwixt nodes, nodes do not depend on

S. Agal (✉)
Dr. V R Godhania College of Engineering and Technology, Porbandar, Gujarat, India
e-mail: sanjayagal@yahoo.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_18

a central node; rather than that, they work together to carry data betwixt nodes that cannot reach each other directly [1] 1,918,802:27,946,431. The provision of a QoS is required by the huge demand for multimedia transmission services. Since MANETs suffer as of external interference as well as internal failure like burst traffic, process failures, link failure, heavy traffic, and battery failure, it is very difficult to guarantee QoS in MANETs nowadays [2] 1,918,802:27,946,423. An efficient routing scheme is needed to meet the demands of design constraints of QoS-enabled MANETs. Routing offers ways for appropriate path selection within the network [3] 1,918,802:27,946,425. Communication betwixt routers is offered by the routing protocol, which also processes the data packet as of source to destination by choosing the suitable path betwixt the sender and receiver (Akhtar et al. 2019) 1,918,802:27,946,442. Most prevailing research works on routing protocols in wireless MANETs focus attention on how to cope with the nodes' mobility, scalability, as well as selecting a minimal path for hop count. However, while taking into account lossy wireless or crowded pathways, it is highly desirable to search for a reliable route that assures high-quality multimedia streaming [4] 1,918,802:27,946,439. Additionally, there are still difficulties for offering satisfactory levels of QoS to end users, which chiefly incorporate bandwidth availability.

An important role was played by BWE with the rapid growth of the Internet in the growing demand for multimedia services like video conferences, streaming video on demand, etc. A good end-to-end transport performance and QoS provisioning in network infrastructure should be contained in the streaming video applications [5] 1,918,802:27,946,434. By assessing the throughput of a path or link devoid of interrupting prevailing flows, the available bandwidth is quantified [6] 1,918,802:27,946,427. During the last few years, several end-to-end BWE algorithms are developed by researchers. These techniques could be fragmented into '2' categories: active probing and passive measurements. The required information as of prevailing data transmissions is inferred by the latter; hence, these methods can be implemented only for network paths that carried traffic recently. On the other hand, Active probing techniques inject end-to-end test traffic into the network. It also evaluates the available bandwidth by observing the cross-traffic's impact on probe packets [7] 1,918,802:27,946,424. As the bandwidth variation causes high delay, jitter, along with packet loss that degrades the performance of streaming video services, real-time video streaming typically imposes stringent bandwidth requirements on the underlying network [8] 1,918,802:27,946,428. Therefore, a novel deep learning-based available BWE approach using FPECM-MFL and GRRSU for adaptive video streaming services in MANET is proposed in this paper.

1.1 Problem Definition

Prevailing research methodologies have some shortcomings, which are described as follows:

- One among the major challenges in bandwidth measurement research is the fact that the presented bandwidth varies as time passes. This is the outcome of cross-traffic's inconsistency over time.
- Probe packet transmission based on available BWE is used by most of the prevailing techniques. The transmission of such packets betwixt diverse nodes and multiple paths led to augmented packet loss and diminished estimation accuracy with elevated time.
- Deep learning (DL)-centered available BWE, incorporating QoS factors like energy-aware route selection, were not regarded in most of the prevailing research works.

Thus, this paper's main objective is to propose a DL-centered available BWE approach utilizing FPECM-MFL and GRRSU for adaptive video streaming services in MANET. The contributions of the proposed work are:

- To overcome the inconsistency of cross traffic over time, an energy-efficient novel FPECM clustering is proposed.
- To choose the optimal paths for efficient transmission of packets, a proficient USA-WHO algorithm is developed.
- To propose an efficient available BWE mechanism by adopting test packet transmission theory to extract the available features and diminish the dimensionality utilizing LIDA.
- DL-centered available BWE scheme is introduced by utilizing a GRRSU classifier.

This Paper structure is organized as: Sect. 2 analyzes Proposed available BWE Framework and Sect. 3 discusses the result and discussion and analyze the performance of the proposed methodologies and finally Sect. 4 the paper with a conclusion.

2 Proposed Available BWE Framework

A novel DL-based available BWE approach utilizing FPECM-MFL and GRRSU for adaptive video streaming services in MANET is proposed in this work. The proposed work includes the following phases, node initialization, clustering, route discovery, optimal route selection, test packet transmission, feature extraction (FE), dimensionality reduction, and finally, bandwidth estimation using the GRRSU model. Figure 1 exhibits the proposed model's block diagram.

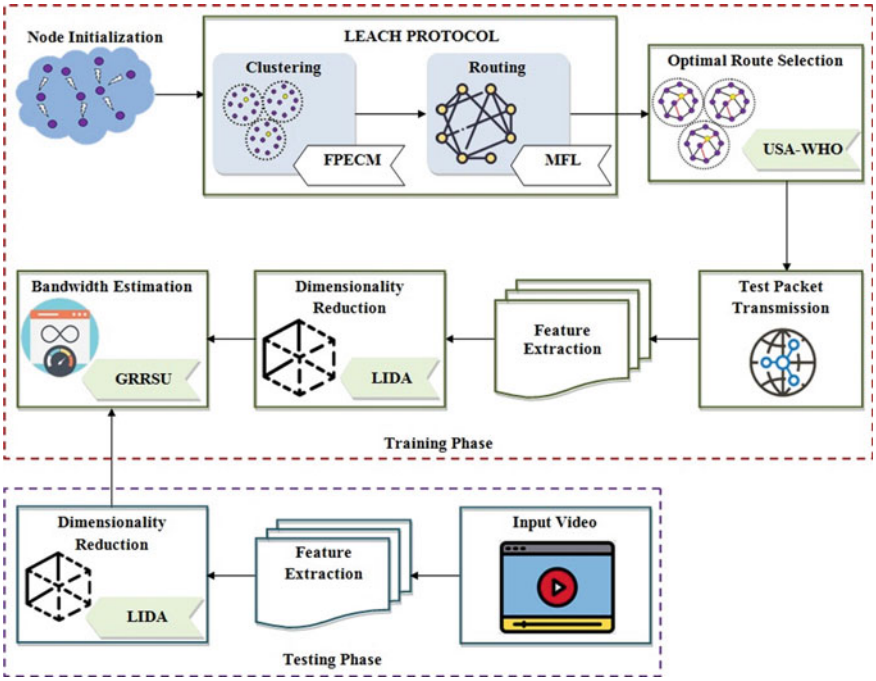


Fig. 1 Proposed methodology's block diagram

2.1 Node Initialization

Initially, the number of nodes needed for communication is randomly initialized by deploying them in the geographical location. The nodes' initialization can be defined as:

Where, the total number of nodes initialized is notated as.

2.2 LEACH Protocol

Subsequent to initialization, grounded on the low-energy adaptive clustering hierarchy (LEACH) protocol, the randomly deployed nodes are grouped. LEACH is an integrated technique with a clustering technique and a routing protocol in MANET. It normally has '2' phases, namely the set-up phase and the steady-state phase (SSP). Clustering is executed in the set-up phase. The communication in the network is established betwixt the source and destination in the SSP. The LEACH protocol's set-up phase and SSP have been modified utilizing FPECM and MFL techniques, respectively, in this paper.

Clustering here, by means of the FPECM clustering model, the randomly deployed nodes.

FPECM algorithm operates by assigning membership to each node in Where, the total number of the cluster center is signified as

Where, the membership of the

The main goal of FPECM is to diminish the objective function regarding the centers and memberships. The objective function

Until the minimum

Where, the final cluster set is signified as

Algorithm 1: FPECM Technique.

Routing The routing protocol is executed with the aid of the MFL technique after the nodes are grouped. To identify the possible routes betwixt the source and the destination via the clusters.

Fuzzification, rule generation, and defuzzification are the '3' steps under which the fuzzy system works. Initially, the training data is fed into the fuzzy control system and mapped by sets of Gaussian membership functions named as fuzzy sets. Fuzzification is defined as the process of altering the input value to a fuzzy value. The fuzzification function can be articulated as,

Where, the Gaussian membership function is signified as Where, the membership function's center is notated as

Where, the minimum number of membership values are notated as

To offer the fuzzy set's most accurate illustration, the inference engine's output is then mapped. Finally, defuzzification is executed, which is explained by the following expression

Where, the generated rules are signified as

2.3 Optimal Route Selection

After discovering the routes.

The USA-WHO algorithm contains '4' main steps, which are explained as follows:

Initial Population and Group Formation: USA-WHO begins by initializing the population randomly. The initialized population is articulated as,

This random population is evaluated repeatedly by an objective function. For selecting the best solution (i.e. optimal path), the minimum distance betwixt the source and destination is considered as the objective function. After that, the random population is partitioned into groups. The number of groups

Where, the stallions' percentage in the total population is signified as Where, the Stallion population is notated as

Grazing Behavior: The stallions are placed in the grazing area's center in the phase of the grazing behavior and the group members search around the leader. This behavior is formulated as,

Where, the group member's current and new grazing position are signified as

Where,

Where, the current iteration is specified as

Mating Behavior: Horse mating behavior is such that foals leave their group before they reach puberty. They join other groups and meet their partner. This is to prevent the father from mating with the daughter or siblings. This mating behavior is explicated by a crossover function

Where,

Group leadership: The leader

Where, the leader's next and current position are signified as Where, the fitness value is illustrated as

Algorithm 2: USA-WHO Technique.

2.4 Test Packet Transmission

After discovering the optimal paths, test packet transmission is executed. Here, the data packets are transmitted as of the source to the destination via each optimal path.

2.5 Feature Extraction

Here, the features of each optimal path as well as the data packets are extracted. The extracted feature set,

Where, the path features are signified as.

2.6 Dimensionality Reduction

Here, by utilizing the LIDA technique, dimensionality reduction is performed in the extracted features.

Finding a linear combination of features that maximize the separation betwixt diverse classes as well as minimize the distance betwixt the similar class is the main idea of LIDA. Primarily, by computing the mean class of the features

Where, the prior probability is notated as

Where,

Where, the number of training samples in class

Where, the first Eigenvector is notated as

2.7 Classifier

Here, to train the system with the features of the data packets and path for BWE, the dimensionality-reduced features.

In GRU, each gate takes '2' inputs at each time, which incorporates the present input as well as the earlier hidden state as vectors. By executing element-wise multiplication betwixt the concerned vector and the respective weights for every gate, the output of each gate is obtained. By using with radial kernel-based weight function, the weights (Fig. 2),

Where, the Euclidean distance betwixt the vectors of features

Reset Gate: The reset gate is calculated by a linear sum betwixt the newly computed state and the prevailing state with the bias parameter. It can be articulated as,

Where,

Update Gate: The update gate determines how much of the prior information as of the earlier time

Where, the update gate is signified as

To pass the relevant information, the current memory content Where, the hyperbolic tangent function is specified as

2.8 Testing Phase

Initially, the data packets (videos) undergo FE in the testing phase. After that, for dimensionality reduction, the features are fed to the LIDA technique. Next, the diminished features are fed as input to the trained GRRSU classifier. The classifier estimates the optimal paths and available bandwidth in that path through which the data transmission betwixt source and destination can occur. Therefore, the BWE was performed effectively by the proposed model. By analogizing the proposed model with some of the prevailing techniques in the following section, the proposed model's efficacy is verified.

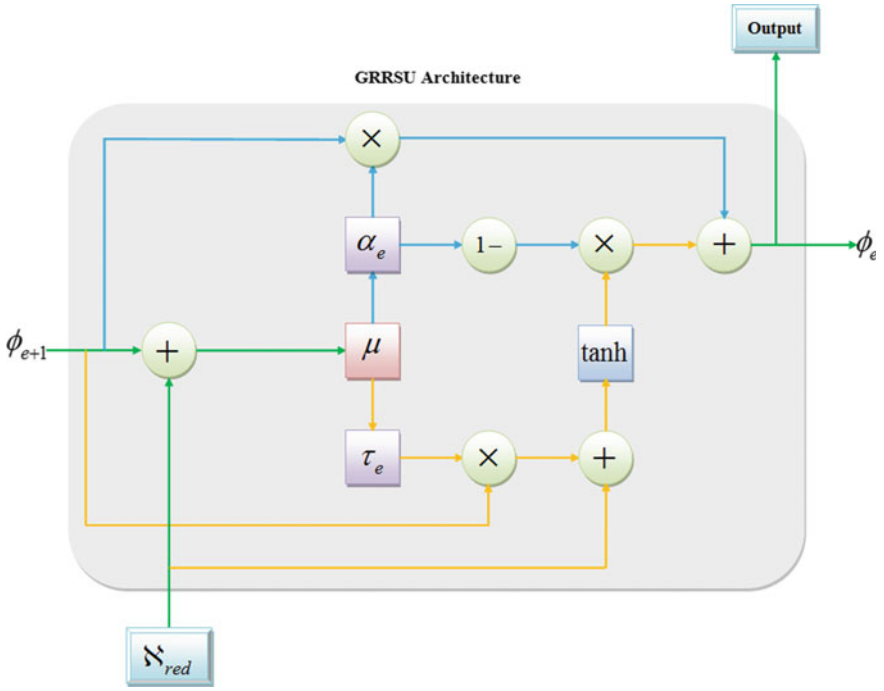


Fig. 2 Architecture of GRRSU

3 Results and Discussion

By comparing the outcomes with some of the prevailing methodologies, the experimental outcomes of the proposed framework are analyzed in this section. MATLAB is the working platform in which the proposed methodology is implemented.

3.1 Performance Measure of the Proposed FPECM

By analyzing the clustering time with some of the existing techniques such as FCM, K-means, and hierarchical clustering (HC), the proposed available BWE technique's performance is verified.

Based on the time consumed for clustering, the proposed and existing model's performance analysis is illustrated in Table 1. To prove the efficiency of a clustering technique, the time consumed for clustering must be low. The clustering time of the proposed FPECM method is 0.052321 s. Conversely, the clustering time increases with a difference of 0.1539, 0.1039, and 0.1060 s for the existing methods like

Table 1 Performance of the proposed FPECM method

Techniques	Clustering time (s)
Proposed FPECM	0.052321
FCM	0.206221
K-means	0.156304
HC	0.158417

FCM, K-means, and HC. Thus, the proposed method's performance is better when analogized with prevailing methods.

3.2 Performance Analysis of USA-WHO

The proposed USA-WHO model's performance is analogized with the prevailing techniques like WHO, grasshopper optimization (GO), PSO, as well as Salp Swarm optimization (SSO) based on the convergence.

The proposed USA-WHO and the prevailing optimization algorithms' convergence speed are analyzed in Fig. 3. By attaining convergence in 20 iterations, the proposed USA-WHO algorithm has exhibited better performance. However, the prevailing models consumed more iterations for attaining convergence. This exhibits the proposed USA-WHO model's superiority. In WHO, the performance was enhanced to a greater extent by the modification of the Ushiki chaotic mapping and self-awareness probability factor.

3.3 Performance Analysis of Classifier

Based on the quality metrics like sensitivity, specificity, F-measure, accuracy, precision, recall, net present value (NPV), false recognition rate (FRR), false positive rate (FPR), false negative rate (FNR), and Mathews correlation coefficient (MCC), the proposed GRRSU classifier's performance is analyzed with the prevailing methods like GRU, deep neural networks (DNN), artificial neural networks (ANN), as well as adaptive neuro-fuzzy inference systems (ANFIS).

Based on sensitivity, specificity, accuracy, and precision, the performance of the proposed and existing methods are analyzed in Fig. 4. The proposed method attained similar sensitivity, specificity, accuracy, and precision value of 96%, whereas the prevailing GRU, ANN, DNN, and ANFIS methods offer similar values of 94.5, 77.5, 93, and 77.5%, respectively. The proposed classifier's performance was enhanced to a greater extent by the modification of the secant activation function and radial kernel-based weight assignment in GRU. It is clear from the attained analysis results that the proposed method performs better regarding BWE.

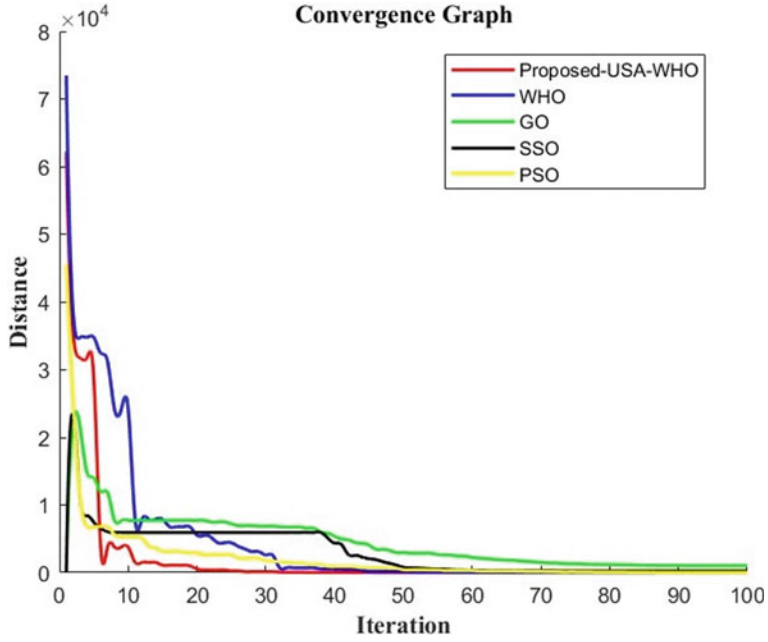


Fig. 3 Performance analysis of the proposed USA-WHO algorithm

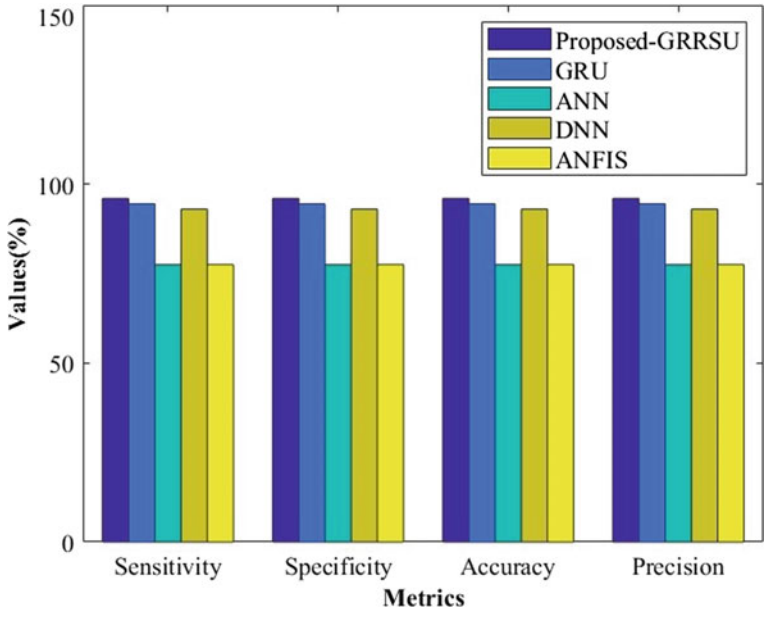


Fig. 4 Performance analysis of the proposed GRRSU

Table 2 Performance comparison of the proposed GRRSU and the existing methods

Techniques/metrics	Recall (%)	F-measure (%)	NPV (%)	FPR (%)
Proposed GRRSU	96	96	96	4
GRU	94.5	94.5	94.5	5.5
ANN	77.5	77.5	77.5	22.5
DNN	93	93	93	7
ANFIS	77.5	77.5	77.5	22.5

Based on recall, F-measure, NPV, and FPR, the proposed GRRSU method's performance is exhibited in Table 2. For efficient performance, the values of recall, F-measure, and NPV should be higher. FPR values contribute to false predictions. The proposed GRRSU method's recall, F-measure, and NPV exhibit the same value of 96%, whereas the values of existing methods provide comparatively lower performance. By achieving a lower FPR value, the proposed method has neglected the false prediction. Overall, the proposed GRRSU method proffered a better performance than prevailing techniques.

The proposed method's analysis with the existing methods is illustrated in Fig. 5. The FNR and FRR values of the prevailing GRU, ANN, DNN, and ANFIS are 5.5, 22.5, 7, and 22.5%, whereas the proposed method achieved a similar FNR and FRR value of 4%, which is lower than the existing methods. By obtaining an MCC value of 92%, the proposed method achieved a better performance, which is higher when analogized to the existing methods. Therefore, it is stated that the proposed method achieves greater performance regarding all metrics.

Figure 6 exhibits the ROC curve of the proposed GRRSU and the existing model. ROC is an effective method of evaluating the quality or performance of a classifier. By plotting the true positive rate (TPR) against FPR, an ROC curve is constructed. It is clear as of the ROC curve that the proposed model exhibited a superior performance than the prevailing models, such as DNN, ANN, ANFIS, and GRU.

3.4 Comparison Analysis of Classifier with and without LIDA

Here, the results attained by the proposed GRRSU model with dimensionality reduction and the outcomes without dimensionality reductions are analyzed. The analysis is grounded on the metrics, such as sensitivity, specificity, accuracy, and precision.

The comparative analysis of the proposed GRRSU classifier with LIDA and the classifier without LIDA is illustrated in Fig. 7. The accuracy attained by the classifier along with the dimensionality reduction technique is 96%, whereas the accuracy of the classifier is 94% when the features are given as input without performing dimensionality reduction. Similarly, a greater result in all metrics was attained by the proposed GRRSU with LIDA. This clearly exhibits that LIDA has enhanced the BWE process.

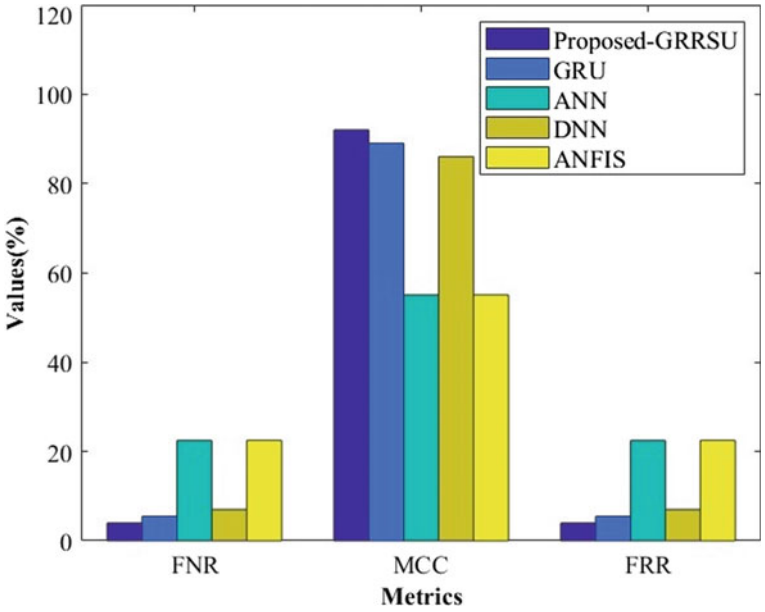


Fig. 5 Performance analysis of the proposed GRRSU

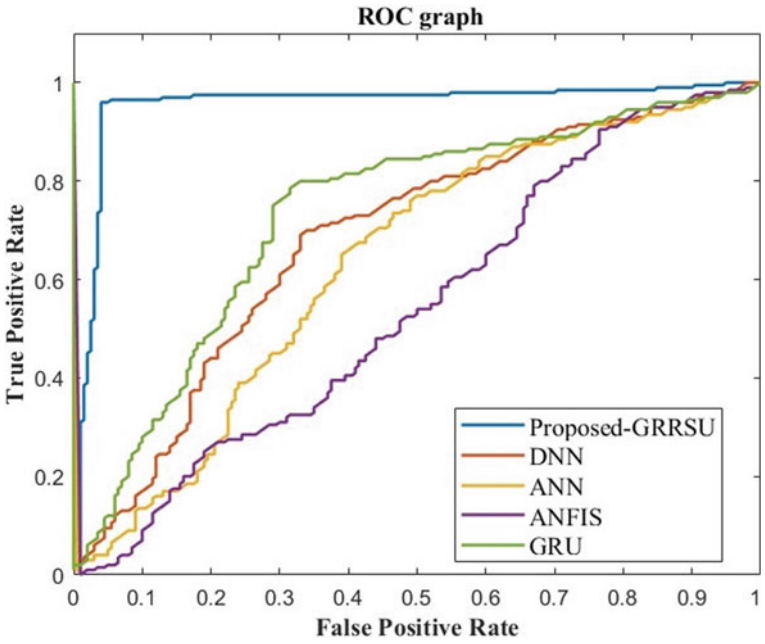


Fig. 6 ROC curve analysis

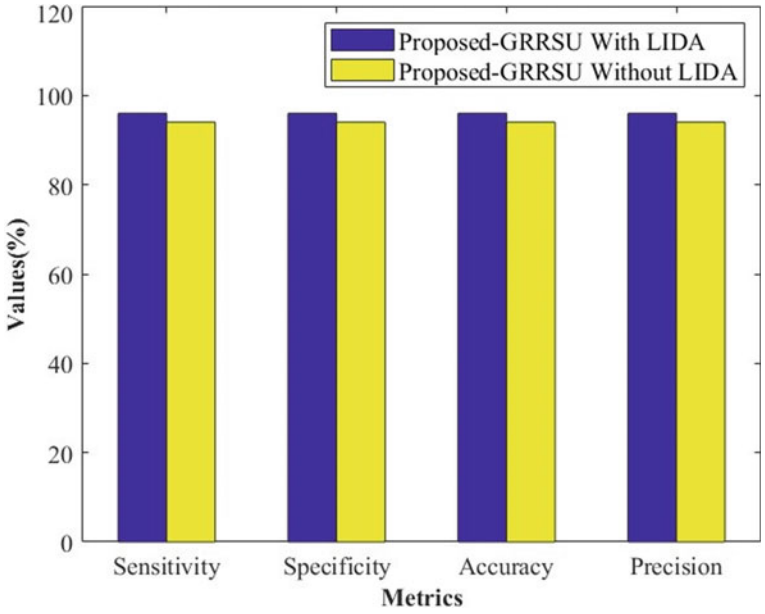


Fig. 7 Performance analysis based on the proposed LIDA

3.5 Comparative Measurement with Literature Papers

Here, the proposed BWE methodology's performance is analogized with traditional approaches like hybrid AEABW 1,918,802:27,946,426, MMEE 1,918,802:27,946,440, and BWE-AODV 1,918,802:27,946,430 grounded on the throughput value.

The proposed framework's performance measure based on its throughput was unveiled in Table 3. Throughput is a gauge of units of information that a node can process in a given amount of time. The proposed method's throughput value exhibits an increased value of 1.62 Mbps compared with the existing MMEE model, which has a throughput of 0.9 Mbps. Similarly, the prevailing AEABW and BWE-AODV models also exhibited lower throughputs of 0.175 and 0.3 Mbps, respectively. Therefore, it is clear that the proposed methodology's overall performance is better than the prevailing techniques.

Table 3 Performance validation of the proposed methodology

Techniques	Throughput (Mbps)
Proposed model	1.62
AEABW	0.175
MMEE	0.9
BWE-AODV	0.3

4 Conclusion

A novel DL-centered available BWE approach utilizing FPECM-MFL and GRRSU for adaptive video streaming services in MANET has been proposed in this work. By using several steps like node initialization, clustering, route discovery, optimal route selection, test packet transmission, feature extraction, dimensionality reduction, and BWE, the BWE is performed in this work. After that, the performance analysis as well as the comparative analysis are executed to evaluate the proposed system's performance. From the experimental analysis, an accuracy of 96% and a throughput of 1.62 Mbps were attained by the proposed framework. Regarding all metrics, the proposed model achieved better performance when analogized with the prevailing system. Therefore, it is concluded that the proposed system is superior to the prevailing techniques. Nevertheless, the model focused only on available BWE but failed to ponder the security level. The work will be extended in the future by detecting the intrusion nodes and secure packet transmission along with available BWE will be executed.

References

1. Saravanakumar P, Sundararajan TVP, Poornimasre J (2021) Bwiee protocol for route discovery and energy efficient in manet. Turk J Comput Math Educ 12(11):3657–3664
2. Chen Z, Zhou W, Wu S, Cheng L (2020) An adaptive on-demand multipath routing protocol with QoS support for high-speed MANET. IEEE Access 8:44760–44773
3. Gunasekaran K, Suresh G, Kavitha V, Lekashri S (2020) Resourceful and reliable routing selection scheme in MANET. In: Proceedings of the 7th international conference on smart structures and systems 2020
4. Miyim A, Ismail M (2021) Dynamic routing protocol for QoS enhancement in wireless mobile Ad-Hoc networks. Jordan J Electr Eng 7(1):29. https://doi.org/10.5455/jjee.204-1600789356
5. Nurcahyani I, Huda MTN (2020) Implementation analysis of available bandwidth estimation for multimedia service on VANET network using A-STAR routing protocol. In: 2020 3rd International Conference on Information and Communications Technology, ICOIACT 2020, May, 446–450. https://doi.org/10.1109/ICOIACT50329.2020.9332050
6. Haveliwala SK, Vyas D, Shah H (2020) Congestion control in MANET by analyzing network status-New ETCP-Westwood
7. Aina F, Yousef S, Osanaiye O (2021) BE-OZR: an efficient available bandwidth estimation in IEEE802.11e network. Bus Comput Sci 14:1–15
8. Li Z, Lei L, Zhang L, Sheng G, Cai S (2019) Available bandwidth estimation for directional CSMA/CA Ad Hoc networks. In: Proceedings of the 2019 15th international conference on mobile Ad-Hoc and sensor networks, MSN 2019, pp 107–112

Online Meeting Summarization Based on Text and Image Processing

Ishita Patel, Vishab Julka, Prutha Kudtarkar, Ishita Sharma, and Bhakti Sonawane

Abstract The rise of online meetings has led to an increase in the amount of audio and visual data that needs to be processed in order to extract relevant information. To address this issue, researchers have proposed a novel approach to online meeting summarization based on text and image processing. This approach involves using natural language processing techniques to extract important information from the audio transcript of the meeting and image processing techniques to extract key visual information from the video stream. The proposed system takes into account various factors, such as speaker turns, sentiment analysis, and key visual elements, and uses these factors to generate a comprehensive summary of the meeting. The system has been evaluated on a dataset of online meetings, and the results show that it is able to generate accurate and concise summaries. This approach has significant potential for improving the efficiency of online meetings by enabling participants to quickly review the most important points discussed during the meeting. Furthermore, it could also be used for automated transcription and captioning, which could help make online meetings more accessible to individuals with hearing impairments. Overall, online meeting summarization based on text and image processing has the potential to greatly enhance the effectiveness and accessibility of online communication. A method to extract keywords from educational video transcripts from Massive Open Online Courses (MOOC) is discussed. The approach is primarily based totally on Regular Expression Grammar Rule technique to identify the Noun Chunks in

I. Patel (✉) · V. Julka · P. Kudtarkar · I. Sharma · B. Sonawane
Shah and Anchor Kutchhi Engineering College, Mumbai, India
e-mail: ishitapatel628@gmail.com

V. Julka
e-mail: vishabjulka17@gmail.com

P. Kudtarkar
e-mail: pruthakudtarkar9168@gmail.com

I. Sharma
e-mail: ishita11094@gmail.com

B. Sonawane
e-mail: bhakti.sonawane@sakec.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_19

the textual content of the transcript. Extracting keywords helps in finding out the specifically important part of the educational material.

Keywords Text · Text summarization semantic analysis · Face detection

1 Introduction

The software will permit the user to automatically summarize relevant information from numerous sources and do sentiment analysis for determining if the meeting was on a positive note or a negative note. Text summarization in biomedical domain has largely been confined to extractive approaches. Abstractive approaches, using deep learning, have recently been a success for summarizing the general-domain documents, such as news articles, however, have not been applied to domain-specific documents due to the problem for neural models to study domain-specific knowledge.

Text summarization is the problem of reducing the number of sentences and words of a document without changing its meaning [1]. Text summarization is a very useful and important part of natural language processing (NLP) [2]. Since there are too many lines of text data in any form, and we have scarcity of time, so we only want the summary of the text. Instead of reading an entire article, we can have the most important statements. Text summarization takes an input of a sequence of words and returns an output of words that describes the summary [3]. In this work, a deep-reinforced, abstractive summarization method is able to study abstracts and generate summaries in the form of a one-sentence headline, or a title. Keyword extraction is the most important task while working with the textual content data [4]. Extracting keywords benefits the reader as to decide the important part of text as an alternative of going through the entire text. A method to extract keywords from educational video transcripts from Massive Open Online Courses is discussed. The approach is primarily based totally on Regular Expression Grammar Rule technique to identify the Noun Chunks in the textual content of the transcript. Extracting keywords helps in finding out the specifically important part of the educational material.

Sentiment analysis (SA) is an intellectual method of extricating the user's emotions and feelings [5]. It is one of the pursued subjects of natural language processing (NLP). The evolution of Internet-based applications has suggested a large quantity of personalized reviews for numerous related data on the web. These reviews exist in distinct forms like social medias, blogs, Wiki, or forum websites. Both tourists and customers locate the records in those reviews to be useful for his or her knowledge and planning processes [6]. The increase of engines like Yahoo and Google has flooded users with copious quantities of relevant reviews about precise destinations, which remains beyond human comprehension [7]. Sentiment analysis poses as an effective device for users to extract the needful facts, in addition to the mixture of the collective sentiments of the reviews. Sentiment analysis with the help of image processing will capture the facial expressions of the participants and create an analysis of whether the meeting was on a positive note or a negative note. It will

help to understand the feelings of the employee and create an analysis whether the employees are happy or not [8, 9].

1.1 Taxonomy for Text Summarization

Text summarization is the process of generating a coherent and brief summary of a large text document while retaining the most important information [10–12].

Text summarization is further divided into three sections as shown in Fig. 1:

1. Based on the input type
   - Single document summarization: This approach involves summarizing a single document or text. The input can be in various formats such as plain text, PDF, HTML. The goal is to extract the most important information from the document and create a shorter version of it.
   - Multi-document summarization: This approach involves summarizing multiple documents on the same topic. The input can be in various formats such as plain text, PDF, HTML. The goal is to extract the most important information from all the documents and create a shorter version of it.
2. Based on the output type
   - Extractive summarization: In extractive summarization, the summary is generated by selecting and extracting the most important sentences or phrases from the original text. The extracted sentences are usually arranged in a logical order to form a coherent summary. Extractive summarization techniques rely on natural language processing and machine learning algorithms to identify the most salient content in the source document. The output of extractive

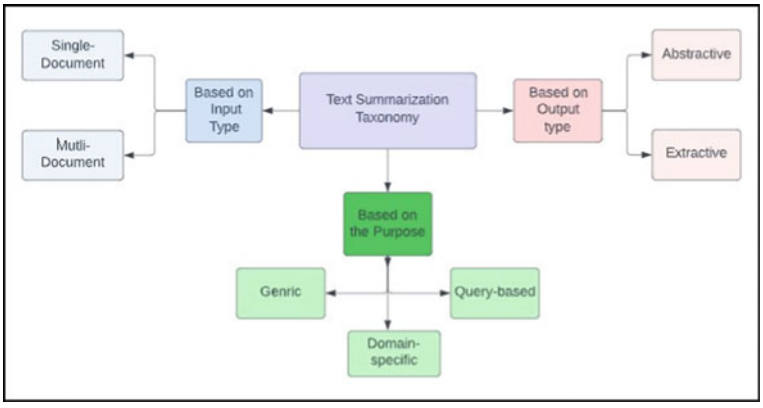


Fig. 1 Taxonomy for text summarization

summarization is usually a subset of the original text, with some modifications to improve coherence and readability.

- Abstractive summarization: Abstractive summarization, on the other hand, generates a summary that is not necessarily a subset of the original text, but rather a new text that captures the most important ideas and concepts from the original document. In abstractive summarization, the algorithm creates a summary by generating new sentences that convey the main ideas from the source document. This requires more advanced natural language processing techniques and machine learning models that can generate coherent and readable text. The output of abstractive summarization is usually shorter than the original text and more concise, while still preserving the main ideas and themes.

3. Based on the purpose

- Generic summarization: Generic summarization refers to the process of summarizing a text document without any specific purpose or context. It is a general-purpose summarization technique that can be applied to a wide range of text documents, such as news articles, research papers, or books. The goal of generic summarization is to provide a brief overview of the most important information in the source document, while retaining its key ideas and themes.
- Domain-specific summarization: Domain-specific summarization is a technique used to summarize text documents that are related to a specific domain or topic. This technique involves using domain-specific knowledge to extract and summarize the most important information from the source document.
- Query-based summarization: Query-based summarization is a technique used to generate a summary of a text document based on a specific query or question. This technique involves analyzing the query or question and identifying the most relevant information in the source document that answers the query. Query-based summarization is commonly used in search engines, where the search query is used to generate a summary of the most relevant search results.

In this system, the extractive summarization technique based on the output type is used to summarize the transcript generated by users in online meetings.

1.2 Taxonomy for Sentiment Analysis

Sentiment analysis is a technique that automatically identifies and extracts the sentiment or emotion expressed in a piece of text as shown in below Fig. 2 [13, 14].

Sentiment analysis can be divided further as shown in Fig. 2.

1. Opinion mining: Opinion mining can be used to analyze customer feedback, social media posts, product reviews, and other forms of user-generated content. Opinion mining can provide valuable insights into customer feedback, social

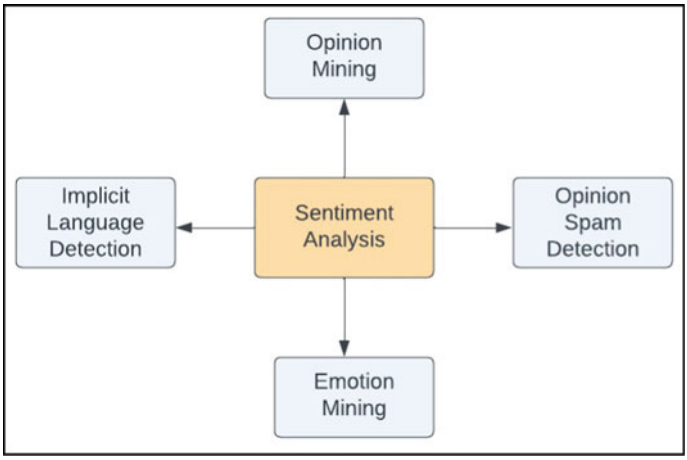


Fig. 2 Taxonomy for sentiment analysis

media trends, and other forms of user-generated content. It can be used to identify areas where improvements are needed, as well as to track the effectiveness of marketing campaigns and other initiatives. However, it is important to note that opinion mining is not always accurate and can be influenced by factors such as sarcasm, irony, and cultural differences.

Opinion mining involves several steps, including:

- Preprocessing: The text is preprocessed to remove stop words, punctuation, and other irrelevant information. The text is also tokenized into individual words.
- Part-of-speech tagging: Each word in the text is tagged with its part of speech, such as noun, verb, or adjective.
- Sentiment classification: The sentiment of each word is determined based on its part of speech and the sentiment it conveys. For example, words like "good" and "excellent" are generally considered positive, while words like "bad" and "poor" are considered negative.
- Opinion aggregation: The sentiment of the entire text is determined by aggregating the sentiment of each word in the text. This can be done using different methods, such as counting the number of positive and negative words, or using machine learning algorithms to classify the sentiment of the text.

2. Opinion spam detection: Opinion spam detection involves analyzing the sentiment and content of reviews to identify patterns that are indicative of spam.

Some of the common features that are used to detect opinion spam include:

- Similarity: Opinion spam often contains similar or identical language across multiple reviews, indicating that the reviews are not authentic.
- Sentiment polarity: Opinion spam often contains extreme sentiment, either positive or negative, and may not be reflective of a genuine customer experience.

- Language complexity: Opinion spam may use overly complex or technical language that is not typical of customer reviews.
- Timing: Opinion spam may be posted in a short period of time, indicating that the reviews are not reflective of genuine customer experiences.
- Profile information: Opinion spam may be posted by users with incomplete or suspicious profile information.

3. Emotion mining: Emotion mining involves analyzing the language used in a piece of text to determine the specific emotions that are conveyed. This can be done using a range of techniques, including natural language processing, machine learning, and deep learning [15, 16]. Some of the common emotions that are detected in emotion mining include happiness, sadness, anger, fear, and surprise. Emotion mining can be useful in a variety of applications, including customer feedback analysis, social media monitoring, and market research. By understanding the specific emotions expressed in customer feedback or social media posts, organizations can gain valuable insights into how their products or services are perceived by their target audience [17, 18]. This information can then be used to improve products, tailor marketing campaigns, and enhance customer experiences.
4. Implicit language detection: Detecting implicit language can be challenging, as it requires analyzing not only the words used in a piece of text but also the context and tone in which they are used. Some of the common techniques used to detect implicit language include natural language processing, machine learning, and deep learning. One of the key challenges in implicit language detection is identifying sarcasm and irony. These forms of implicit language are often used to express a sentiment that is opposite to what is stated explicitly. For example, a tweet that says "I just love sitting in traffic for hours on end" may actually express a negative sentiment toward traffic, despite the use of the word "love". Implicit language detection can be useful in a variety of applications, including social media monitoring, customer feedback analysis, and market research. By understanding the sentiment expressed in implicit language, organizations can gain valuable insights into how their products or services are perceived by their target audience. This information can then be used to improve products, tailor marketing campaigns, and enhance customer experiences. However, it is important to note that implicit language detection is not always accurate, and the results may be influenced by factors such as cultural differences, context, and tone. Therefore, it is important to have a human review process in place to ensure the accuracy of the sentiment analysis results.

In this system, the emotion mining part of sentiment analysis is used for face detection module of online meetings. It captures the faces of the users in the meeting and determines the emotions such as neutral, happy, sad, angry, fearful, disgusted, and surprised.

2 Proposed System

The proposed system consists of online meetings application in which the speech of users is detected and converted into a transcript of text. This transcript is further summarized by using text summarization of natural language processing. Simultaneously, the system captures the faces of the users in the meeting. Sentiment analysis of emotions is then generated by this face detection.

2.1 Algorithm and Working

For the server part, Agora.io is used. So, the user just needs to join the meeting and all its data will be given on to the Agora server for further processing [19].

The above block diagram shown in Fig. 3 gives the clear view of the working of the Agora server. First, a token is created before the users join the meeting. Then, for joining the meet, the user needs to join a channel. Before joining the channel, an alert message is displayed on the user's screen asking to give access to the user's camera and microphone options. Then this access information is sent to the Agora server, and the server respectively. Hence, users communication is possible by sending and receiving video and audio in the channel (Fig. 4).

The user has to first either create a meeting or join a meeting. All the participants start their meeting and will simultaneously view their transcript. After the meeting is over, the participants can get a summarized text of what was discussed in the meeting. This is the minutes of the meeting which is automatically generated after the completion of the meeting. The system also gives sentiment analysis which helps to analyze whether the meeting was on a positive note or on a negative note.

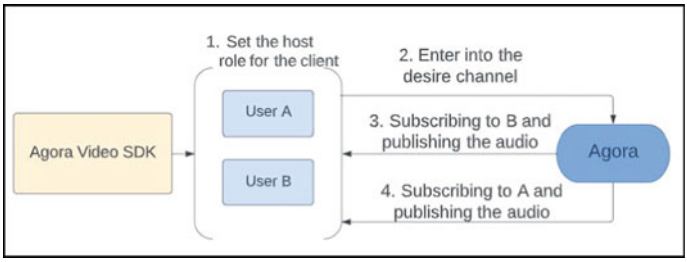


Fig. 3 Block diagram describing the flow of simulation

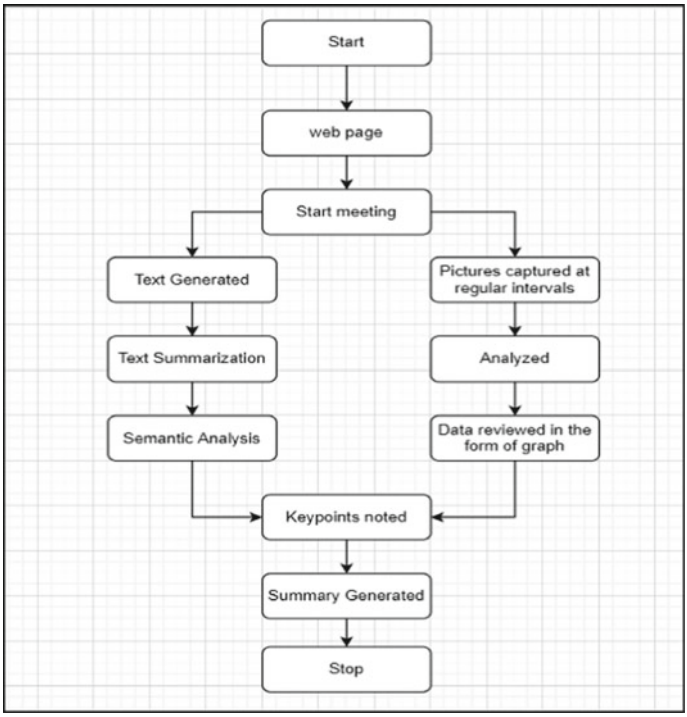


Fig. 4 Block diagram

2.2 Implementation

The implementation of Agora video calling involves a series of steps that are necessary to set up the video call and establish communication between the participants. Hence, it involves the following steps for the call sequence:

1. Initialize the Agora SDK: Start by executing the 'initialise' function to initialize the Agora SDK. This procedure initializes the SDK using the App ID and configures the required video calling resources.
2. Build an Agora RTC Engine: After initializing the SDK, the execution of the 'create' method leads to the construction of the Agora RTC Engine. A visual and audio communication instance of the RTC Engine is created using this method.
3. Set up the local video and audio: The third step is to set up the local video and audio by calling the 'setupLocalVideo' and 'setupLocalAudio' methods. The local user can record and broadcast video and audio streams using these techniques.
4. Join a channel: After setting up the local audio and video, you must join a channel by invoking the "joinChannel" method. Using this technique, communication is initiated and the local user is connected to a certain channel.

5. Set up the remote video: After joining the channel, you must use the 'setupRemoteVideo' method to set up the remote video. This technique renders the incoming video stream and configures the video display for the remote user.
6. Manage user events: During the video conference, you must manage user events such users joining or leaving the channel, network outages, and problems with the audio or video quality. To handle these events, you can make use of the SDK's event listeners.
7. End the video call and exit the channel: After the video call has ended, exit the channel by executing the 'leaveChannel' method. By using this technique, the local user is removed from the channel, freeing up resources for the RTC Engine.

The call sequence for implementing Agora.io Video Calling involves initializing the SDK, creating an RTC engine, setting up local and remote video and audio, joining a channel, handling user events, and leaving the channel.

2.3 Layout

- Link for Joining the Meeting

After linking on the link provided by the host of the meeting, the users will be redirected to this page where they will find a button to join the meeting. The participants have to click on the button to join the meeting.

- Meeting Interface for Four Participants

After the users have clicked on the button, they will be able to join the meeting and communicate with each other through the video chatting application. The users will be able to see the transcript of what each of the participants is speaking. Each user will also be given a taskbar where he will be given buttons to leave the stream, switch on or off the mic, and switch on or off the camera.

- Participants Switching off their Camera

After a participant switches off their camera, the layout of the meeting changes, and everyone in the meeting will be prompted with a new layout.

- Two Participants Switching off their Camera

Similarly, when two participants switch off their camera, the interface changes, and every participant is prompted with a changed interface.

- Switching off the Mic

The users also have the ability to switch off their mic. They can simply click on the turn-off mic button as seen in Fig. 5(d). They will still be able to hear and see the transcript of what the other participants are speaking.

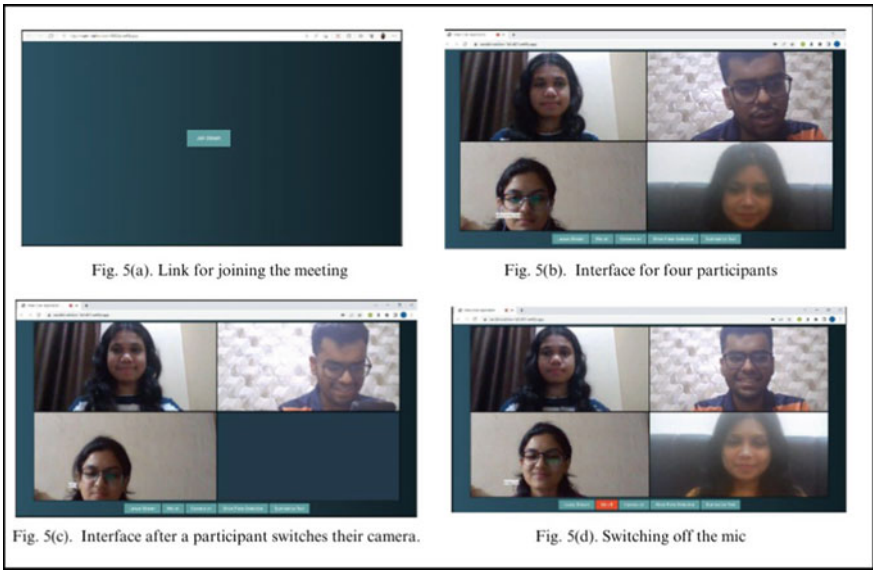


Fig. 5(a). Link for joining the meeting

Fig. 5(b). Interface for four participants

Fig. 5(c). Interface after a participant switches their camera.

Fig. 5(d). Switching off the mic

Fig. 5 Interface of the system

- Mobile View of the System

This system can also be accessed through a mobile phone. When a user joins the meeting through the mobile phone, a slightly different layout is prompted to them (Fig. 6).

- Dialog Box of Face Detection

The system has a dialog box that shows the face recognition points on the computer screen of individuals. Using machine learning and image processing algorithms, it can capture the facial features of the person and determines their mood. It captures features such as neutral, happy, sad, angry, fearful, disgusted, and surprised. As shown in Fig. 7a, the dialog box captures the feature 'happy' with an accuracy of 0.99. Figure 7b shows the dialog box that captures the feature 'happy', 'neutral', and 'sad' with an accuracy of 0.54, 0.31, and 0.13, respectively. So, the system will take into account the feature that depicts the highest accuracy. So here, it will consider 'happy' as the facial feature. Figure 7c shows the dialog box that captures the feature 'neutral' and 'sad' with an accuracy of 0.64 and 0.31, respectively. So, the system will capture 'neutral' as the feature.

- Text Summarization of audio

Upon clicking on the 'Summarize Text' button, the transcript generated by speech-to-text conversion of participants' audio input is summarized and displayed on a separate window.

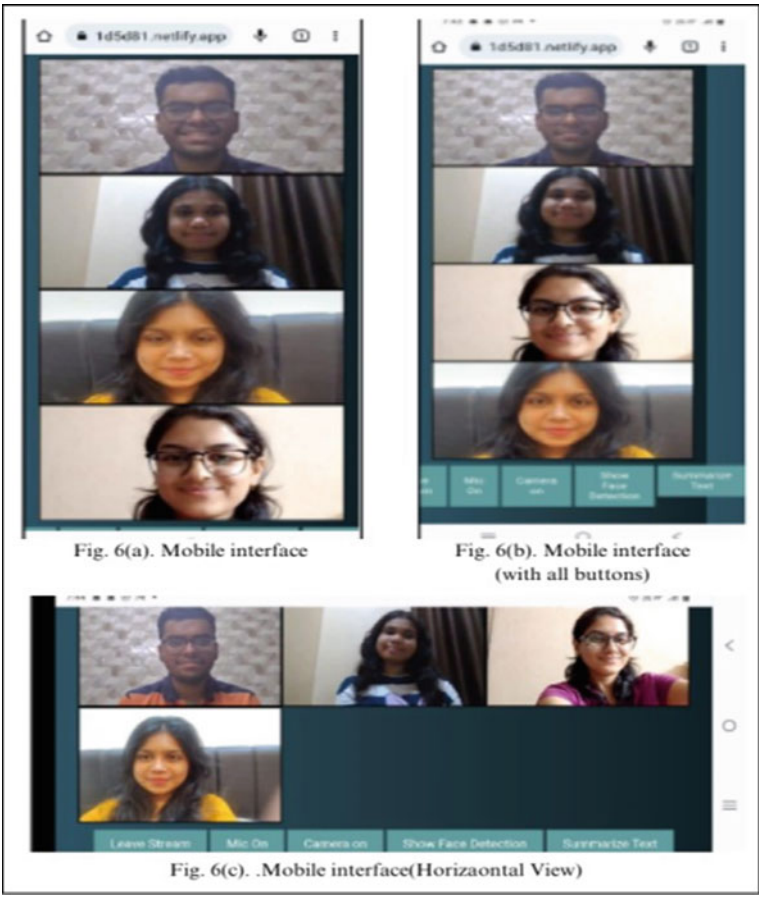


Fig. 6(a). Mobile interface

Fig. 6(b). Mobile interface (with all buttons)

Fig. 6(c). .Mobile interface(Horizaontal View)

Fig. 6 Mobile interface

3 Conclusion and Future Scope

3.1 Conclusion

Text summarization is the task of extracting salient information from original text document. In this process, the information gained is generated as condensed report and presented as a concise summary to the user. It is quite arduous for humans to understand and interpret the content of text. Hence, to get the quick gist of the meeting, an application is created to convert the text to speech and to summarize the textual content. Sentiment analysis is done on the images captured using image processing. This will help us in determining the mood of the meeting. It will also help in employee development as it understands whether the employee is motivated or

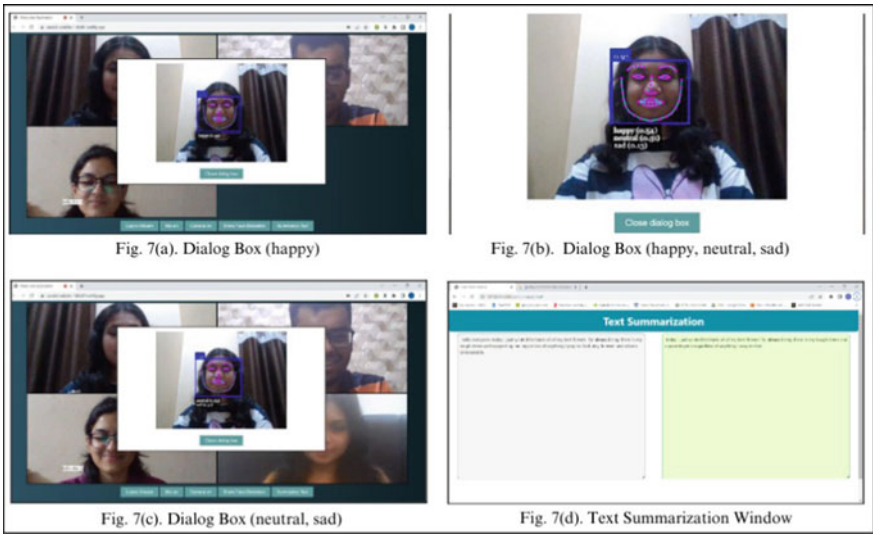


Fig. 7(a). Dialog Box (happy)

Fig. 7(b). Dialog Box (happy, neutral, sad)

Fig. 7(c). Dialog Box (neutral, sad)

Fig. 7(d). Text Summarization Window

Fig. 7 Dialog box with face detection and text summarization

depressed through sentiment analysis. This will also help in achieving quality work and meeting deadlines.

3.2 Future Work

In the future, the summarized text which we have gained as an output will be further processed or mapped with a dictionary dataset to produce a valuable outcome. Further, we will keep varying this dataset as per the new use cases and the applications. So that, with every new dataset, the model will give that domain-specific analysis and application areas will be expanded. This increases the overall scope of the project. Also, the output of the face detection model generated will be analyzed for the whole sentiment analysis of the meeting. And, finally, we will generate the graphs for our respective sentiment analysis of both text and image.

References

1. Awasthi KG, Bhogal PS, Anand SS, Soni PK (2021) Natural language processing (NLP) based text summarization: a survey. In: Proceedings of the 2021 6th international conference on inventive computation technologies (ICICT), pp 1310–1317. https://doi.org/10.1109/ICICT50816.2021.9358703
2. Boorugu R, Ramesh G (2020) A survey on NLP based text summarization for summarizing product reviews. In: Proceedings of the 2020 second international conference on inventive

research in computing applications (ICIRCA), pp 352–356. https://doi.org/10.1109/ICIRCA48905.2020.9183355
3. Jain A, Bhatia D, Thakur MK (2017) Extractive text summarization using word vector embedding, pp 51–55. https://doi.org/10.1109/MLDS.2017.12
4. Shukla H, Kakkar M (2016) Keyword extraction from educational video transcripts using NLP techniques. In: Proceedings of the 2016 6th international conference: cloud system and big data engineering (Confluence), pp 105–108. https://doi.org/10.1109/CONFLUENCE.2016.7508096
5. Devika MD, Sunitha C, Ganesh A (2016) Sentiment analysis: a comparative study on different approaches. Proced Comput Sci 87:44–49
6. Tandel A, Modi B, Gupta P, Wagle S, Khedkar S (2016) Multi-document text summarization: a survey. In: Proceedings of the 2016 international conference on data mining and advanced computing (SAPIENCE), pp 331–334. https://doi.org/10.1109/SAPIENCE.2016.7684115
7. Gigioli P, Sagar N, Rao A, Voyles J (2018) Domain-aware abstractive text summarization for medical documents. In: Proceedings of the 2018 IEEE international conference on bioinformatics and biomedicine (BIBM), pp 2338–2343. https://doi.org/10.1109/BIBM.2018.8621539
8. Nasukawa T, Yi J (2003) Sentiment analysis: capturing favorability using natural language processing, pp 70–77. https://doi.org/10.1145/945645.945658
9. Rajput A (2019) Natural language processing, sentiment analysis and clinical analytics
10. Andhale N, Bewoor LA (2016) An overview of text summarization techniques. In: Proceedings of the 2016 international conference on computing communication control and automation (ICCUBEA), pp 1–7. https://doi.org/10.1109/ICCUBEA.2016.7860024
11. Janjanam P, Reddy CP (2019) Text summarization: an essential study. In: Proceedings of the 2019 international conference on computational intelligence in data science (ICCIDS), pp 1–6. https://doi.org/10.1109/ICCIDS.2019.8862030
12. Biswas S, Rautray R, Dash R, Dash R (2018) Text summarization: a review. In: Proceedings of the 2018 2nd international conference on data science and business analytics (ICDSBA), pp 231–235. https://doi.org/10.1109/ICDSBA.2018.00048
13. Huang Y, Li G (2010) A semantic analysis for internet of things. In: Proceedings of the 2010 international conference on intelligent computation technology and automation, pp 336–339. https://doi.org/10.1109/ICICTA.2010.73
14. Choudhary S, Godara J (2021) Semantic analysis on social media. In: Proceedings of the 2021 international conference on computing sciences (ICCS), pp 239–243. https://doi.org/10.1109/ICCS54944.2021.00054
15. John A, Abhishek MC, Ajayan AS, Sanoop S, Kumar VR (2020) Real-time facial emotion recognition system with improved preprocessing and feature extraction. In: Proceedings of the 2020 third international conference on smart systems and inventive technology (ICSSIT), Tirunelveli, India, pp 1328–1333. https://doi.org/10.1109/ICSSIT48917.2020.9214207
16. Chandraprabha KS, Shwetha AN, Kavitha M, Sumathi R (2021) Real time-employee emotion detection system (RtEED) using machine learning. In: Proceedings of the 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV), Tirunelveli, India, pp 759–763. https://doi.org/10.1109/ICICV50876.2021.9388510
17. IJCRT (2022) Real time facial recognition system using facial expressions
18. Elliott ML (2022) Sentiment analyzer, tool which analyzes facial emotion in real time. Int J Res Publ Rev 3(4):2072–2077
19. Agora real-time voice and video engagement. https://www.agora.io

A Byte Transformation Technique for Lossy Image Compression

Tanuja R. Patil and Vishwanath P. Baligar

Abstract With the advancement of IT field and social media communications, the transaction of images is increasing day by day. Hence, it is required to do the image compression for storage as well as transmission. In this paper, we discuss a transformation-based technique for grayscale image compression. In this approach, a block of four pixel data is transformed in to three bytes of data. Later, additional bits are added by suitable encoding method to achieve high-quality reconstructed images. Though it is lossy, the accuracy achieved at same PSNR levels compared to JPEG is higher. Compression ratio achieved is also comparable.

Keywords Lossy image compression · Transformation technique for image compression · Quality of reconstructed images · Comparison with JPEG

1 Introduction

In recent years, because of the advancement in social media, daily transactions involve large number of digital images, requiring large amount of memory for storage and also transmission. Most of the image capture hardware, support 8-bit images, which are stored as grayscale images. There is a high redundancy present in the storage of grayscale images, and by reducing this redundancy, we can achieve compression. The image compression may be lossless or lossy type. Lossless image compression is required in medical images, satellite imagery, etc. where perfect reconstruction is required. Earlier lossless compression techniques are run length coding, arithmetic coding, LZW coding, Huffman coding, etc. Recent techniques are being developed using predictive-based coding. Some of the lossy compression

T. R. Patil (✉) · V. P. Baligar
K.L.E. Technological University, Hubballi, India
e-mail: tanuja_p@kletech.ac.in

V. P. Baligar
e-mail: vpbaligar@kletech.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_20

techniques are transformation-based codes like discrete cosine transform (DCT, used in JPEG) or wavelet transform (used in JPEG 2000) and many others.

In this paper, a unique method is discussed, which uses an octal encoding technique, which will result in slightly lossy compression. Some additional bit data is added to improve the quality, so that near lossless results can be achieved. Organization of the paper is as follows. Section 1 is introduction, Sect. 2 describes literature survey, Sect. 3 discusses about the methodology, and Sect. 4 deals with results.

2 Review of Literature

Singh and Kirar [1] discuss about various image compression techniques like k-means clustering, 2D-DWT, and fuzzy logic-based image compression.

Balle et al. [2] discuss about a method which uses uniform quantizer, and a nonlinear synthesis transformation. The transforms use three stages of convolutional linear filters and nonlinear activation functions. They have chosen a joint nonlinearity to implement a form of local gain control by which the visual effects are improved.

Rastogi and Sengar [3] discuss about the Performance Comparison of Lossless Compression Techniques for Text Data. They say that run length coding is best algorithm to compress text data.

Patidar et al. [4] provide a review on various image compression techniques, which are used in medical image compression. Performance analysis and existing research are compared on medical image compression.

Goel and Garg [5] discuss about a method of spatial space reduction with an idea of data folding. It uses shading images with various sizes. A row-folding and column-folding method which is used iteratively is discussed here. This method gives an improved compression ratio.

Mentzer et al. [6] discuss that deep neural networks trained as image auto-encoders are the recent advancements in image compression. They have used the technique of rate distortion trade-off for an image compression auto-encoder. They have used context model to directly model the entropy.

Ilic et al. [7] say that using JPEG approach, at lower values of PSNR, there appear contours by which the image appears blurred. These structures are not comfortable for the human vision, and hence, improvement is required.

By this, we came to know that there is a necessity to improve the visibility of lossy images at low PSNR values. Hence, we have come up with a new technique named a byte transformation approach by which the quality of images can be increased.

3 Proposed Method

Byte transformation approach is a novel method of image compression which results in high-quality reconstructed images. Here, we discuss about the compression and decompression techniques. The grayscale image is processed using windowing method by considering a block of four pixels at a time, and these four pixels are transformed into three bytes by a unique transformation approach. Later, some additional bits are added. In this way, the whole image data are converted to produce a transformed image. Later, image reconstruction is done by suitable decompression algorithm. These algorithms are discussed in detail.

3.1 Compression Algorithm

1. Read the gray scale image and divide into blocks of four pixels.
2. Write the decimal value of the pixel intensity in 8 bit binary form.
3. Divide 8 bits of all four pixels into three groups.
4. First group G1 is formed by considering bit7 and bit6 of all 4 pixels.
5. Second group G2 is formed by considering bits at bit5, bit4 and bit3 positions.
6. Third group G3 is formed by considering bits at bit2, bit1 and bit0 positions.
7. Create First byte B_1 by combining group1 bits.
8. Bits 5, 4, 3 indicate an octal value. Mark '1' in the bit position indicated by this octal value in the bit position table for all 4 pixels. Later combine this 8 bit data of all 4 pixels to generate second byte B2.
9. Bits 2, 1, 0 indicate an octal value. Mark '1' in the bit position indicated by this octal value in the bit position table for all 4 pixels. Later, combine this 8 bit data of all 4 pixels to generate third byte B3.
10. Add additional bits by reading the count of '1's present in the bytes B2 and B3.
11. Finally all the files are compressed by using Huffman compression for further compression.

Example: Consider a sample of Lena image as shown in Table 1.

1. Let us consider the first block consisting of four pixels as per Table 1, P1, Q1, R1, S1.
2. Write these four pixels into binary form, i.e., P1 = 162 = 10100010, Q1 = 164 = 10100100, R1 = 161 = 10100001, S1 = 163 = 10100011.

Table 1 Sample Lena image

162 (P1)	164 (Q1)	160 (P2)	158 (Q2)	160
161 (R1)	163 (S1)	158 (R2)	160 (S2)	161
163 (P3)	164 (Q3)	162	160	162
161 (R3)	162 (S3)	162	161	162

3. Now consider two most significant bits of these pixels. $P1_{[7,6]} = 10$, $Q1_{[7,6]} = 10$, $R1_{[7,6]} = 10$, $S1_{[7,6]} = 10$ then concatenate the 2 bits of all four pixels to form B1 = 10101010. This byte will be stored in a file 'B1'.
4. Second byte B2 is generated by logical OR of bit positions of group2 bits. The bit positions of these four pixels are determined as follows. For example, in pixel P1, group 2 bits = 100 = 4 (in octal). So '1' is marked at bit position 4 and '0' is entered in all other bit positions. Similarly, bit positions for other pixels are also determined and found to be 4 in all these cases.

Byte B2 is formed by logical OR of all the above-formed bytes as shown below.

B2 = (00010000) OR (00010000) OR (00010000) OR (00010000) = 00010000.

This is how second byte is generated and this can be stored in a file and can be named as 'B2'.

5. Third byte is generated as shown. The bit positions of group 3 bits, i.e., bits at 2, 1, 0 positions are checked and '1' is marked in these bit positions. In above example, bits at bit positions 2, 1, 0 can be written as follows.

$P1_{[2,1,0]} = 010$ i.e 2, $Q1_{[2,1,0]} = 100$ i.e. 4, $R1_{[2,1,0]} = 001$ i.e. 1, $S1_{[2,1,0]} = 011$ i.e. 3

Hence, at these positions, '1's are entered to generate byte3. Byte3 is generated by logical OR of these bit position indicating bytes. Byte3 = (000001000) OR (00010000) OR (00000010) OR (00001000) = 00011110.

This is how third byte is generated.

6. Additional Bit Generation

In byte2 and byte3, count the number of '1's, which will be either 1, 2, 3, or 4. If the number of '1's is one, all pixels indicate the same bit position value, and hence, reconstruction will be perfect, but if the count of '1's is two or three, additional bits are to be added to achieve better results. It will result in lossy compression only, but reconstructed images will be of higher quality. Hence, in this section, we have shown how to add the additional bits, if the count of '1's is 2 or 3 and how the quality has been improved.

6.1 How to Generate the Bits if the Count is 2?

In this section, generation of additional bits for two different blocks is explained where, the count of number of '1's is 2 in one block and 3 in another block. For the first block of Lena image sample shown in Table 1, count of number of '1's is 1, and hence, we will consider the next block indicated by P2, Q2, R2, S2. In this block, P2 = 160 = 10 100 000, Q2 = 158 = 10 011 000, R2 = 158 = 10 100 011, S2 = 160 = 10 011 101. In these four pixels, number of '1's is two in byte2. The bits at fifth, fourth, and third positions of these four pixels are $P2_{[5,4,3]} = 100$, $Q2_{[5,4,3]} = 011$, $R2_{[5,4,3]} = 100$, $S2_{[5,4,3]} = 011$. For this block, bytes B1, B2, B3 are generated as per the above steps. Now, bit generation for the count of '1's = 2 is explained.

The bit positions in four pixels of this block are either 3 or 4. Now to generate the additional bits, the bit positions shown in Table 2 are given the codes as follows. The code '0' is assigned to the lower bit position (which is 3) here. The code '1' is assigned for the higher bit position (which is 4). Now, the additional bit data is generated by checking the bit positions of P2, Q2, R2, S2 and the code is assigned as shown.

Bit Position of Pixels

For P2, bit position is 4, for Q2, it is 3, for R2, it is 4, and for S2, it is 3. Hence, a new 4-bit code is generated as 1 0 1 0 and stored in a separate file say 'Bit1' file.

6.2 How to Generate Bits if the Count = 3?

Consider another block of four pixels, in which the number of '1's is three. This block has P3 = 163, Q3 = 169, R3 = 176, S3 = 163. Their binary representation can be given as P3 = 10 100 011, Q3 = 10 101 001, R3 = 10 110 000, S3 = 10 100 011. Here, the bits at bit positions five, four, and three can be written as follows. $P3_{[5, 4, 3]} = 100$, $Q3_{[5, 4, 3]} = 101$, $R3_{[5, 4, 3]} = 110$, $S3_{[5, 4, 3]} = 100$, and '1' is marked in bit positions.

Next, we explain encoding of bit positions. Initially, bit positions are identified and written in ascending order. In this case, bit positions identified are 4, 5, 6. The assignment of the codes is as shown. Code '00' to lower bit position, i.e., for 4. Code '01' to mid bit position 5 and '10' code is given to higher bit position 6. Hence, Bit2 = 00 01 10 00. Thus, the bit-code is developed for the count of 3 and stored in a file say 'Bit2' file.

7. In the same way, for the bits at position two, one, zero, same procedure as mentioned in 6.1–6.2 are followed to generate the extra bits and saved in 'Bit3' and 'Bit4' files.
8. These above-generated files are compressed using Huffman compression to achieve further compression.

Table 2 Bit position table of fifth, fourth, third bits of four pixels

	Bit positions							
$Bits_{[5, 4, 3]}$	7	6	5	4	3	2	1	0
$P2_{[5, 4, 3]}$	0	0	0	1	0	0	0	0
$Q2_{[5, 4, 3]}$	0	0	0	0	1	0	0	0
$R2_{[5, 4, 3]}$	0	0	0	1	0	0	0	0
$S2_{[5, 4, 3]}$	0	0	0	0	1	0	0	0

3.2 Decompression Algorithm

Decompression can be achieved with the help of stored three bytes and additional bit information. This algorithm is explained below.

1. Read the files 'B1', 'B2', 'B3', 'Bit1', 'Bit2', 'Bit3', 'Bit4'.
2. Get the two most significant bits of four pixels from B1 file.
3. Read the 'B2' file, and get the bits located at fifth, fourth, and third positions of a byte as described below.
    i. B2 file contains the information regarding count of '1's, whether this count is two or three and also bit positions. Mark the lower and higher bit positions as L and H.
    ii. If the count value = two, refer the 'Bit1' file and get the code to reconstruct bits at 5, 4, 3 positions. For example, if code value is (01 01), it shows which bit position has to be assigned for P, Q, R, S pixels, i.e., in this example, lower bit position code (L) is assigned to pixel $P_{[5, 4, 3]}$ and $R_{[5, 4, 3]}$, and higher bit position code (H) is assigned to $Q_{[5, 4, 3]}$ and $S_{[5, 4, 3]}$. This is how reconstructions of bits 5, 4, 3 of all four pixels are done.
    iii. For the number of '1's = Three, read the 'Bit2' file to know the stored code. For example, the code value 00 10 01 01 infers that the fifth, fourth, and third bit values of pixel P have lower bit position value and pixel Q has higher bit position value, and R and S have mid bit position value. Thus, we can reconstruct bits 5, 4, 3 of all pixels.
4. In the next step, bits at positions two, one, zero also can be obtained by following the above steps by reading 'Bit3' and 'Bit4' file.
5. For a count value = four, again we can add the additional bits. If we add, we will get lossless compression, but compression ratio slightly reduces.

4 Discussion of Results

A set of standard IEEE images of size (512 × 512), as shown in Figs. 1, 2, 3, and 4, and some non-standard images of size (720 × 580) shown in Fig. 5 and (320 × 280) shown in Fig. 6 are used for testing, to show that this algorithm works for any size image. Python tool is used to implement the algorithms. Original images are shown at left side and reconstructed images are on the right side. Here, we can see that the quality of reconstructed images is very high. We have used our metric called as correctness ratio to find the number of pixels having same value as that of original image after reconstruction shown in Table 3. The compression ratio achieved is comparable. The correctness ratio is higher compared to JPEG, at same PSNR values [8].

Original and Reconstructed Images

Fig. 1 Lena image

Fig. 2 Barbara image

Fig. 3 Tank image

Fig. 4 Cameraman image

Fig. 5 Hill image

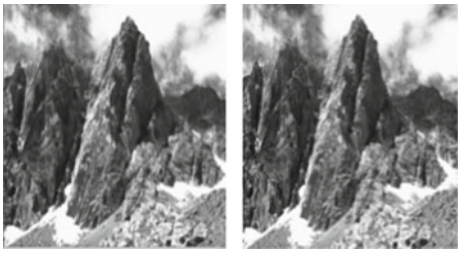


Fig. 6 Cat image

Table 3 Comparison of equal pixels with JPEG approach

Inputs	PSNR (db)	Total number of pixels	Equal pixels using JPEG technique	Correctness ratio by JPEG approach	Equal pixels using our technique	Correctness ratio by our approach	Compression ratio C.R
Lena	41.1	262,144	50,484	0.19	179,631	0.68	1.45
Barbara	38.5	262,144	40,486	0.15	166,163	0.63	1.38
Tank	41.1	262,144	55,565	0.21	194,505	0.74	1.43
Cameraman	42.1	262,144	64,761	0.24	208,071	0.79	1.43
Hill	38.1	3,07,200	73,872	0.24	108,789	0.35	1.42
Cat	39.6	50,440	9080	0.18	29,223	0.57	1.45

5 Conclusion and Future Scope

Byte transformation technique is a unique and a very simple technique for lossy image compression. Reconstructed images will be of very high quality. The accuracy obtained by this method is higher, i.e., the number of pixels with same value as compared to original using this method are more than JPEG approach measured at same PSNR values. Also, lossless image compression is possible by adding additional bits. Computation complexity is also lesser. It uses very simple approach for compression as well as decompression. Further, there is a scope to achieve lossless compression and also to increase compression ratio.

References

1. Singh A, Kirar KG (2017) Review of image compression techniques. In: Proceedings of the 2017 international conference on recent innovations in signal processing and embedded systems (RISE), pp 172–174. https://doi.org/10.1109/RISE.2017.8378148
2. Balle J, Laparra V, Simoncelli EP (2017) End-to-end optimized image compression. In: Proceedings of the 5th international conference on learning representations (ICLR)
3. Rastogi K, Sengar K (2014) Analysis and performance comparison of lossless compression techniques for text data. Int J Eng Technol Comput Res (IJETCR) 2(1):16–19

4. Patidar G, Kumar S, Kumar D (2020) A review on medical image datacompression techniques. In: Proceedings of the 2nd international conference on data, engineering and applications (IDEA), pp 1–6. https://doi.org/10.1109/IDEA49133.2020.9170679
5. Goel R, Garg S (2015) Lossless image compression using data folding followed by arithmetic coding. IOSR J Comput Eng 17(2):67–70
6. Mentzer F, Agustsson E, Tschannen M, Timofte R, Van Goo L (2018) Conditional probability models for deep image compression. https://openaccess.thecvf.com/content_cvpr_2018/papers/Mentzer
7. Ilic S, Petrovic M, Jaksic B, Spalevic P, Lazic L, Milosevic M (2013) Experimental analysis of picture quality after compression by different methods. Przegląd Elektrotechniczny
8. Patil TR, Baligar VP, Huilgol RP (2018) Low PSNR high fidelity image compression using surrounding pixels. In: proceedings of the 2018 international conference on circuits and systems in digital enterprise technology (ICCSDET), Kottayam, India, pp 1–6. https://doi.org/10.1109/ICCSDET.2018

Antenna at Terahertz Band: Challenges, Fabrication and Specification

Arpita Patel and Poonam Thanki

Abstract In the recent years, due to massive advances in the technology, there is a continuous demand for high data rate and bandwidth. The band in the EM spectrum that covers frequency 0.1–10 THz is known as THz band. Due to attractive features of terahertz antenna, they become popular amongst researchers. An exhaustive review of latest THz work related to THz band antenna, fabrication technology, design, THZ antenna challenges along with the measurement techniques is addressed in this paper.

Keywords Terahertz · Fabrication · Measurement · Design

1 Introduction

To fulfil the requirement of increased data traffic of future wireless communication system, terahertz frequency spectrum (0.1–10 THz) is discovered by researchers. The band from 0.1 to 0.3 THz is known as sub terahertz band [1]. The terahertz spectrum is also named as sub mm wave, near mm wave and far infrared. Figure 1 shows location of terahertz band in EM spectrum. This spectrum has good features such as high data rate, low power requirement and less prone to interference. Terahertz band shows rich potential in explosives, and biohazards scanning, screening, imaging and sensing applications. To incorporate all these features, various antenna structures are explored with the different shapes and materials to achieve miniaturization and good performance [2].

A. Patel (✉) · P. Thanki
Charotar University of Science and Technology, Changa, Gujarat, India
e-mail: arpitapatel.ec@charusat.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_21

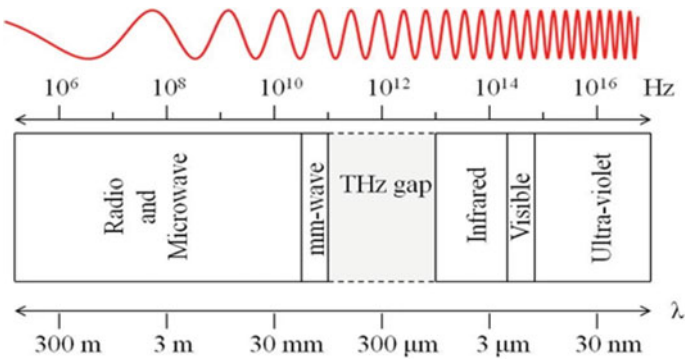


Fig. 1 THz band in EM spectrum

This paper explores various THz and sub THz band antennas with their shapes and structure. This paper also offers the recent work of various fabricated antenna and their measurement methods. In addition, literature review of various shaped antenna is also presented.

2 Terahertz Antenna

An antenna plays vital role in any communication system. THz antennas are small in size but provide wide bandwidth with higher gain and thus are important devices for transmitting and receiving THz EM waves in THz systems. They come in a wide variety of shapes and sizes. THz antennas are classified into three types: dielectric, metallic and modern material antennas. THz antennas are superior to conventional antenna in providing wide bandwidth, high resolution and strong directivity. It should be noted that, THz antenna has many challenges compared with conventional antenna. Since THz band antenna operates at high frequency, its size is greatly reduced. Compressed size, extensive bandwidth and high data rate are the major challenges in the design of antenna at THz band [2]. Also its packaging is limited by material and processing technology. The designing of THz antenna which can radiate more effectively is another challenge. In literature, various THz antennas are designed for various THz applications [3]. Micro strip patch antennas are suitable for THz communication, due to its non-planner structure and can be easily integrated with other IC.

The choice of polarization mainly depends upon the medium. For THz and satellite communication, circular polarization is preferable as they are less sensitive to multipath fading in contrast to linearly polarized wave. Additionally, due to smaller wavelength of THz antenna, very high accuracy and location system are required [4]. Also the orientation between transmitter and receiver is difficult because of reduced size of THz antenna and so circular polarization is very popular amongst researchers

as to fulfil the requirement of high gain and data rate. A high gain antenna radiating at 300 GHz was proposed for a secure wireless communication system [5]. Various fractal biological antennas topologies like retina, Wasps and DNA shape are also proposed by researchers [6]. Moreover, various shapes like *V*, *Z*, horn, hexagonal, triangle, circular, plant shape, inverted *k* shape, elliptical ring shape antennas are proposed in literature [7, 8]. Graphene is a very promising candidate for THz antenna as its conductivity is dependent on frequency, antenna based on graphene surface are expected to radiate well at THz frequency. Various terahertz antennas with graphene as a conductor have been proposed for numerous applications in literature [9].

3 THz Band Antenna Challenges

High pass loss at THZ frequency requires extremely directional antenna for future wireless communication for short distance transmission [10]. The design of ultra-wideband highly directional antenna is itself a challenge at terahertz band. Similarly, there is also a challenge of designing high gain antenna at UWB with low cost. Moreover, the shape and material affect the manufacturing cost. Moreover low-cost antenna is required for THz applications in the future. Traditional antennas manufacturing process is not useful to serve the tera band of frequency, because fabrication and installation process of THz band antenna need to be accurate with smooth surface finishing. So, low-cost materials are required to fulfil the requirement of future wireless communication system. Although graphene is a suitable candidate for THz band, it also possesses some challenges for fabrication of various shapes of antenna along with the cost. In addition, measurements of THz band antenna are again a challenge due to its small size. Figure 2 shows various THz band antenna challenges.

4 Literature Reviеw of Latest Terahertz Antenna

The latest literature review of THz band antenna and its fabrication method is presented in this section. According to latest studies of THZ antenna in [11–20], few antennas have achieved super wide bandwidth and few have achieved UWB but with low gain. The gain with above 18 dBi is required for 6G communication. Some of the works in [21–25] have achieved gain of above 18 dBi with large antenna array. There are few terahertz antennas which cover dual, tri and multi-band that are reported in literature [26–29]. Also, several antenna geometries like horn, log-periodic, SIW antenna and micro strip antenna have been reported in literature. Out of all these geometries, micro strip structures have performed better in terms of wide bandwidth and gain, but they are complex. Few micro strip antenna structures for

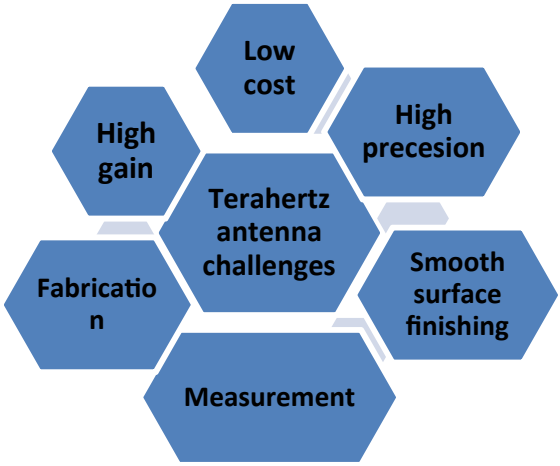


Fig. 2 THZ band antenna challenges

THz frequency spectrum that have been reported are either single band or narrow band and wideband in terahertz frequency spectrum.

Graphene is a most suitable applicant for THz radiator. It has a very high electrical conductivity compared with gold as well as silver. So, they are likely to radiate at THz frequencies. A graphene-based 2 × 2 MIMO antenna is presented in [30]. The antenna use polyimide substrate. The circular patched shape Yagi like antenna is operated over three bands in 1–30 THZ range as shown in Fig. 3. The antenna offers ultra-wide bandwidth of 10.96THz with isolation of − 26 dB, which can be used for short distance with high speed tera hertz communication.

A dual polarized micro strip patch antenna is proposed in [31]. The antenna element consists of graphene patch and has dual orthogonal ports. The antenna provides peak gain of 9.6 dB and radiates at 3.98 THz. The antenna has a stable gain and efficiency of 90%. Another graphene-based micro strip patch antenna fabricated on Arlon 1000 substrate is designed in [32]. The antenna radiates from 0.53 to 0.84 THz. Another, graphene-based circular patch antenna is proposed in [33]. The antenna radiates around 7 THz and has a gain of 7.2 dB.

Right and left hand circular polarization is achieved with *y*-shaped antenna presented in [34] as shown in Fig. 4, by adjusting the graphene conductivity the polarization rotation. Two symmetric orthogonal arms are the main key points of the design. By adjusting the electrical length of the arm, the return loss achieved is more than 12 dB.

Various shapes of THz antenna are reported in literature nowadays. A jasmine-shaped MIMO antenna using coplanar waveguide feed is designed in [35] as shown in Fig. 5.The antenna covers a band of 0.65–100 THz. It exhibits the omnidirectional radiation pattern and has a peak gain of 11 dB and isolation of more than 30 dB.

A sierpinksi carpet-shaped ultra-wideband antenna is proposed in [36] which operates over 2.2–10.5 THz range is depicted in Fig. 6. The antenna uses polyimide substrate and finds its application for weapon screening and skin imaging area.

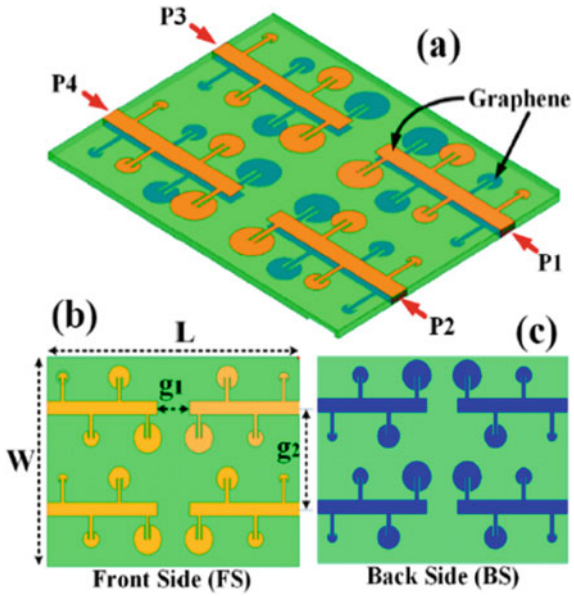


Fig. 3 Circular patch shape Yagi like MIMO antenna [30]

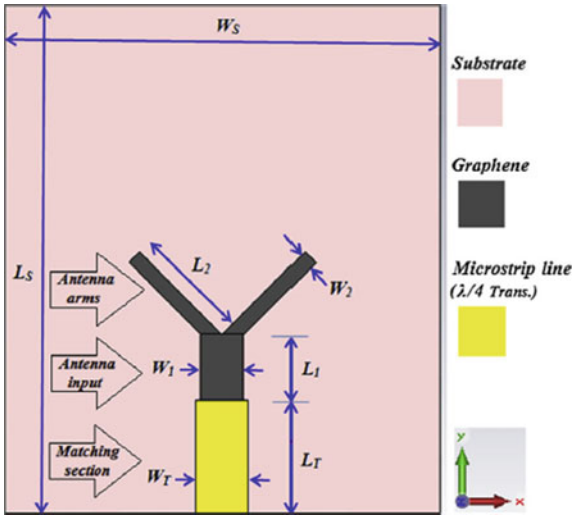


Fig. 4 Y-shaped THz antenna [34]

In [37], a low-cost butterfly-shaped antenna is proposed as shown in Fig. 7 [37]. An independent fractal antenna unit simplifies the fabrication process and so cost is very low. The antenna operates 0.1–10 Thz band and achieves a high gain of 16.95 dB. An inverted k-shaped antenna for WBAN application is proposed in [5, 38] with defected ground structure as shown in Fig. 8. It radiates at 8.8 THz and exhibits a gain of 22.1 dB.

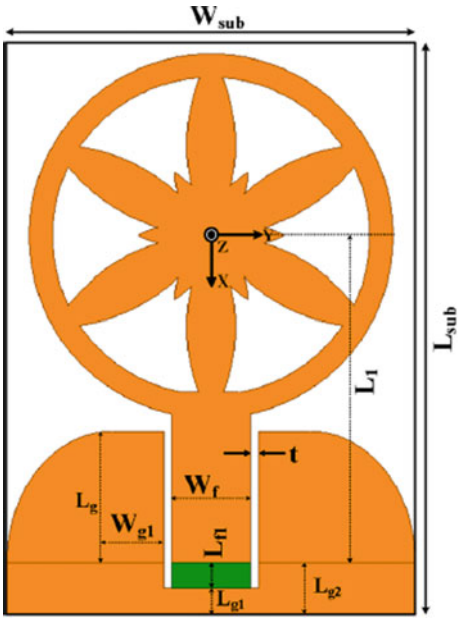


Fig. 5 Jasmine-shaped antenna [35]

A nature-inspired plant-shaped antenna for wideband application is proposed in [39] as shown in Fig. 9. The antenna has six leaves and they are sinusoidal tapered. The antenna is designed on polyimide substrate and attains wide-ranging bandwidth starting with 1.12–10.02 THz. The achieved gain is of 18.66 dBi, which makes it applicable for 6G communication.

A super wideband MIMO antenna in windmill shape is proposed in [18] as shown in Fig. 10. The antenna element is operated from 3.1 to 60 THz. Very low ECC is achieved without using any decoupling mechanism. The isolation is 30 dB in whole band. The antenna finds its application in water content imaging and weapon screening.

Another super wideband MIMO antenna with two elements is proposed in [40] as shown in Fig. 11. The antenna can be used for high-speed applications. A high isolation is achieved by *L*-shaped decoupling structure. The antenna uses RT5880 substrate and is of compact size. The radiation efficiency is of 70% and gives bandwidth from 0.33 to 10 THz.

A super wideband fractal antenna with a shape of hexagonal is proposed in [41] as shown in Fig. 12. The antenna operates in 0.2–11.5 THZ range with peak gain of 10.82 dB.

A mushroom-shaped wideband antenna with half-done ground is proposed in [42] as shown in Fig. 13. Mushroom shape is achieved by cutting semi sinusoidal in the rectangular patch. A wide bandwidth of 0.37–2.80 THz is achieved with a gain of 10.74 dBi at 1.95 THz (Fig. 14).

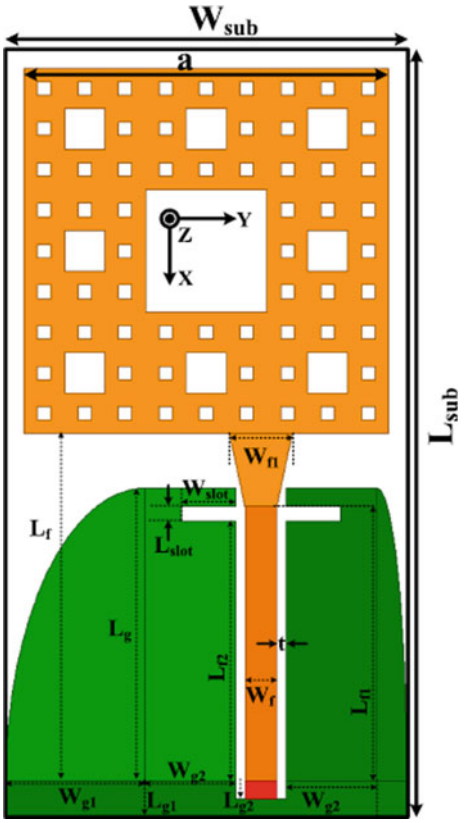


Fig. 6 Sierpinksi carpet-shaped antenna [36]

Fig. 7 Butterfly-shaped antenna [37]

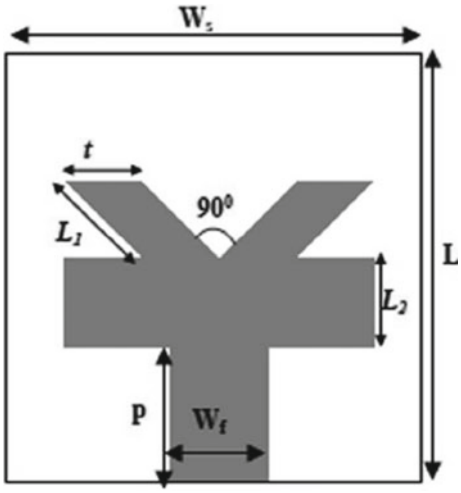


Fig. 8 Inverted K-shaped antenna [38]

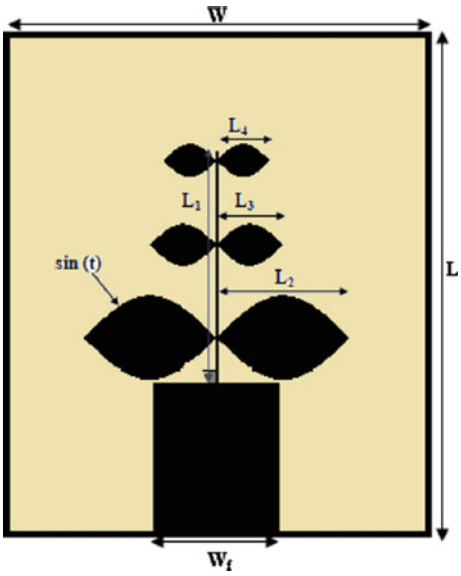


Fig. 9 Nature-inspired plant-shaped antenna [39]

A wideband antenna of rhombus shape is proposed in [43] for THZ wave. A gold metal-based section is created inside the square slot to achive the high radiation efficiency of 97.3% and wide impedance bandwidth of − 10 dB. This antenna can be used for indoor wireless communication and for defends and biomedical imaging application in THz range. A dual band fractal antenna operates in sub terahertz and tera hertz band proposed in [44]. The antenna sturcture is designed by adding a single coil as patch and then by adding four fractal coils in the shape of DNA. The antenna

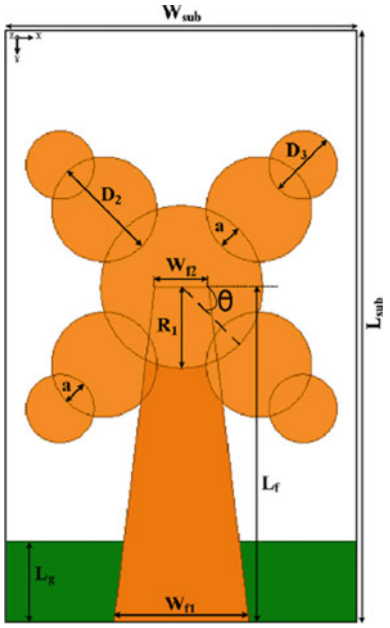


Fig. 10 Windmill-shaped antenna [18]

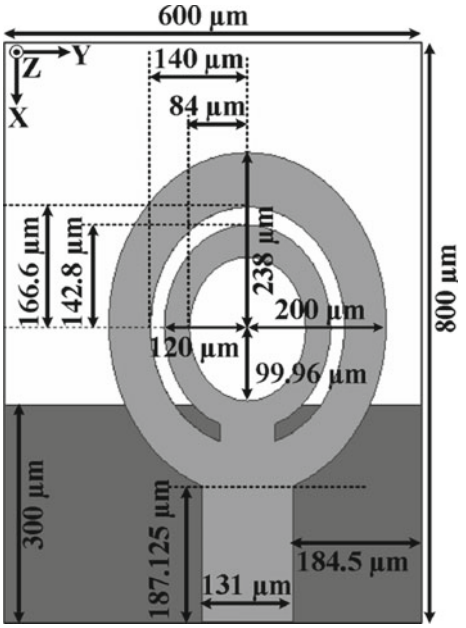


Fig. 11 Elliptical ring-shaped antenna [40]

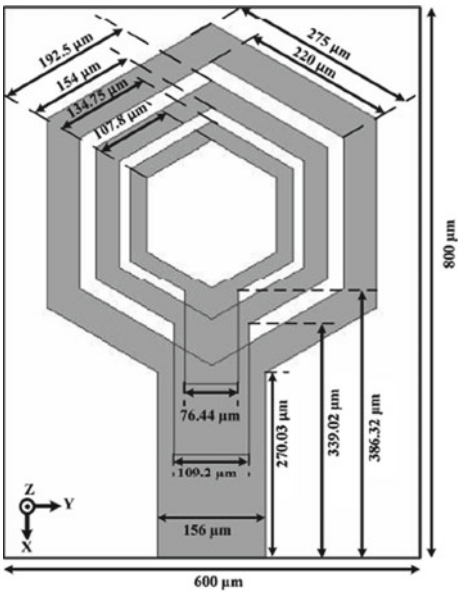


Fig. 12 Hexagonal-shaped antenna [41]

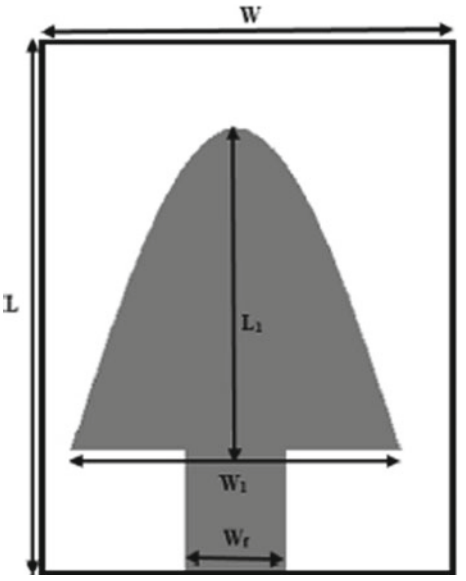


Fig. 13 Mushroom-shaped antenna [42]

radiates at 2.71 THz and exhibits gain of 9.03 dBi. The bandwidth is enhanced by the use of partial ground.

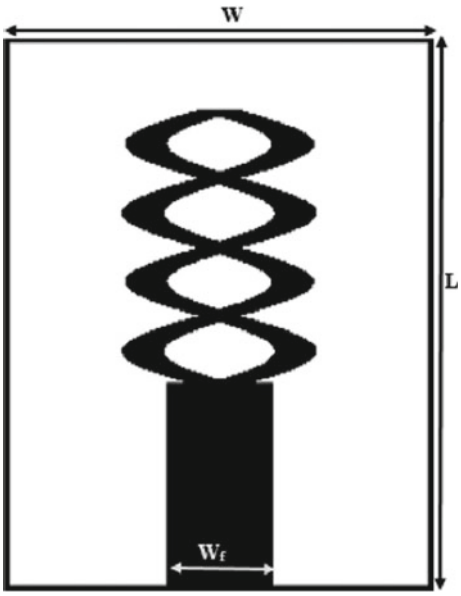


Fig. 14 DNA-shaped antenna [44]

An ultra-wideband elliptical patch antenna is proposed in [8] as shown in Fig. 15. The antenna provides a bandwidth of 0.46–5.46 THz and exhibits a gain in of 12 dBi.

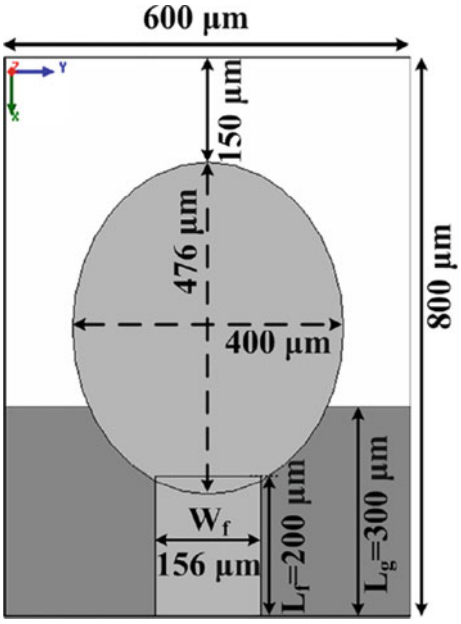


Fig. 15 Elliptical patch antenna [8]

Reconfigurable type of antenna is tuneable antenna and can be realized by applying an external voltage to change the conductivity of graphene material. A reconfigurable antenna on polyimide substrate is proposed in [45]. The five resonance mode can be excited by regulating the potential of graphene. The proposed antenna achieves radiation gain of 5.82 dB at resonance frequency of 695.55 GHz and 6.92 dBi at 705.05 GHz. Terahertz antenna can fulfil the trans receiver system requirement in satellite communication. Moreover, transparent antennas are suitable for satellite payload as size and space are of concern there. A broadband antenna suitable for transceiver wireless communication in THz band is proposed in [3]. A transparent patch antenna using Ta4C3 MXene operates at 800 GHz and can achieve a broad impedance bandwidth of 10 dB. The use of Ta4C3 MXene helps to achieve the better radiation performance compared with graphene-based transparent antenna. The radiation efficiency is > 89% in 788–822 band with the gain is of greater than 7.24 dB. A circular polarized single element patch antenna for terahertz wave is proposed in [46]. The patch of the antenna is made up of graphene. The antenna uses silicon-dioxide substrate. It radiates at 2.05 THz and has gain of 5.02 dB and achieves efficiency of 95%. Another, circularly polarized lens antenna is proposed and fabricated in [47]. The antenna is fed by pyramidal horn; the antenna is fabricated using 3D printing technique to reduce the cost. The antenna radiates at 300 GHz and achieves gain of 30 Dbic. A circular shape on chip antenna on polyimide substrate is proposed in [48] for S THz applications. The antenna element is excited by electromagnetic energy via gap of dielectric ring in ground plane sing micro strip feedline on bottom layer. The antenna exhibits large bandwidth of 0.35–0.385 Thz with an efficiency of 65.71%. A five micro strip antenna with different photonic bandgap substrate and crystals is designed in [49]. The antenna radiates from 0.5 to 0.8 THz. The high-radiation characteristics are achieved at 0.65 THz for sensing and communication application. The another best characteristic was achieved with the alternate structure gives return loss of − 83.73 dB and a wide bandwidth of more than 230 GHz with gain of 9.19 dB and radiation efficiency of 90.84%, respectively. A Vivaldi antenna which can operate for mm and terahertz band application is proposed in [50]. The performance is improved by cross-shaped resonator and rectangular slots at radiating arms. The antenna radiates around 76 GHZ and has a return loss of − 58.83 dB (Fig. 16).

5 Fabrıcatıon and Measurement Method for Tera Hertz Antenna

In Table 1, list of THz antenna which are fabricated with low-cost material is presented. These designs match their simulation results with the fabricated antenna. It is observed from the table that the fabrication has become possible for the antenna having frequency above 1 THZ only with the geometry like patch, horn lens and parabolic. Moreover, fabrication of antenna at THZ band requires great accuracy

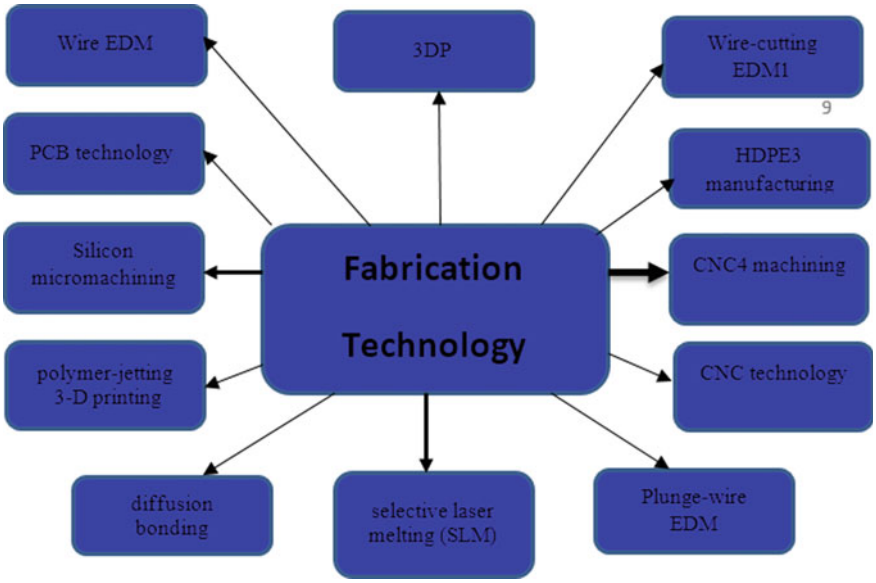


Fig. 16 THZ antenna fabrication technologies

[11]. Also, the equipment requirement is low in low terahertz band whilst maintaining the characteristics of THz wave. Figure 17 shows some latest THz band antennas which are fabricated as given in Table 1.

5.1 Measurement Method

The measurement of THz antenna is limited to 1.1 THZ as, the equipment above this range is not available due to its higher cost. Moreover, some studies in the 0.1–0.4 THz have used the traditional antenna measurement system so as to agree upon their own simulation results. As there is no anechoic chamber available for the range above 0.3 THZ, outdoor facility is considered for future study. Figure 18 shows few measurement methods of latest works of THz antenna.

Table 1 Summary of recently fabricated THZ antenna

References	Antenna type	Antenna working band	Fabrication technology
Campo et al. [21]	Conical horn	0.27–0.33 THz	Wire-cutting electrical discharge machine
Kosogor and Tikhov [22]	Elliptical lens antenna	0.14–0.22 THz	HDPE3 manufacturing process
Kushwaha and karuppanan [49]	Cassegrain	0.22–0.3 THz	CNC4 machining
Yu et al. [51]	Vivaldi patch antenna	0.56–0.74 THz	Traditional PCB etching procedure
Bhardwaj and Volakis [52]	Horn	0.22–0.32 THz	Computerized numerical control—CNC
Aji et al. [53]	Horn	0.09–0.140 THz	Plunge-wire electrical discharge machine
Pan et al. [54]	Planar bow-tie	1 THz	Antenna scaling method (antenna is fabricated at 0.001 THz)
Zhang et al. [55]	Lens	0.320–0.380 THz	Not mentioned
Tajima et al. [56]	Three horns	E-band, D-band and H-band	Selective laser melting (SLM)
Kim [57]	Step-pro_led pyramid horn antenna	300 GHz	PCB and LTCC layers
Kim et al. [58]	Corporate-feed waveguide slot array	120 GHz	Diffusion bonding technology
Tekkouk et al. [59]	Corporate-feed slotted waveguide array 16 × 16 array	350 GHz	Silicon and diffusion bonding
Wu et al. [60]	Lens antenna		Polymer-jetting 3D printing
Miao et al. [61]	Lens antenna	265–320 GHz	3D printing (in house) method with rapid prototyping
Wu et al. [62]	Wideband transmit-array	150 GHz	PCB and LTCC technology
Nayeri et al. [63]	Wideband reflect array	220 GHz	3DP technology
Carrasco and Perruisseau-Carrier [64]	Different dielectric RAs	100–600 GHz	3D printer (Objet Eden 350)
Attia et al. [65]	Wideband RA	1.3 THz	Not mentioned

(continued)

Table 1 (continued)

References	Antenna type	Antenna working band	Fabrication technology
Aqlan et al. [66]	Resonant cavity antennas (RCAs)	55.4–66.6 GHz	PCB technology
Liu et al. [67]	Fabry–Perot cavity (FPC) antenna	H-band	3DP
Zhao et al. [68]	Antipodal curvedly tapered slot	500 GHz	Silicon micromachining
Bhardwaj and Volakis [52]	Double fan-shaped slot	500 GHz	Silicon micromachining
Wu et al. [69]	Hexagonal waveguide-based horn antenna	110 GHz	Wire electrical discharge machine
[78]	Discrete dielectric lens	300 GHz	3D printing

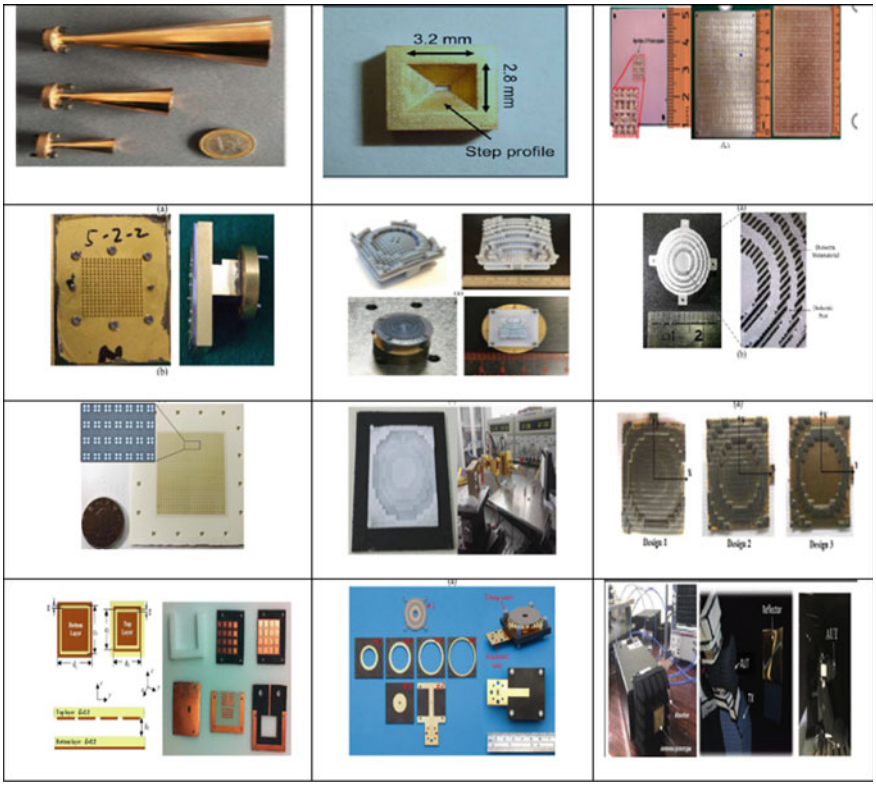


Fig. 17 Fabricated antenna at THz band

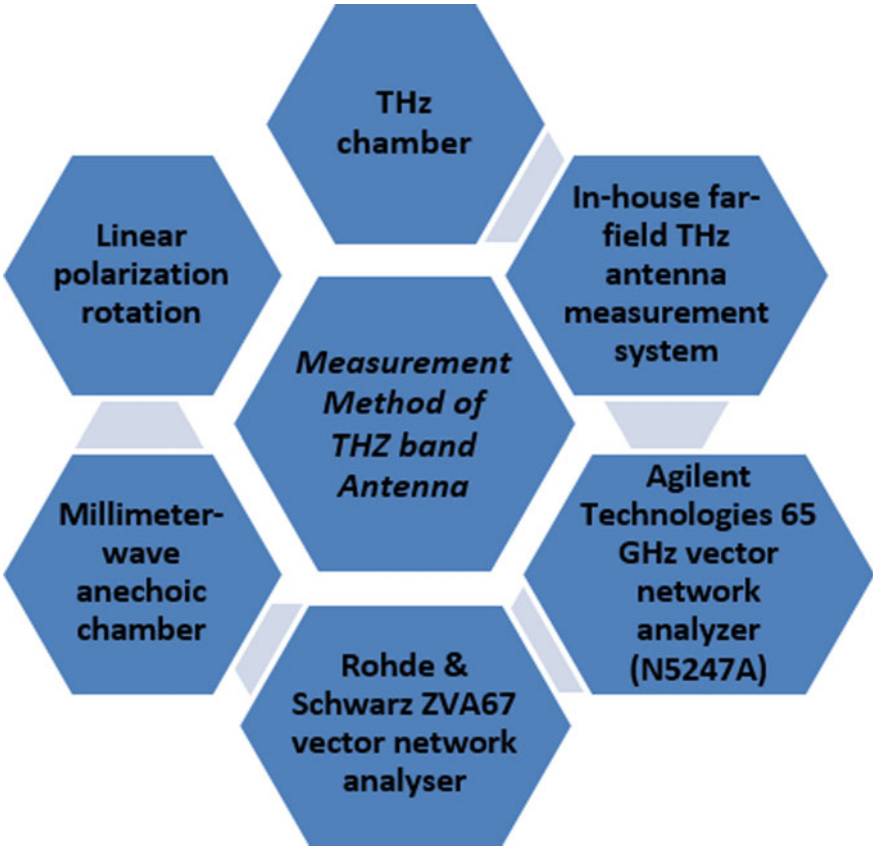


Fig. 18 Measurement method of THz band antenna

6 Conclusion

This review presents a general idea of antenna design and fabrication for terahertz and sub terahertz antenna. The antenna of various shapes, substrates and configurations were reviewed and tabulated with their performance parameters. Though these antennas achieve good performance, they are still not adequate for new THz systems, so they must be further optimized to make them suitable for emerging applications.

References

1. Hajiyat ZRM et al (2021) Antenna in 6G wireless communication system: specifications, challenges, and research directions. Optik 231:166415

2. He Y, Chen Y, Zhang L, Wong SW, Chen ZN (2020) An overview of terahertz antennas. China Commun 17(7):124–165
3. Shalini M (2021) A compact antenna structure for circular polarized terahertz radiation. Optik 231:166393
4. Zhu HT, Xue Q, Che WQ (2016) An in-house far-field THz antenna measurement system. In: Proceedings of the 2016 IEEE international workshop on electromagnetics: applications and student innovation competition (iWEM). IEEE
5. Keshwala U, Rawat S, Ray K (2020) Inverted K-shaped antenna with partial ground for THz applications. Opt Int J Light Electr Opt 219:537
6. Singh P et al (2018) Fractal and periodical biological antennas: hidden topologies in DNA, wasps and retina in the eye. Soft computing applications. Springer, Singapore, pp 113–130
7. Ananth Kumar MS (2021) A review of design and analysis for various shaped antenna in terahertz and sub-terahertz applications. Turk J Comput Math Edu 12(8):2053–2071
8. Saraereh OA et al (2022) Design and analysis of a novel antenna for THz wireless communication. Intell Autom Soft Comput 31(1):607–619
9. Ma X, Chen Z, Chen W, Chi Y, Li Z, Han C et al (2020) Intelligent reflecting surface enhanced indoor terahertz communication systems. Nano Commun Netw 24:100284
10. Poorgholam-Khanjari S, Zarrabi FB (2020) Reconfigurable Vivaldi THz antenna based on graphene load as hyperbolic metamaterial for skin cancer spectroscopy. Opt Commun 157:126482
11. Aqlan B, Himdi M, Le Coq L, Vettikalladi H (2020) Sub-THz circularly polarized horn antenna using wire electrical discharge machining for 6G wireless communications. IEEE Access 8:117245–117252
12. Strinati EC, Barbarossa S, Gonzalez-Jimenez JL, Ktenas D, Cassiau N, Dehos C (2019) 6G: the next frontier
13. Singhal S (2020) Compact superwideband terahertz antenna. In: Proceedings of the 2020 URSI regional conference on radio science (URSI-RCRS). IEEE, pp 1–4
14. Vettikalladi H, Sethi WT, Abas AFB, Ko W, Alkanhal MA, Himdi M (2019) Sub-THz antenna for high-speed wireless communication systems. Int J Antennas Propag 2019:14245
15. Singhal S (2019) Hexagonal fractal antenna for super wideband terahertz applications. Optik 128:163615
16. Hocini A, Temmar M, Khedrouche D, Zamani M (2019) Novel approach for the design and analysis of a terahertz microstrip patch antenna based on photonic crystals. Photon Nanostruct Fundam Appl 36:100723
17. Davoudabadifarahani H, Ghalamkari B (2019) High efficiency miniaturized microstrip patch antenna for wideband terahertz communications applications. Optik 194:163118
18. Singhal S (2019) Elliptical ring terahertz fractal antenna. Optik 194:163129
19. Singhal S (2019) Ultrawideband elliptical microstrip antenna for terahertz applications. Microw Opt Technol Lett 61(10):2366–2373
20. Geetharamani G, Aathmanesan T (2019) Metamaterial inspired THz antenna for breast cancer detection. SN Appl Sci 1(6):595
21. Campo MA, Carluccio G, Blanco D, Litschke O, Bruni S, Llombart N (2020) Wideband circularly polarized antenna with in-lens polarizer for high-speed communications. IEEE Trans Antennas Propag 69:43–54
22. Kosogor A, Tikhov Y (2020) A 220–300 GHz offset dual-reflector antenna for point-to-point radio. In: Proceedings of the 2020 14th European conference on antennas and propagation (EuCAP), IEEE, pp 1–3
23. Wu GB, Zeng Y-S, Chan KF, Qu S-W, Chan CH (2019) High-gain circularly polarized lens antenna for terahertz applications. IEEE Antennas Wirel Propag Lett 18(5):921–925
24. Sawada H, Kanno A, Yamamoto N, Fujii K, Kasamatsu A, Ishizu K, et al (2017) High gain antenna characteristics for 300 GHz band fixed wireless communication systems. In: Proceedings of the 2017 progress in electromagnetics research symposium-fall (PIERS-FALL), IEEE, pp 1409–1412

25. Bansal G, Marwaha A, Singh A (2020) A graphene-based multiband antipodal Vivaldi nanoantenna for UWB applications. J Comput Electron 69:1–10
26. Badr NS, Moradi G (2020) Graphene-based microstrip-fed hexagonal shape dual band antenna. Optik 202:163608
27. Zhang H-F, Liu G-B, Huang T, Zeng L (2020) Design of a frequency reconfigurable broadband THz antenna based on the vanadium dioxide. Plasmonics 89:1–7
28. George JN, Ganesh Madhan M (2017) Analysis of single band and dual band graphene based patch antenna for terahertz region. Phys E Low-Dimens Syst Nanostruct 94:126–131
29. Alharbi AG (2022) Sorathiya V (2022) Ultra-wideband graphene-based micro-sized circular patch-shaped yagi-like MIMO antenna for terahertz wireless communication. Electronics 11(9):1305
30. Shalini M, Ganesh-Madhan M (2019) Design and analysis of a dual-polarized graphene based microstrip patch antenna for terahertz applications. Optik 194:163050
31. Shamim SM et al (2021) Design and implementation of miniaturized wideband microstrip patch antenna for high-speed terahertz applications. J Comput Electr 20(1):604–610
32. Khan MD et al (2020) High-performance graphene patch antenna with superstrate cover for terahertz band application. Plasmonics 15(6):1719–1727
33. Chashmi MJ, Rezaei P, Kiani N (2020) Y-shaped graphene-based antenna with switchable circular polarization. Optik 200:163321
34. Singhal S (2021) CPW fed jasmine shaped superwideband terahertz antenna for pattern diversity applications. Optik 231:166356
35. Singhal S (2021) Asymmetrically CPW fed square Sierpinski Carpet ultra wideband terahertz antenna. Optik 242:167056
36. Shi Y et al (2021) (2021) Design of terahertz detection antenna with fractal butterfly structure. IEEE Access 9:113823–113831
37. Keshwala U, Rawat S, Ray K (2020) Inverted K-shaped antenna with partial ground for THz applications. Optik 219:165092
38. Keshwala U, Rawat S, Ray K (2020) Plant shaped antenna with trigonometric half sine tapered leaves for THz applications. Optik 223:165648
39. Singhal S (2020) Four arm windmill shaped superwideband terahertz MIMO fractal antenna. Optik 219:165093
40. Singhal S (2020) Hexagonal fractal antenna for super wideband terahertz applications. Optik 206:163615
41. Keshwala U, Ray K, Rawat S (2021) Ultra-wideband mushroom shaped half-sinusoidal antenna for THz applications. Optik 228:166156
42. Krishna CM et al (2021) A micro-sized rhombus-shaped THz antenna for high-speed short-range wireless communication applications. Plasmonics 16(6):2167–2177
43. Keshwala U, Rawat S, Ray K (2021) Design and analysis of DNA shaped antenna for terahertz and sub-terahertz applications. Optik 232:166512
44. Jha KR, Singh G (2010) Analysis and design of rectangular microstrip antenna on two-layer substrate materials at terahertz frequency. J Comput Electron 9(2):68–78
45. Anisha A, Sriram-Kumar D (2022) Performance analysis of Ta_4C_3 MXene based optically transparent patch antenna for terahertz communications. Optik 260:168959
46. Wu GB et al (2019) High-gain circularly polarized lens antenna for terahertz applications. IEEE Antennas Wireless Propag Lett 18(5):921–925
47. Alibakhshikenari M et al (2020) High-gain metasurface in polyimide on-chip antenna based on CRLH-TL for sub-terahertz integrated circuits. Sci Rep 10(1):1–9
48. Temmar MNE et al (2019) Analysis and design of a terahertz microstrip antenna based on a synthesized photonic bandgap substrate using BPSO. J Comput Electr 18(1):231–240
49. Kushwaha RK, Karuppanan P (2019) Design and analysis of Vivaldi antenna with enhanced radiation characteristics for mm-wave and THz applications. Opt Quant Electr 51(9):309
50. Khan MAK, Shaem TA, Alim MA (2020) Graphene patch antennas with different substrate shapes and materials. Optik 202:163–700

51. Yu H-Y, Yu J, Yao Y, Liu X, Chen X (2019) Wideband circularly polarised horn antenna with large aspect ratio for terahertz applications. Electron Lett 56(1):11–13
52. Bhardwaj S, Volakis JL (2018) Hexagonal waveguide based circularly polarized horn antennas for sub-mm-wave/terahertz band. IEEE Trans Antennas Propag 66(7):3366–3374
53. Aji AP, Apriono C, Zulkifli FY, Rahardjo ET (2017) Radiation pattern validation of a THz planar bow-tie antenna at microwave domain by scaling up technique. In: Proceedings of the 2017 ınternational conference on radar, antenna, microwave, electronics, and telecommunications (ICRAMET), IEEE, pp 108–111
54. Pan W, Zeng W, Yu X, Zhang J (2016) A novel high-gain directional lens antenna for terahertz band. Prog Electromagn Res 68:107–117
55. Zhang B, Zhan Z, Cao Y, Gulan H, Linner P, Sun J, Zwick T, Zirath H (2016) Metallic 3-D printed antennas for Millimeter- and submillimetre wave applications. IEEE Trans THz Sci Technol 6(4):592–600
56. Tajima T, Song H-J, Ajito K, Yaita M, Kukutsu N (2014) 300-GHz step-pro_led corrugated horn antennas integrated in LTCC. IEEE Trans Antennas Propag 62(11):5437–5444
57. Kim D (2013) 4–4-element corporate-feed waveguide slot array antenna with cavities for the 120 GHz-band. IEEE Trans Antennas Propag 61(12):5968–5975
58. Kim D, Hirokawa J, Ando M, Takeuchi J, Hirata A (2014) 64–64-element and 32–32-element slot array antennas using doublelayer hollow-waveguide corporate-feed in the 120 GHz band. IEEE Trans Antennas Propag 62(3):1507–1512
59. Tekkouk K, Hirokawa J, Oogimoto K, Nagatsuma T, Seto H, Inoue Y, Saito M (2017) Corporate-feed slotted waveguide array antenna in the 350-GHz band by silicon process. IEEE Trans Antennas Propag 65(1):217–225
60. Wu G-B, Zeng Y-S, Chan KF, Qu S-W, Chan CH (2019) 3-D printed circularly polarized modied fresnel lens operating at terahertz frequencies. IEEE Trans Antennas Propag 67(7):4429–4437
61. Miao Z-W, Hao Z-C, Luo GQ, Gao L, Wang J, Wang X, Hong W (2018) 140 GHz high-gain LTCC-integrated transmit-array antenna using a wideband SIW aperture-coupling phase delay structure. IEEE Trans Antennas Propag 66(1):18190
62. Wu MD, Li B, Zhou Y, Guo DL, Liu Y, Wei F, Lv X (2018) Design and measurement of a 220 GHz wideband 3-D printed dielectric re-ectarray. IEEE Antennas Wireless Propag Lett 17(11):2094–2098
63. Nayeri P, Liang M, Sabory-Garcia RA, Tuo M, Yang F, Gehm M, Xin H, Elsherbeni AZ (2014) 3D printed dielectric re_ectarrays: Lowcost high-gain antennas at sub-millimeter waves. IEEE Trans Antennas Propag 62(4):2000–2008
64. Carrasco E, Perruisseau-Carrier J (2013) Re-ectarray antenna at terahertz using graphene. IEEE Antennas Wireless Propag Lett 12:253–256
65. Attia H, Abdelghani ML, Denidni TA (2017) Wideband and high-gain millimeter-wave antenna based on FSS Fabry_Perot cavity. IEEE Trans Antennas Propag 65(10):5589–5594
66. Aqlan B et al (2020) Sub-THz circularly polarized horn antenna using wire electrical discharge machining for 6G wireless communications. IEEE Access 8:117245–117252
67. Liu Y, Lu H, Wu Y, Cui M, Li B, Zhao P, Lv X (2016) Millimeter-wave and terahertz waveguide-fed circularly polarized antipodal curvedly tapered slot antennas. IEEE Trans Antennas Propag 64(5):1607–1614
68. Zhao P, Liu Y, Lu H, Wu Y, Lv X (2017) Experimental realization of terahertz waveguide-fed circularly polarized double-fan-shaped slot antenna. IEEE Antennas Wireless Propag Lett 16:2066–2069
69. Wu GB, Zeng YS, Chan KF, Qu SW, Chan CH (2019) High- gain circularly polarized lens antenna for terahertz application. IEEE Antennas Wireless Propag Lett 18(5):921–925

Support Vector Machine Classifier for Wheat Grain Identification Based on Grid Search Optimization Technique

Nabin Kumar Naik, Prabira Kumar Sethy, Millie Panigrahi, and Santi Kumari Behera

Abstract India is one of the biggest places that grows and eats wheat grains, so it makes sense that they would want to know the quality and variety of wheat to meet the needs of the food industry. Moreover, because the prices of different types of grains vary in different parts of the country, it is important for both customers and farmers to be able to identify and sort grains based on the end-products, demand, and costs of each type. For quality review in the grain supply chain, it is important to identify and verify wheat varieties. In this paper, an approach based on machine learning and optimization techniques is proposed. Using a grid search method, the support vector machine (SVM) was able to classify three types of wheat grains (Kama, Rosa, and Canadian) with a validation accuracy of 100% and a test accuracy of 100%.

Keywords Wheat grain · SVM · Identification · Grid search algorithm · Optimization technique

1 Introduction

According to the FAO's preliminary prediction, worldwide wheat production will fall year over year in 2023 but will still be the second largest on record at 784 million tons. High prices in North America have encouraged farmers to plant more winter wheat in 2023 than they did since 2015. Despite the fact that dry weather continues to be an issue in the Central Plains [1], an increase in overall wheat production from the lower outturn in 2022 to 51 million tons in 2023 would be the largest output in three years. Such a large production of wheat is a matter of concern for people to

N. K. Naik · P. K. Sethy (✉)
Department of Electronics, Sambalpur University, Burla 768019, India
e-mail: prabirsethy.05@gmail.com

M. Panigrahi
Department of ECE, Trident Academy of Technology, Bhubaneswar 751014, India

S. K. Behera
Department of CSE, VSSUT, Burla 768018, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_22

categorize seeds manually [2]. Most crop classification is based on human intuition. Seed categorization plays a crucial role and should be improved. Determining the wheat variety by hand requires time and skilled judgment. It is difficult to determine the differences between seeds in a collection when they all appear identical [3]. The development and widespread application of computer-aided data analytics are a result of recent advances in information science [4]. The ability to examine an object several times and analyze its constituents at any point inside a sample is made possible by the three-dimensional nature of CT scanning. Recent investigations have shown that X-ray images offer an alternative method for the non-destructive measuring methodology [5, 6].

In the past, machine learning methods were implemented effectively in numerous supply chain processes to categorize seeds and grains and aid the productivity of food grains. To separate the seven different types of dry beans, Koklu and Ozkan [7] developed an automated system based on machine learning. Four classification models were trained using the same set of 12 dimensions and four shape features (MLP, SVM, KNN, and DT). The SVM classifier achieved the highest accuracy (93.13%) among the tested classifiers. Machine learning and computer vision-based automatic categorization systems offer fast and high-throughput capabilities [3]. K-nearest neighbor (KNN), regression trees (RegTree), and Gaussian NB were used for classification with accuracy calculations of 92, 94, and 92%, respectively. The ensemble classifier, which is 95% accurate, is based on hard voting conclusions. For the advantage of farmers, Agrawal and Dahiya [8] combined nonlinear (KNN, CART, NB, and SVM) and basic linear (LR and LDA) seed classification techniques. Fazel et al. [9] suggested an ML-based categorization technique for seven different wheat types. The linear SVM, quadratic SVM, and cubic SVM algorithms were all developed with this specific goal in mind. With the use of form, color, and texture data, QSVM achieved the highest accuracy of any technique, 97.6%. When classifying rice varieties in Asian nations, Qadri et al. [10] used five different machine vision techniques, considering ROI, and the LMT-tree classifier reached a maximum overall accuracy of 97.4%. In [11], the classification of maize seeds used a tenfold cross-validation method for model optimization. To achieve an accuracy of 96.46% with the SVM, five different types of maize seed images were used in a GUI-developed varietal classification model. Six different varieties of maize seeds have been categorized by Ali et al. [12] with an accuracy rate of 98.93%, and MLP has the highest potential among the four classifiers studied (Random Forest, BayesNet, LogitBoost, and MLP).

The goal of this work is to create a system based on machine learning that can identify wheat varieties from X-ray images and categorize them into three primary types, making the duties of farmers simpler and easier. Therefore, a classifier that can accurately predict the category of unknown or new data and be supervised trained with a high number of attributes estimated for a limited set of examples is required. The primary goals of this work are to extract geometrical features from wheat grain samples in order to properly anticipate the class of a new sample and to identify which classifier best matches the model generated from those features with grid search optimization evaluating the performance of the classifier for three-class classification.

These goals can be accomplished by determining the classifier that best matches the model created using these features.

The three key findings of this study are as follows:

- Wheat grains have access to fully automated end-to-end grading systems.
- A supervised machine learning framework was developed that uses a grid search algorithm to optimize the prediction model.
- A performance evaluation was performed based on the accuracy of predicting the three varieties of wheat grains.

2 Materials and Methodology

2.1 About Dataset

The UCI [13] seed dataset contained 210 wheat kernel samples. In addition to the class attribute, each instance included seven features. All samples shared the same values for the seven characteristics. The sample size of 70 grains included three varieties of wheat: Canadian, Kama, and Rosa. Soft X-rays were employed to obtain high-resolution images of the inner structure of the kernel. An X-ray image of a wheat kernel is shown in Fig. 1 [14]. It is not as invasive or costly as other imaging methods such as scanning microscopy or laser technology. "KODAK X-ray plates" are used to store the images.

The X-ray photograms were scanned using an "Epson Perfection V700 table photo-scanner with a built-in transparency adaptor, 600 dpi resolution, and 8-bit grayscale levels". The methods used to assess the bitmap graphics files that were collected were built using the computer software program GRAINS, which was created specifically for X-ray diagnostics of wheat kernels [15, 16]. To construct the

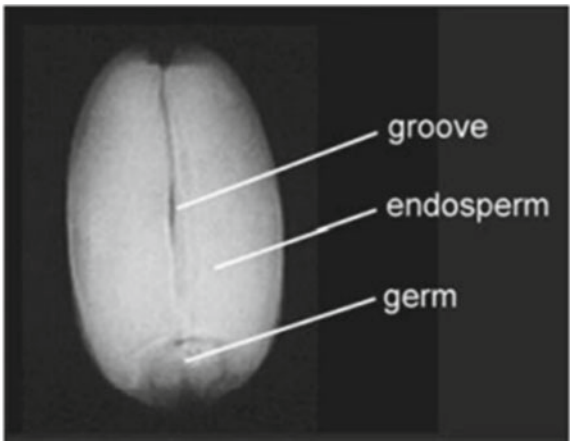


Fig. 1 X-ray image of a wheat kernel

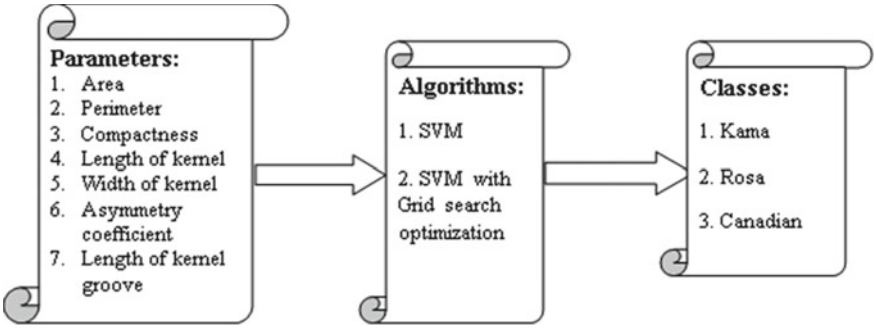


Fig. 2 Block diagram of the classification process

data, seven geometric characteristics of the wheat kernels were measured from 210 samples. The block diagram in Fig. 2 shows the factors considered when classifying wheat into three groups.

2.2 Methodology

This study proposes a method for the classification of three categories of wheat using a supervised machine learning algorithm with grid search optimization. The data input given to the model includes the seven geometrical features and statistical parameters of the SVM, minimum value, maximum value, number of steps, and scaling methods. A grid search space was built for parameter selection by training an SVM. Here, cross-validation was implemented by considering 70% training, 20% validation, and 10% testing on the entire dataset with three classes. Then, all parameter combinations were evaluated using grid search optimization, and the best features were considered for retraining the model.

The performance evaluation is conducted for multi-class classification using SVM and accuracy is calculated for two experiments:

- Performance of SVM on validated and test data.
- Performance of SVM with grid search algorithm on validated and test data.

The flowchart is shown in Fig. 3. Detailed step-by-step procedure for the prediction of wheat types and performance evaluation process of Proposed Model.

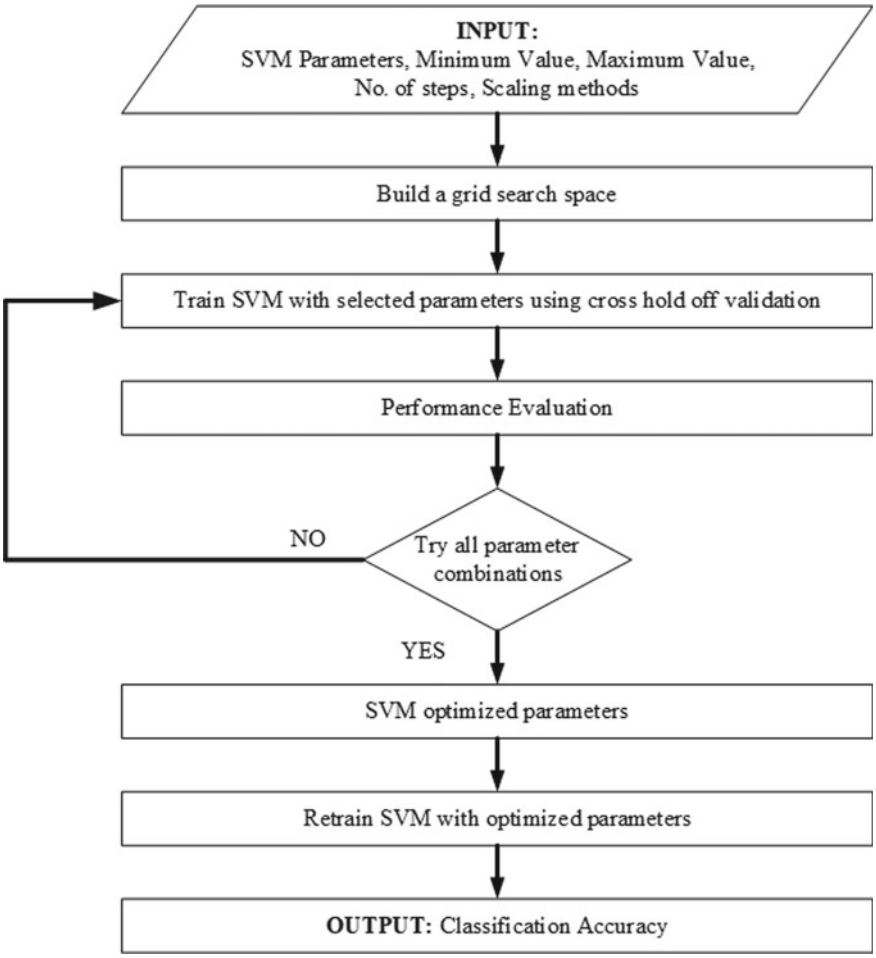


Fig. 3 Flowchart of the proposed model

3 Result and Discussion

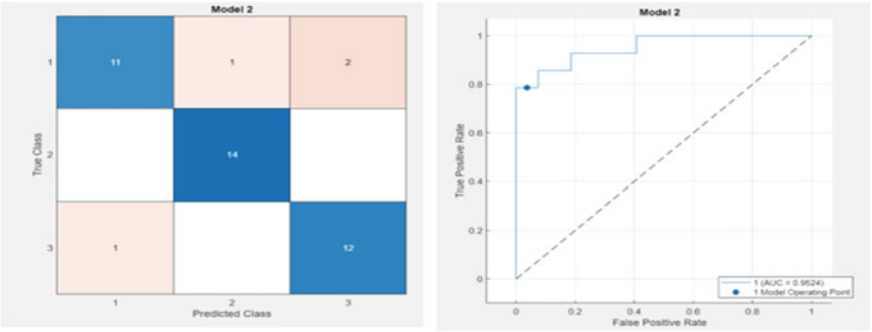
The findings of the suggested method are presented in this section, and the same dataset was used to investigate the different machine learning models. Model performance was assessed using the test data. The proposed model was constructed in MATLAB 2022b on an Intel(R) Core (TM) i7 CPU of the 12th generation. Cross-hold-off validation methods were used to assess the effectiveness of each model. To analyze the effectiveness of the grid search algorithm, an SVM without any optimization and an SVM with grid search optimization were evaluated separately. The dataset is split into training:validation:testing which is 70:20:10. The confusion matrix and

AUC of the SVM on the validated data and test data are illustrated in Figs. 4 and 5, respectively.

To improve the classification accuracy, a grid search algorithm was introduced with SVM. The hyperparameter optimization using the grid search algorithm for wheat grain classification is illustrated in Fig. 6, and the best parameters are listed in Table 1.

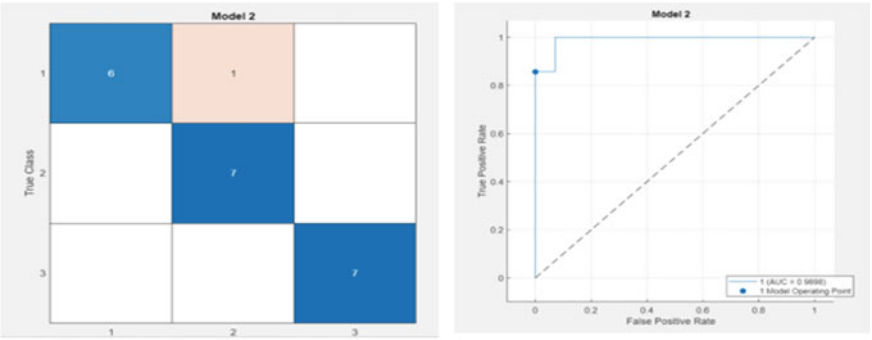
The performance of the SVM with grid search algorithm on validated data and test data is illustrated in Figs. 7 and 8, respectively.

It was observed from the findings that the support vector machine has (accuracy of 95.2% and an AUC of 0.9524) for the validated data and (accuracy of 95.24% and AUC of 1) on the test data. Further, the support vector machine with the grid search algorithm has (accuracy of 100% and AUC of 1) on validated data and (accuracy of 100 and AUC of 1) on test data. This implies a 5% enhancement in accuracy with the inclusion of the grid search algorithm.



(a) (b)

Fig. 4 Performance of support vector machine on validated data **a** Confusion matrix, **b** AUC



(a) (b)

Fig. 5 Performance of support vector machine on test data. **a** Confusion matrix, **b** AUC

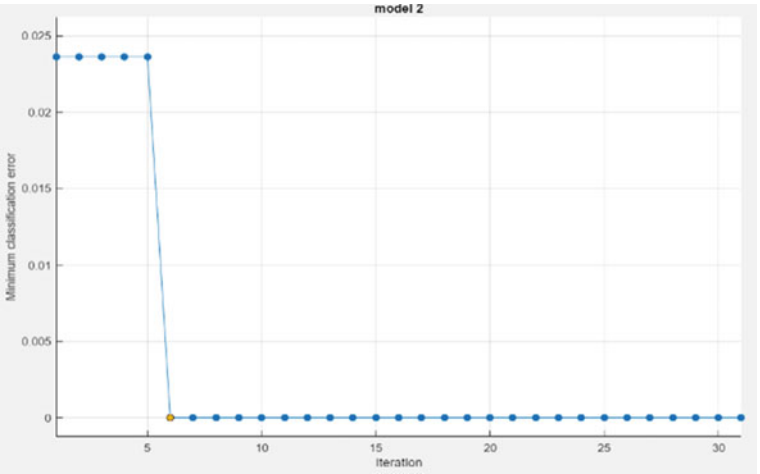


Fig. 6 Hyperparameter optimization of SVM using grid search algorithm for wheat grain classification

Table 1 Hyperparameter setting of grid search algorithm

Hyperparameter	Value
Multiclass method	One-Vs-One
Box constraint	215.4435
Kernel frame	Linear

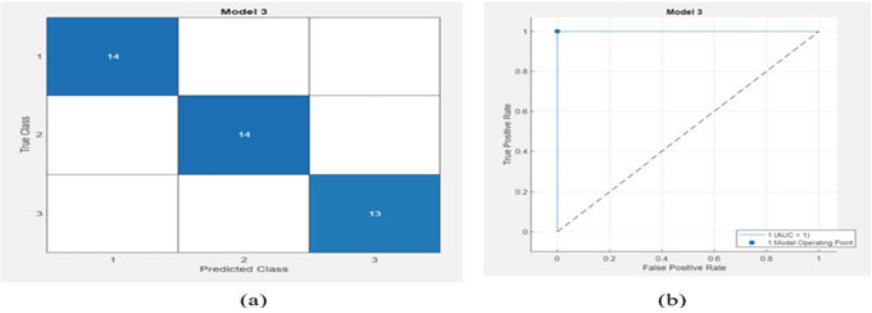
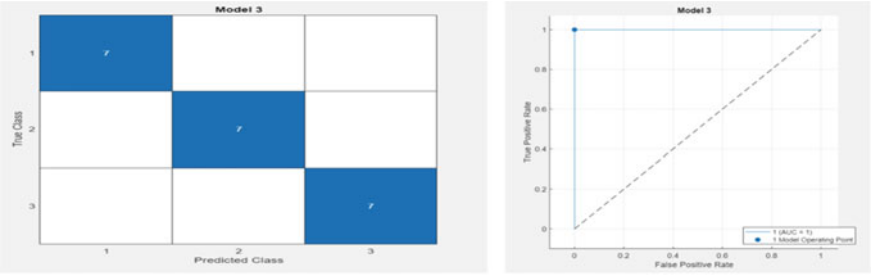


Fig. 7 Performance of SVM with grid search algorithm on validated data **a** Confusion matrix **b** AUC

The efficiency of the classifiers in the categorization of different varieties of grains using ML is shown in Table 2 in a comparison analysis. The majority of cutting-edge ML techniques have been created for the classification of various species or types of seed samples. In contrast, this study examined the effectiveness of several



(a) (b)

Fig. 8 Performance of SVM with grid search algorithm on test data **a** Confusion matrix **b** AUC

Table 2 Comparison of classification models for grain varieties' identification in existing work

Reference	Seeds	Varieties	Classification technique used	Accuracy with best classifier
Agrawal and Dahiya [7]	Dry beans	7	MLP, SVM, KNN, DT	SVM with 93.3% accuracy
Khatri et al. [3]	Wheat	3	KNN, DT, NB, Ensemble	Ensemble classifier with 95% accuracy
Qadri et al. [9]	Wheat	7	LDA, QDA, LSVM, QSVM, CSVM	QSVM with 97.6% accuracy
Lingwal et al. [17]	Wheat	15	Deep CNN	94.88% training accuracy and 97.53% test accuracy
Ali et al. [11]	Maize	5	MLP, DT, LDA, NB, SVM, KNN, AdaBoost	SVM with 96.46% accuracy
Seeds Data Set [12]	Maize	6	Random Forest, Bayes Net, Logit Boost, and MLP	MLP with 98.93%
Proposed	Wheat	3	SVM, SVM with grid search optimization	SVM with grid search optimization with 100% validation and test accuracy

classifiers in the classification task of rating seed quality, thereby simplifying the job of agriculturalists.

4 Conclusion

Wheat grain identification is an important aspect of consumer demand for high-quality wheat flour. Furthermore, it is also important for seed culture. In this study, a support vector machine with a grid search optimization technique-based classification method is proposed for the classification of three varieties of wheat grains. The

three varieties are Canadian, Kama, and Rosa. The introduction of a grid search algorithm for hypertuning the SVM has enhanced accuracy by 5%. In the SVM (accuracy of 95.2%, AUC of 0.9524) on validated data and (accuracy of 95.24%, AUC of 1) on the test data. Further, the SVM with the grid search algorithm has (accuracy of 100% and AUC of 1) on validated data and (accuracy of 100 and AUC of 1) on test data.

References

1. FAO (2022). World food situation. https://www.fao.org/worldfoodsituation/csdb/en/. Accessed 04 Feb 2023
2. Chitra HSH, Suguna S, Sujatha SN (2016) A survey on image analysis techniques in agricultural product. Indian J Sci Technol 9(12):1–13
3. Khatri A, Agrawal S, Chatterjee JM (2022) Wheat seed classification: utilizing ensemble machine learning approach. Sci Program 2022:1–9
4. Giraldo P, Benavente E, Manzano-Agugliaro F, Gimenez E (2019) Worldwide research trends on wheat and barley: a bibliometric comparative analysis. Agronomy 9(7):352
5. Charytanowicz M, Niewczas J, Kulczycki P, Kowalski PA, Łukasik S (2016) Discrimination of wheat grain varieties using X-ray images. In: Pietka E, Badura P, Kawa J, Więcławek W (eds) Information technologies in biomedicine, advances in intelligent systems and soft computing, vol 471. Springer, New York, pp 39–50
6. Koklu M, Ozkan IA (2020) Multiclass classification of dry beans using computer vision and machine learning techniques. Comput Electr Agricult 174:105507
7. Agrawal D, Dahiya P (2018) Comparisons of classification algorithms on seeds dataset using machine learning algorithm. Compusoft 7(5):2760–2765
8. Fazel-Niari Z, Afkari-Sayyah AH, Abbaspour-Gilandeh Y, Herrera-Miranda I, Hernández-Hernández JL, Hernández-Hernández M (2022) Quality assessment of components of wheat seed using different classifications models. Appl Sci 12:4133
9. Qadri S, Aslam T, Nawaz SA, Saher N, Razzaq A, Ur-Rehman M, Furqan Qadri S (2021) Machine vision approach for classification of rice varieties using texture features. Int J Food Propert 24(1):1615–1630
10. Xu P, Yang R, Zeng T, Zhang J, Zhang Y, Tan Q (2021) Varietal classification of maize seeds using computer vision and machine learning techniques. J Food Process Eng 44(11):e13846
11. Ali A, Qadri S, Mashwani WK (2020) Machine learning approach for the classification of corn seed using hybrid features. Int J Food Prop 23(1):1110–1124
12. Seeds Data Set (2023) An open-source online data repository hosted at UCI machine learning repository. https://archive.ics.uci.edu/ml/datasets/seeds. Accessed 02 Feb 2023
13. Charytanowicz M, Niewczas J, Kulczycki P, Kowalski PA, Łukasik S (2016) Discrimination of wheat grain varieties using x-ray images. In: Information technologies in medicine: 5th international conference, ITIB Kamień Śląski, vol 1. Proceedings Springer International Publishing, Poland, pp 39–50
14. Sharma A, Jain A, Gupta P, Chowdary V (2020) Machine learning applications for precision agriculture a comprehensive review. IEEE Access 9:4843–4873
15. Mohan A (2022) Machine learning based crop yield prediction using spectral images. Doctoral dissertation, National Institute of Technology Karnataka, Surathkal
16. Charytanowicz M, Niewczas J, Kulczycki P, Kowalski PA, Łukasik S, Żak S (2010) Complete gradient clustering algorithm for features analysis of x-ray images. Inform Technol Biomed 2:15–24
17. Lingwal S, Bhatia KK, Tomer MS (2021) Image-based wheat grain classification using convolutional neural network. Multimedia Tools Appl 80:35441–35465

RTOS Schedulers for Periodic and Aperiodic Taskset

Dhruva R. Rinku, M. Asha Rani, and Y. Krishna Suhruth

Abstract A scheduler is a kernel object that allocates processor time between various jobs depending on any particular strategy, including but not limited to deadlines, frequency of occurrence, reaction time, throughput, etc. For environments with single-core and multi-core processors, the scheduler operates differently. In overload conditions where single core is failed to schedule available tasks, multicore gives guarantee to schedule tasks without missing deadlines. Programmers such as nearest deadline first and monotonic rate are well-accepted in real-time operating systems. The most common real-time scheduling techniques described in this article can run on multiple CPUs. Using single-cost and deadline-nearest-first scheduling approaches, these approaches solve the problem of ordering periodic and early task groups. To evaluate and assess the effectiveness of the algorithms, computer tests are carried out. In this research, we suggested and witnessed the modeling of various real-time scheduling algorithms in single core and multicore, we compared different functions such as response time and deadline, CPU usage, etc., and observed that the scheduler schedules input tasks in multicore without missing deadline when single core was failed to schedule available tasks.

Keywords Multicore · Response time · Deadline · Utilization factor · Scheduling algorithm

D. R. Rinku (✉)
Department of Electronics Communication and Engineering, CVR College of Engineering, Hyderabad, India
e-mail: dhruva.rinku@cvr.ac.in

M. Asha Rani
Department of Electronics Communication and Engineering, JNTUH, Hyderabad, India

Y. Krishna Suhruth
Department of Computer Science Engineering, CVR College of Engineering, Hyderabad, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_23

1 Introduction

Real-time systems are those whose proper function relies not only on the conclusions drawn from reasoning, but also on the instant in which those conclusions are reached. These systems, which are of a high level of intricacy, are used in a variety of fields, including automation, avionics, networked systems, and multimedia. Multiple CPUs are needed to handle the processing needs of these apps. The scheduler is a component of the operating system [1–3] that decides which task and which function will execute next. The task's schedulers make sure it receives the system tools it needs to finish the given job. Schedulers are created to maintain optimum utilization of all computer resources, efficiently distribute system resources among numerous apps, or accomplish a particular service result. Scheduling is an essential aspect of any computing system that allows a computer to perform multiple activities simultaneously on a single CPU unit [1, 3–6]. A scheduler may be created to achieve any number of objectives, such as increasing throughput, reduce response time or latency, and ensure fairness when allocating resources or between processes in general, running at the right time depending on the workload and priority of each process [4]. Because it is common for these criteria to be implemented in disagreement with one another, a planner must find an appropriate compromise while taking the application's most important requirements into consideration. Depending on the goal prerequisite of the application, any one of the aforementioned worries is given preference. The planner must ensure that processes can make goals in real-time settings in order to maintain system stability. Without the scheduler, activities might not run in the sequence that a user or the operating system itself desires [3], making it one of the most crucial operating system components. Task scheduling for general OS may be done by selecting one of the available scheduler programs, such as Round-Robin, based on priority, first come first served, least task first, etc. However, for a real-time OS as the name suggests, scheduler should give guarantee of meeting deadline for each task. Scheduler for real-time system decides priority of task to be executed and depends on temporal parameters such as deadline, response time, occurrence rate of periodic tasks, etc. The study of real-time scheduling in a single processor is the subject of a large number of research projects nowadays [2, 6–10].

In this paper, we evaluate and contrast how various types of jobs perform when scheduling using the rate monotonic (RM) and earliest deadline first (EDF) methods. In over loaded condition which the scheduler needs to select by operating system is a question of the hour.

Here's how the remainder of the text is laid out: Article classification and sample procedures are discussed in Sect. 2 of this document. Section 3 addresses the usage of simulators. In Sect. 4 RM scheduling technique is analyzed for the periodic taskset. In Sect. 5 EDF approach is discussed for different type of jobs. Section 6 discussed handling sporadic tasks by scheduler. Section 7 discussed schedulability condition-based scheduler.

2 System Model

In this study, we focus on analyzing a particular collection of real-time functions $\tau = \{\tau 1, \tau 2, \ldots, \tau n\}$, in the treatment group, $P = \{P1, P2, \ldots, Pm\}$. In computer systems, a function is usually a string or a process. Period *Ti*, end time *Di*, and execution time *Ci* are the parameters characterizing the function. Period and execution time τ satisfy $Ti > 0$ and $0 < Ci\ Ti = Di$ $(i = 1, \ldots, n)$.

$Ui = Ci/Ti$ is selected as the factor of utilization for activity *i*. If we define maximum utilization of each periodic job in the system as, then the utilization factor of any given group of tasks is just the sum of the utilizations of the tasks within that set $i = 1, \ldots, n(Ci/Ti)$. In this study, we consider a strict physical timing scheme, in which *i* is released at time *t*, completion of work done must take place after *t* or in time period (t, t) in many units (d).

The following limits are applicable to the paradigm used in this study. The jobs are stand alone. In other words, the arrival of any other tasks in the system has no impact on the arrival of any particular task. The expense of changing the jobs' contexts is regarded as insignificant. Other than the CPU, no resources are shared between jobs. Access control systems cost zero. Each task is assigned to no more than one CPU at a moment, and the expense of transfer is regarded as zero.

For similar multiprocessors (processors running at the same speed), efficiency study of various scheduling methods is done in this work. Task priorities are given statically (using RM) and dynamically (using EDF).

3 Implementation and Analysis of Schedulers

Implemented and analyzed schedulers for periodic and aperiodic tasks with the use of SimSo Simulator. SimSo is a scheduling simulator for real-time multiprocessor architectures that uses statistical models to consider scheduling overheads (scheduling choices, context changes, and the influence of caches). Based on a discrete event simulator, it enables rapid scheduling policy development using Python.

4 Rate Monotonic Scheduler (RM)

The rate monotonic algorithm is a method of giving jobs set priority in order to maximize their "schedulability." If every job in a group meets its goals consistently, it is said to be schedulable. Each task's priority is determined by the algorithm based on how long it has been since it was last completed; hence, the shorter the duration, the greater the task's priority [11].

A fixed-priority method that works best is RM [12, 13]. If the RM algorithm cannot be used to schedule a taskset, no fixed-priority method can be used to schedule it.

Condition for Scheduling in RM by Liu and Layland: Liu and Layland suggested scheduling criteria for groups of tasks planned in RM based on processor utilization [7]. In this case, a utilization limit based on the total number of tasks in the system is used to measure the utilization of the task pool. If the following criteria is met, a job group will not skip any deadlines:

$$\sum_{i=1}^{n} Ci/Ti \leq n\left(2^{1/n} - 1\right) \tag{1}$$

In Eq. 1, C is execution time, T is period of occurrence, and n is number of tasks.

To assess this, scheduler program has been written and three periodic tasksets T1 (2, 0.5), T2 (3, 1.2), T3 (6, 0.5), generated for rate monotonic scheduler where, 2, 3, and 6 are the respective periods of task 1, task 2, and task 3. 0.5, 1.2, and 0.5 are the execution time of task 1, task 2, and task 3. Among these task groups, task 1 gets the highest priority according to the RM scheduler because it has the shortest duration of the three tasks.

The utilization factor for this set of tasks can be calculated as follows:

$$U = 0.5/2 + 1.2/3 + 0.5/6 = 0.73.$$

The schedule associated with $U_{\mathrm{RM}}(3)$ is $n = 3 = 3\ (2^{1/3} - 1) = 0.75$.

$U \leq n(2^{1/n} - 1)$, this is the condition to be satisfied for optimal scheduling, for the given taskset $U < U_{\mathrm{RM}}$ condition is satisfied. Therefore, no task will miss the deadline.

The fact that it is not always feasible to use the CPU to its maximum potential is a significant drawback of fixed-priority scheduling. RM is the best fixed-priority system, but despite this, it has a worst-case time limit of:

$$Wn = n\left(2\left({}^{1/n}\right) - 1\right) \tag{2}$$

where n represents how many tasks are in a system.

Figure 1 shows the simulation of RM scheduler for three periodic tasks.

From the Gantt chart, it is observed that no task is missing its deadline as for the given example U is 0.73 which is less than U_{RM}, i.e., 0.75 for $n = 3$.

5 Earliest Deadline First (EDF) Scheduler

A dynamic priority scheduling method called earliest deadline first (EDF) or least time to go is used in real-time operating systems to prioritize tasks. When a scheduling event (such as a task finishing or a new job being released) takes place, the queue will be checked for the task that is closest to its deadline [14, 15]. When compared with

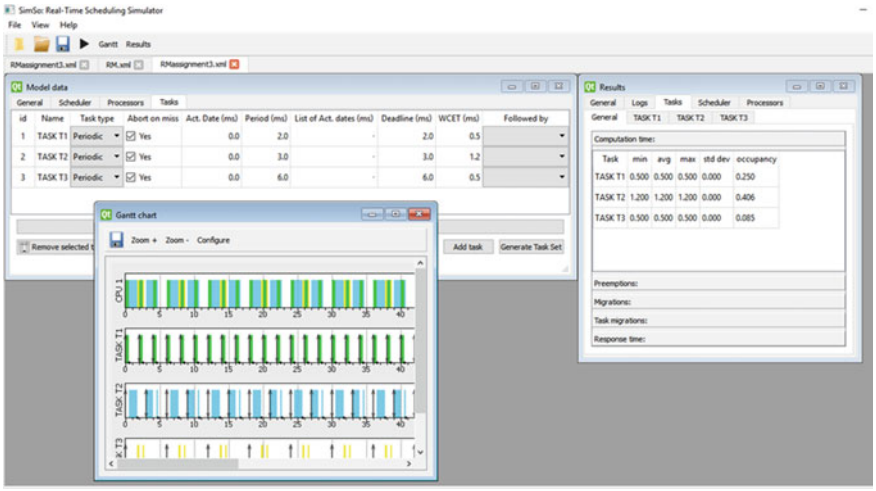


Fig. 1 Simulation of RM scheduler for three periodic taskset

other scheduling methods for real-time systems, EDF is efficient. While maintaining the job deadlines, it might increase CPU utilization to around 100%.

Schedulability Condition for EDF

Liu and Layland [7] proposed a necessary and sufficient utility constraint of the EDF algorithm in 1973.

Theorem (EDF) [7]: The least upper limit is 1 if a collection of τ of n tasks can be scheduled using the earliest deadline first method. Theorem proves that any collection of periodic, independent, preemptive tasks can be scheduled using EDF only if the total utilization of the task group is less than 1. Note that $Di = Ti$ in this scenario. As shown by Dertouzos [16] EDF is good for both periodic and non-periodic function groups.

The EDF scheduler has been implemented and four periodic tasks T1 (2, 0.5, 1.9), T2 (5, 2) have been executed.

T3 (1, 0.1, 0.5), T4 (10, 5, 20) for single and double cores.

Utilization factor for given four tasks:

$U = 0.5/2 + 2/5 + 0.1/1 + 5/10 = 1.25$ which is > 1 so, for this taskset EDF won't provide feasible scheduling if it schedules on single core.

As shown in Fig. 2 when it is simulated on a single core, task 4 missed its deadline when utilization factor crosses 1.

This situation can be avoided by using more than one core.

Figure 3 shows that all tasks meet their deadlines when scheduled on dual cores using the EDF scheduler.

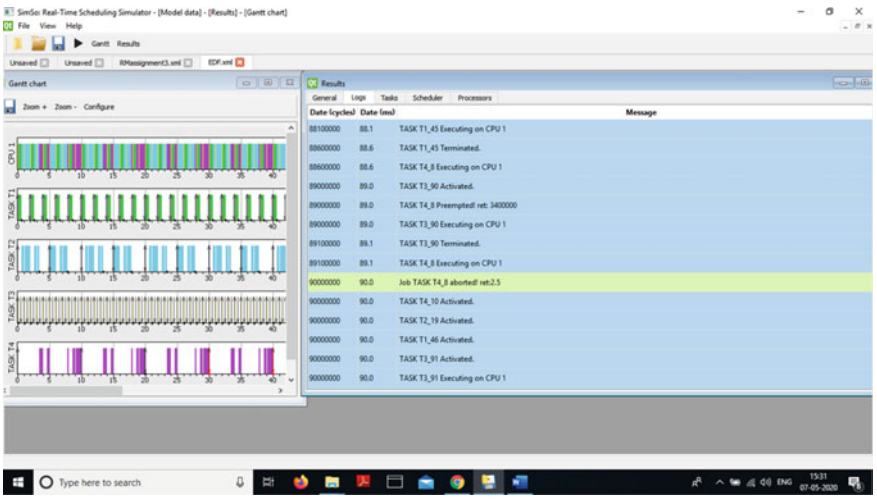


Fig. 2 Single CPU—four periodic tasks EDF

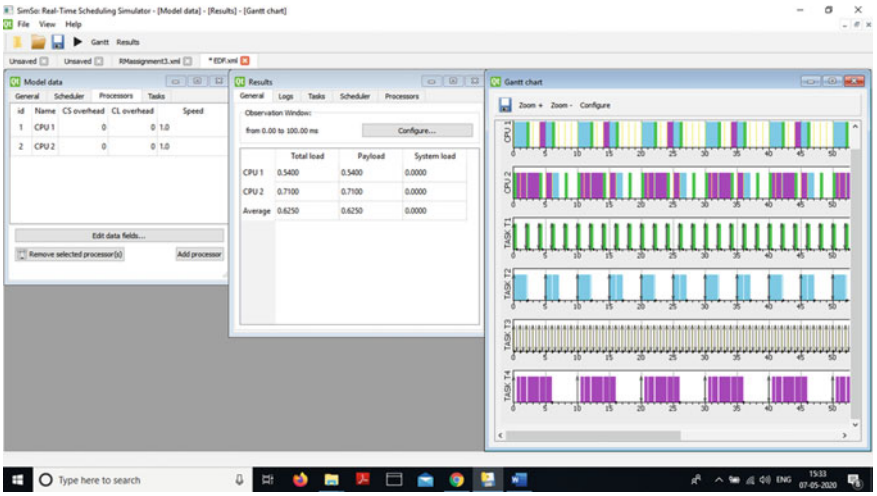


Fig. 3 EDF scheduler with two CPU four periodic tasks

So, it is observed that in the overload condition, i.e., when utilization factor is crossing 1, single core cannot schedule given taskset, but multicore architecture can be able to give optimal output.

6 Scheduling Analysis of Sporadic Task with EDF and RM Algorithms

As shown by Dertouzos [16] EDF is good for both groups of periodic and aperiodic activities. Sporadic tasks may make a request at any time, but two successive requests must be separated in time by at least p "time units."

To analyze scheduling algorithms, the synthetic data for the taskset has been generated with inputs as, T1 (3, 0.5), T2 (4, 1.5, 3), T3 (7, 1.0, 5). A sporadic job T4 has been generated with arrival time $t = 50$, having the execution time of 10 and a relative deadline of 30. The simulation is shown in Fig. 4 that if scheduling done with RM sporadic task T4 is missing its deadline, as RM gives feasible output for the periodic taskset, but for the unpredictable task RM scheduler is not optimal. Figure 4 shows that sporadic task misses deadline when using RM strategy.

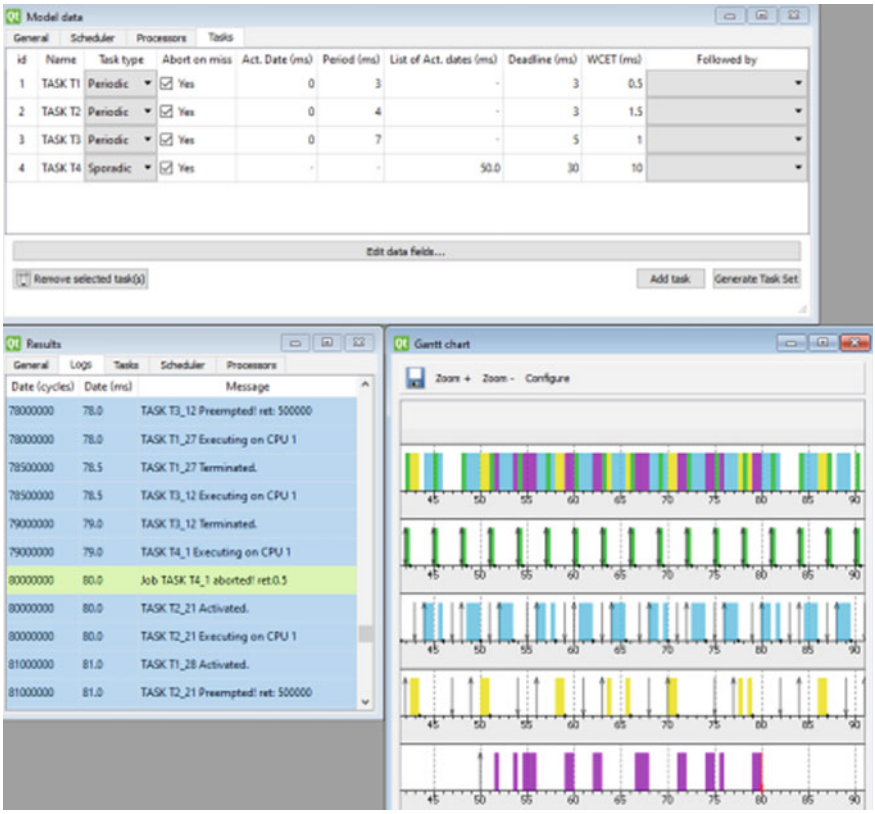


Fig. 4 Scheduling of sporadic task with RM strategy

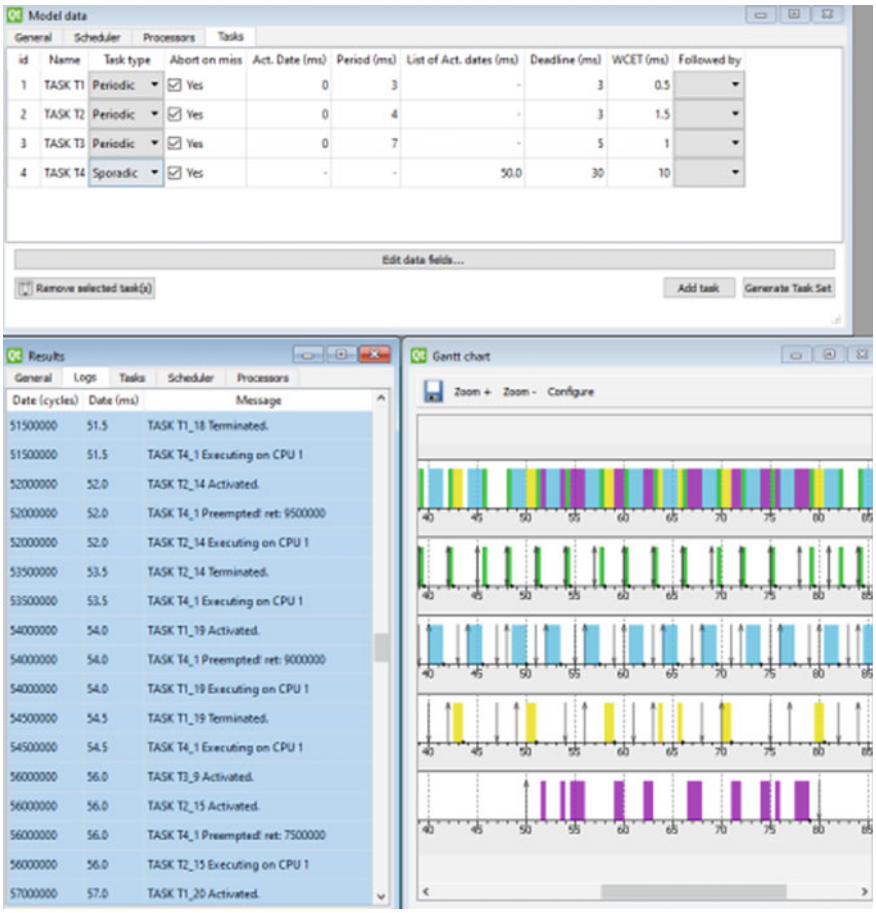


Fig. 5 Scheduling of sporadic task with EDF algorithm

Same taskset scheduled with EDF algorithm; simulation shown in Fig. 5—Gantt chart that if scheduling has been carried out with EDF strategy, sporadic task is not missing its deadline.

7 U_{RM} Schedulability Test-Based Scheduler

Implemented new scheduler. As per this scheduler, U_{RM} test has been carried out at run time. Considered condition to achieve high priority for the task. U_{RM} and U are calculated for each task. At runtime, a task whose utilization factor is lower than the task's U_{RM} gets higher priority.

For every task, which arrives in ready queue U_{RM} value has been calculated as per Eq. 3,

$$U_{\mathrm{RM}} = (\mathrm{taskNUM}[i] + 1.0) * ((\mathrm{pow}(2.0, 1/(\mathrm{taskNUM}[i] + 1.0))) - 1.0) \quad (3)$$

Utilization of the task has been calculated as per Eq. 4

$$U = (c[1] + (\mathrm{task.wcet/task.period})) \quad (4)$$

Condition has been set that, if present utilization factor of the task is less than U_{RM} value, then task will be scheduled on CPU.

With this strategy, 11 periodic taskset created and scheduled on three CPUs, observed every single assignment is completed on time.

Tasks 1–11 input: The following taskset will be scheduled using the modified algorithm on three processors: T1 (2, 1), T2 (2.5, 0.1), T3 (3, 1), T4 (4, 1), T5 (4.5, 0.1), T6 (5, 1), T7 (6, 1), T8 (7, 1), T9 (8, 1), T10 (8.5, 0.1), T11 (9, 1).

From Gantt chart as shown in Fig. 6 all the 11 tasks can schedule on three cores without missing deadline with proposed scheduler.

Fig. 6 Gantt chart for 11 tasks with 3 CPUs U_{RM} schedulability test-based algorithm

8 Conclusions

We have introduced real-time scheduling techniques for multiple parallel processors in this article and described the algorithm's schedulability requirements.

This paper's primary objective is to review popular multiprocessor scheduling techniques and evaluate their performance. Evaluated how well RM and EDF scheduling strategies performed.

In addition to introducing the circumstances in which some algorithms perform better than others, this performance assessment study was created to serve as a future guide for the creation of new algorithms. The efficacy and processing efficiency of the methods were assessed through extensive modeling tests.

The schedulability evaluations were assessed for the number of jobs in computer trials. In the future, we intend to evaluate and contrast the efficacy of additional online scheduling methods.

Acknowledgements The author is thankful to Mr. Ravikanth for his constant support in carrying out this work.

References

1. Rinku DR, Asha Rani M (2020) Evaluation of scheduling algorithms on Linux OS. In: ICETE 2019. LAIS, vol 4. Springer, pp 210–217. https://doi.org/10.1007/978-3-030-24318-0_25
2. Khan R, Kakhani G (2015) Analysis of priority scheduling algorithm on the basis of FCFS & SJF for similar priority jobs. Int J Comput Sci Mob Comput 4(9)
3. Rinku DR, Asha Rani M, Suhruth Krishna Y (2022) Exploring the scheduling techniques for the RTOS. In: ICT infrastructure and computing. Lecture notes in networks and systems, vol 520, pp 11–18. https://doi.org/10.1007/978-981-19-5331-6_2
4. Baker TP (2003) Multiprocessor EDF and deadline monotonic schedulability analysis. In: IEEE real-time systems symposium, Dec 2003
5. Oh Y, Son SH (1993) Preemptive scheduling of periodic tasks on multiprocessor: dynamic algorithms and their performance. Technical report CS-93-26. University of Virginia, CS Department, May 1993
6. Rinku DR, AshaRani M (2020) Reinforcement learning based multi core scheduling (RLBMCS) for real time systems. IJECE 10(2):1805–1813
7. Liu CL, Layland W (1973) Scheduling algorithms for multiprogramming in a hard real-time environment. J ACM 20(1):46–61
8. Baruah S (2005) The limited-preemption uniprocessor scheduling of sporadic task systems. In: Proceedings of euro micro conference on real-time systems, July 2005
9. Matarneh RJ (2009) Self-adjustment time quantum in round robin algorithm depending on burst time of the now running processes. Am J Appl Sci 6(10):1831–1837
10. Akhtar M, Hamid B, ur Rehman I, Humayun M, Hamayun M, Khurshid H (2015) An optimized shortest job first scheduling algorithm for CPU scheduling. J Appl Environ Biol Sci 5(12):42–46. PDCA12-70 data sheet. Opto Speed SA, Mezzovico, Switzerland
11. Kato S, Takeda A, Yamasaki N (2008) Global rate-monotonic scheduling with priority promotion. IPSJ Trans Adv Comput Syst (ACS) 2(1):64–74
12. Dong J, Zhang Y (2009) A modified rate monotonic algorithm for scheduling periodic tasks with different importance in embedded systems. In: International conference on electronic measurement & instruments, Sept 2009. IEEE Xplore, pp 4-606–4-609

13. Oh D-I, Baker TP (1998) Utilization bounds for N-processor rate monotonic with static processor assignment. Real Time Syst 15(2):183–193
14. Baruah S (2002) Robustness results concerning EDF scheduling upon uniform multiprocessor. In: Euromicro conference on real-time systems
15. Baker TP (2003) Multiprocessor EDF and deadline monotonic schedulability analysis. In: IEEE real-time systems symposium, Dec 2003
16. Dertouzos ML (1974) Control robotics: the procedural control of physical processes. In: Proceedings of IFFP congress, pp 807–813

Neural Transfer of Style Learning and Its Application

Harsh Mody, Harsh Parikh, Neeraj Patil, and Shanthi Therese

Abstract Neural transfer of style learning is the technique of gathering styling technique such as brush strokes, color palette, etc., in case of images and transferring it to another image creating a whole new image. Such a technique can be also used upon video and audio formats allowing users with very low artistic capabilities to create advanced styled content such as creating new paintings transferring style from famous painters, and similarly creating cool styled VFX backdrops without need of costly VFX artists or creating custom audio filters to sound like popular artists, etc. Neural transfer of learning technique has many applications which are yet to be explored by users. The primary problem with it is lack of awareness and good quality models, which generate excellent results to replace actual artists.

Keywords Neural style transfer · Machine learning · Image processing · Audio processing · Video processing · Convolutional neural networks · Applications of neural style transfer

1 Introduction

Neural art transfer was introduced and discovered first in the paper "A Neural Algorithm of Artistic Style" by Leon Gatys, and subsequently accepted by the peer reviewed CVPR conference in 2016. Neural art transfer is the transfer of the style from one image to another image, while keeping the other image intact. It is used to adopt the appearance, style, brush strokes, color palette of the style image and using those artistic ideas on the content image. This makes a combination of the semantic content of one image and the style/texture of another image. The principal problem of neural art transfer revolves around the way of computing the content loss and the style loss of the image separately using deep neural networks. There are a lot of a pretrained convolutional neural net that will compute the gram matrix of the image which will be used to compute the loss function and learn from it [1].

H. Mody (✉) · H. Parikh · N. Patil · S. Therese
Thadomal Shahani Engineering College, Mumbai 400050, India
e-mail: harsh.mody@icloud.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_24

2 Convolutional Neural Network

Convolutional neural net or CNN is an algorithm which takes an image as an input and extracts its prominent features. The preprocessing which is required by the ConvNet is very low in comparison with the other algorithms of classification. Convolution takes in input images through a set of convolutional filters, each of which triggers certain features from the images. While in the primitive techniques, the kernels are hand-engineered, with enough training, ConvNets can learn these kernels/characteristics [2].

The ConvNet reduces the pictures into a format that is easier and faster to process, without sacrificing important features for generating a reasonable prediction. This information is critical when we are developing an architecture that is good at learning features and is largely scaled to huge datasets. Rectified linear unit (ReLU) is used for quicker and more active training by mapping the values less than or equal to nothing and keeping the positive values unchanged. This is also called activation, because only the activated features are carried to the next step of the model. Pooling layers make things easier for the image by nonlinear down sampling, minimizing the number of features that the network would need to train for.

As the algorithm was developed for a two-dimensional (2-D) image with different channels, the operations are executed between an array of image data and 2-D array of weights which are known as kernel or filter. The kernel is generally smaller in size when compared with the input image and the type of multiplication operation applied between kernel-sized part of the input image and the kernel is a dot product. Dot product is element-wise multiplication between each of the kernel-sized part of the input image and kernel, which is then added and which is always resulting in a single value. As it results a single value obtained from dot product, the operation is frequently referred to as the "scalar product" [3].

The filter can be of any size but keeping it smaller than the input image is intuitional as it will make the same kernel that is the array of weights to be multiplied by the image many times at various parts of the image. The kernel is specifically applied to the overlapping patch or kernel-sized parts of the image, string from the left to the end and from the top portion to the bottom [4].

The image, we get from the multiplication of the kernel weights with the input image array once is a single numerical value. Since the kernel will be multiplied many times to different parts of the array, the output image is a two-dimensional vector of resultant values that depict the filtering of the image. This 2-D array from the convolutional net is also referred to as "feature map". After creating the feature map, we can also apply to each value in the feature map array a nonlinearity, such as tanh, similar as we do for the dense layers of ANN [5].

3 Transfer of Style for Images

Neural style transfer for images takes in two images such as: content image and style image. This approach facilitates in the creation of the content image in the pattern of the style image. It uses neural networks to transfer the aesthetic style from one image to another. Designing, content generating and the development of creative tools all benefit from neural style transfer. The process of merging style from one picture into another while keeping the content intact is known as neural style transfer. The only difference is the image's style which provides an aesthetic touch to your image. The content picture displays the layout or a rough drawing, whereas the style image depicts the painting or colors [6] (Fig. 1).

It is a computer vision application connected to image processing methods and deep convolutional neural networks. Like a traditional neural network, a ConvNet contains neurons having weights and biases. The algorithm learns these weights and biases during the training and learning process, and it constantly updates them with each new training data.

However, in the matter of ConvNet, all the hidden neurons in each convolutional layer share the weights and bias values. This means that the same features, such as a point, line or edge are being detected in all the hidden neurons, in various parts of the image. This tends to make the net quite tolerant to the capturing of different objects in the input image. For example, a network to detect face of a person will be able to detect them throughout the image multiple times.

3.1 Content Loss

Content loss is the loss calculated to compare the features of the generated image to the used content image. It is calculated in order to make sure that the output image from the neural net contains the same content as the input content image, while algorithm changes the style. In this way, the authenticity of the input content array isn't taken away from it and from the input style image, the style gradients get added.

Two values are needed for calculating the content loss function. The content features of the generated output image as well as the content image using the loss convolutional networks [7]. Then, we calculate the mean squared error or L2 norm

Fig. 1 Neural art transfer for images

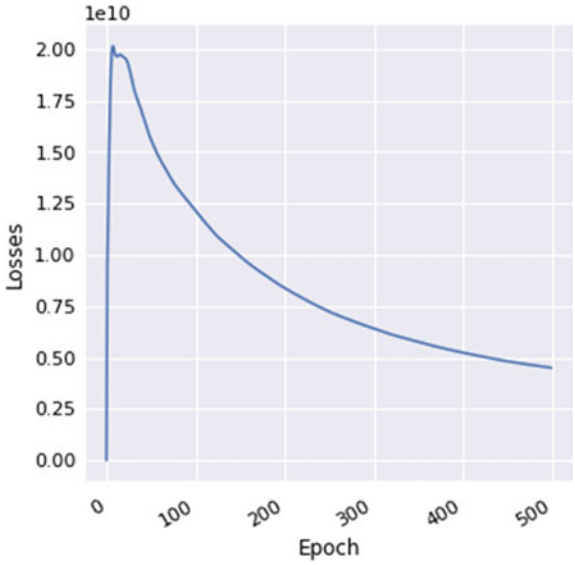


Fig. 2 Content loss graph

error. To perform the mean squared error, first we calculate the error of the generated output image features to the content image features. We calculated the element-wise subtraction [8]. Then we square the errors calculated to get the squared errors of the element-wise values then take the mean of the squared errors (Fig. 2).

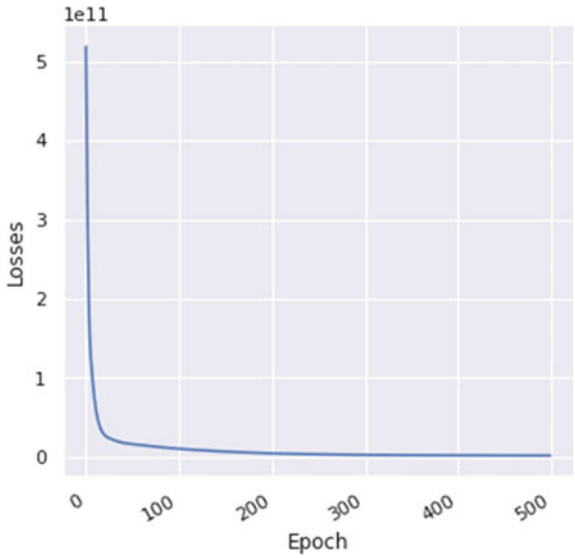
The graph shows the content loss in the generated image. As we see, the initial loss is zero which increases suddenly to a huge value and then keeps on decreasing as the model proceeds. This is because the sudden introduction of different style at the first epoch reduces the content to huge level but from there, we try to minimize both losses.

3.2 Style Loss

The style loss of generated image is generated image from all the layers in the model, while the content one is generated from the higher levels of the networks only. The deepest layers of the network are generally able to distinguish between the style image and the generated image, since we wouldn't want to lose the images value and its real meanings.

Now, for calculating the style loss, we create a gram matrix. First, a loss network is used to calculate the loss. Both the generated output image and the original style image is fed to the style loss network, generally extracted from a pretrained model. This produces their activations [9]. Next, these outputs from the loss network are averaged over every single value in the feature map to generate what is called a gram matrix. The gram matrix measures the covariance and hence it tends to get

Fig. 3 Style loss graph



the important information and features about the parts of the image which tend to activate similarly. The advantage of the gram matrix technique is that it is responsible for enabling the coexistence of different features in various regions of the image.

After the calculation of the gram matrix, the squared distances between the gram matrix generated by the output image and the gram matrix generated by the original style image is calculated. We use the Frobenius norm to measure distances (Fig. 3).

The graph plots the style loss in the generated image for the epochs. As we see, the initial loss is very high since the style is not present in the content image. But as the epochs proceed, the style loss minimizes to almost zero and we get an almost straight line toward the end.

3.3 Adam Optimizer

Adam optimizer is used in the model for optimizing the model and making it more accurate. The Adam optimizer is a combination of two different optimization techniques. The gradient descent with momentum is used in combination with the RMSP in Adam optimizer technique. The "exponentially weighted average" of the gradients and exponential moving average are taken into consideration to boost the gradient descent algorithm in this technique [10].

Adam optimizer has two more tunable hyperparameters beta one and beta two which can be varied to increase the model's accuracy score. The Adam optimizer algorithm provides much better performance than the gradient descent with

momentum and RMSP and outperforms them by great extent into providing an optimized gradient descent. The optimizer leverages the power of adaptive learning rates methods to search for the individual learning rates for each of the parameter.

4 Transfer of Style for Images

The music industry could profit from neural style transfer on audio. Nowadays, AI generated music is increasingly fashionable. These algorithms may be used by both beginners and by industry experts to develop new music. This means that anyone can record audio that has some unique musical style to it and can transfer those unique tones to another audio file and create something novel. We have seen how the concept of transfer of style is applied to images and how they transform. CNNs have been utilized for image neural style transfer. Images are represented as a set of pixels; an image's "style" could be defined by its colors and localized structures, resulting in high-quality style transfers. Previous attempts at applying the same notion to audio (i.e., audio style transfer) have failed [11].

Music can be represented in multiple ways which varies form sheet music to a spectrogram. It all depends on whether we rely on the score, sound, or control, and no end-to-end system can manage all levels of music representation simultaneously in a graceful manner. The problem of representing the music or lyrics is essential because deep learning must be applied to the fields of music and acoustics. The main motive of a successful model would also create attractive and engaging outcomes that doesn't look like it was superimposed [12].

We utilize spectrogram to solve this problem. In other words, we split our original waveform into overlapping chunks, extracting the magnitude of the frequency in each chunk (using Fourier Transformation), and the resultant vector becomes a peak on our final spectrogram. The spectrogram's x-axis indicates time, while the y-axis shows frequency. To render these spectrograms more effectively, we convert each "pixel" to a decibel scale by calculating the logarithmic of each value. Finally, we convert spectrograms to the mel scale using a mel filter bank, resulting in mel-spectrograms [13] (Fig. 4).

While previous network designs were exclusively made of stacked convolutional layers, newer architectures examine novel ways of generating convolutional layers to increase learning efficiency. These designs give various architectural measures that machine learning professionals can adjust to solve myriad of computer vision problems. These architectures may be used to extract most prominent features for image segmentation, object identification, image classification, and a range of other sophisticated applications.

One of the major applications in audio is the ability to create new and unique soundtrack by amalgamation of Lacrimosa by Wolfgang Amadeus Mozart and a violin soundtrack like Beethoven violin sonata into a model. Future scope for this algorithm is increasing the number of convolutional layers to extract more high-level features to facilitate better style transfer. Once the new spectrogram is generated the

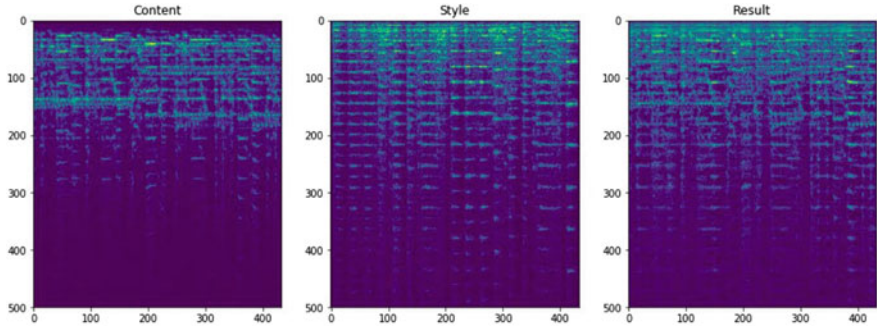


Fig. 4 Neural art transfer in audio

quality of the image is reduced and to make the output audio file as clear as possible multiple images enhancing filters are used to make sure that the output spectrogram is as clear as possible which in turn improves the quality of the sound [14] (Fig. 5).

As mentioned above in the transfer of style in image file, there is a way to get the optimum number of epochs the model needs to create a decent output. We plot the losses and find a place where the curve flattens for the content and style loss. The first is to minimize as much noise as possible in the output, and the second is to handle audio files. Furthermore, if the total number of epochs are increased, then there is a chance to improve the output audio quality, but this risks over-extraction and making the file worthless [15].

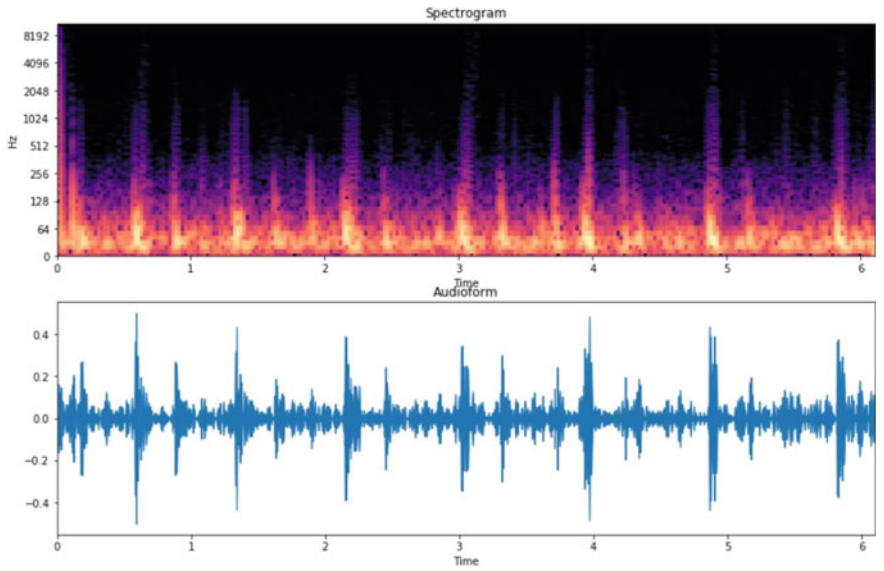


Fig. 5 New spectrogram generated

5 Transfer of Style for Video Files

A collection of images played at a very high frequency is nothing but a video. Each image is called a frame and can be joined with other frames to build a video. For the neural transfer of style for a video, the first step is applying transfer of style algorithm to use and load the pretrained model such as VGG-16, RESNET, etc. The shape is important because each model takes the input image with shape, e.g., VGG-16 needs an input with shape of $224 \times 224 \times 3$. The next step in the process is to define the content and style cost function. We configure the pooling layer by throwing some unwanted information by using average pool instead. Then we pick a deep layer out of all layers as output to define the model up to that layer [16].

The preprocessed content and styled image to compute the feature and activation map at the output layer corresponding to the model and model output corresponding to any random disturbance. We then also compute mean squared error loss and gradients for the image network which will help us update the input image in the reverse direction of the gradient and allow content loss to reduce so that produced image matches the input code [17] (Fig. 6).

The gram matrix computes at each layer, i.e., two images with same feature maps, we expect they will have same style but not same content. So, activation maps in early layers need to be programmed in such a way that they can capture finer textures and low-level features, whereas activation maps in deeper layers would capture with higher-level style elements in the image. For best results, hence, we ideally take a mix of shallow and deep layers and we compile multi-output model. Next, we combine style and content cost to obtain total loss of the network. We optimize our model for this quantity on each stage. Then, we solve the optimization problem and improve the loss minimization function. After that we run the optimizer function on the input content of a style image as per the defined symbolic computation graph. The network then does the job which it was designed for, i.e., minimize overall loss and generate an image which looks closer to content and style image.

In the video case, the network inference next is then applied to all the frames in video, e.g., for a 60 FPS video, the whole above video is done 60 times per second video length [18].

6 Conclusion

Apart from personal and artistic usage, neural style transfer has potential to transform so many great industries. It can be used by small motion picture houses to make great cool VFX looking scenes which would traditionally be only possible by big motion picture houses like Disney, etc., due to the sheer number of VFX artists they have at their disposal. Using their creativity and AI, they would be able to build as good scenes as them [19] (Fig. 7).

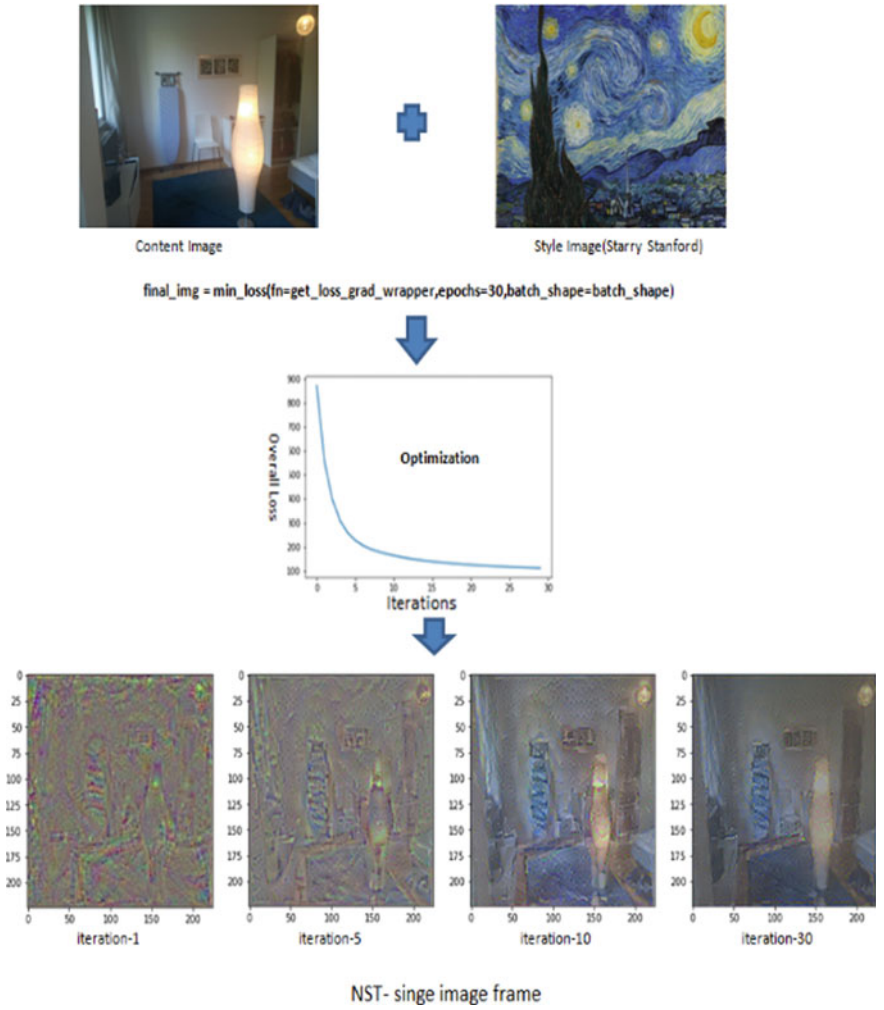


Fig. 6 Neural art transfer for video files

Fig. 7 Use in VFX

Fig. 8 Filter creator

Neural style transfer also has potential to transform e-commerce business by allowing users to customize their garment in a unique personalized style and texture. It can also be used in automobile industries to design new car wrap and paints, etc. This can help evolve fashion designs faster and quicker. Neural style transfer also has great potential to be a good social media filter creator tool allowing users to create unique customized filters to up their social media game (Fig. 8).

With all such promising applications, we are very sure neural transfer style learning is the future of customized e-commerce, social media filters and with further refinements will even replace VFX artists for small motion picture houses and help them build content easily [20].

References

1. Olu-Ipinlaye O (2022) Understanding & interpreting convolutional neural network architectures. In: Paperspace blog. https://blog.paperspace.com/understanding-interpreting-convolutional-neural-network-architectures/. Accessed 8 Nov 2022
2. Convolutional neural networks' mathematics. In: Towards data science. https://towardsdatascience.com/convolutional-neural-networks-mathematics-1beb3e6447c0. Accessed 8 Nov 2022
3. (PDF) Research on neural style transfer. https://www.researchgate.net/publication/356022464_Research_on_Neural_Style_Transfer. Accessed 8 Nov 2022
4. IBM Cloud Education. What are convolutional neural networks? In: IBM. https://www.ibm.com/cloud/learn/convolutional-neural-networks. Accessed 8 Nov 2022
5. A review of the automobile industries waste water treatment methodologies. https://www.irjet.net/archives/V6/i6/IRJET-V6I6260.pdf. Accessed 8 Nov 2022
6. Style transfer guide. In: Fritz. https://www.fritz.ai/style-transfer/. Accessed 8 Nov 2022
7. Johnson M (2022) What is content loss in neural style transfer? In: nnart. https://nnart.org/what-is-content-loss/. Accessed 8 Nov 2022
8. Dey S. Hands-on image processing with python. In: O'Reilly Online Learning. https://www.oreilly.com/library/view/hands-on-image-processing/97817893437 31/5ef8eed3-fdca-4ee0-9e50-292a6bc0ea10.xhtml. Accessed 8 Nov 2022
9. Neural style transfer: using deep learning to generate art. In: Neural style transfer: everything you need to know. https://www.v7labs.com/blog/neural-style-transfer. Accessed 8 Nov 2022
10. Vicky J (2021) Adam optimizer explained. In: Machine learning journey. https://machinelearningjourney.com/index.php/2021/01/09/adam-optimizer/. Accessed 8 Nov 2022
11. Ramachandran V (2019) CSCI 599 final blog post. In: Medium. https://medium.com/csci599-the-robinsons/csci-599-final-blog-post-7a773a144da9. Accessed 8 Nov 2022
12. Voice translation and audio style transfer with GANs. https://towardsdatascience.com/voice-translation-and-audio-style-transfer-with-gans-b63d58f61854. Accessed 8 Nov 2022

13. Getting to know the mel spectrogram. In: Towards data science. https://towardsdatascience.com/getting-to-know-the-mel-spectrogram-31bca3e2d9d0. Accessed 8 Nov 2022
14. CS230 deep learning. In: CS230. https://cs230.stanford.edu/. Accessed 8 Nov 2022
15. Fatima SM, Shehzad M, Murtuza SS, Raza SS (2020) Neural style transfer based voice mimicking for personalized audio stories. In: Proceedings of the 2nd international workshop on AI for smart TV content production, access and delivery. ACM conferences. https://doi.org/10.1145/3422839.3423063. Accessed 8 Nov 2022
16. Using CycleGAN to perform style transfer on a webcam. https://towardsdatascience.com/using-cyclegan-to-perform-style-transfer-on-a-webcam-244142effe7f. Accessed 8 Nov 2022
17. Neural style transfer on real time video (with full implementable code … https://towardsdatascience.com/neural-style-transfer-on-real-time-video-with-full-implementable-code-ac2dbc0e9822. Accessed 8 Nov 2022
18. CVF open access. https://openaccess.thecvf.com/. Accessed 8 Nov 2022
19. Neural style transfer. In: Engati. https://www.engati.com/glossary/neural-style-transfer. Accessed 8 Nov 2022
20. Neural style transfer | TensorFlow core. In: TensorFlow. https://www.tensorflow.org/tutorials/generative/style_transfer. Accessed 8 Nov 2022

Identity-Based Encryption (IBE)-Based Improved Data Possession Techniques

Parth Shah and Priteshkumar Prajapati

Abstract Asymmetric key cryptography is used by Provable Data Possession (PDP) schemes to confirm the accuracy of files that are stored remotely. The size and authenticity of the public key are the main drawbacks of asymmetric key cryptography based on the RSA algorithm. In this work, Identity-Based Encryption—Dynamic Provable Remote Data Possession (IBE-RDP) solutions have been proposed. IBE solves the issue as it uses the user's personal identity (e.g., phone number) as a public key. Hence, the problems related to the public key size and certification management are solved. The proposed solutions are efficient in terms of key generation and verification in spite of the more number of users. In addition, the proposed solutions solve the threats like (k, θ)—cheatability and server unforgeability. Two alternatives have been proposed IBE-RDP-I and IBE-RDP-II. IBE-RDP-I is efficient but provides a weaker security guarantee whereas IBE-RDP-II is less efficient with stronger security. The proposed solutions are implemented and proved very efficient. Computational Diffie–Hellman (CDH) hardness assumption is taken as a base for the security of the algorithm.

Keywords Cryptography · Security · Remote storage · Asymmetric key cryptography · Private Key Generator (PKG)

1 Introduction

Public Key Infrastructure (PKI) provides very promising security for communications systems. The use of paired keys strengthens it, which are mathematically related where one encrypts the plain text and another one decrypts the cipher text;

P. Shah · P. Prajapati (✉)
Chandubhai S. Patel Institute of Technology (CSPIT), Charotar University of Science and Technology (CHARUSAT), Changa, India
e-mail: pritesh.pnp.007@gmail.com

P. Shah
e-mail: parthshah.ce@charusat.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_25

one key is public, and another one remains private to the creator of a key. Deployment of a public key cryptosystem for the large-scale system usually employs the certifying authority (CA) service, which is authorized for producing public key certificates. Such a certificate for a user consists of the user's public key and identification information, along with authentication information of CA in the form of a signature. Anyone in possession of the key can verify the signature contained within the certificate using an authenticated copy of the public key of the CA.

Though a PKI cryptosystem is very popular, still it has a few disadvantages. The receiver is supposed to create private keys and public keys, the sender requires finding out and knowing the receiver's public key. Furthermore, the sender requires querying a CA to get the target receiver's public key. Though this may seem trivial in very small premises, however, the senders who are outside the organization not having access to the central repository may face a delay in response while getting a public key. Hence, it may be time-consuming for the sender to send an encrypted message. Moreover, the process works only if the recipient has already decided to use it and has made the key available.

Additionally, the sender may not know how to get the receiver's certificate. Also, the sender should have the assurance that the receiver's public key is valid, i.e., the receiver's certificate has not been withdrawn by the CA on account of the receiver has left the place of employment, or because of the compromise of the receiver's private key. An enterprise key management scheme requires doing numerous jobs. In the enterprise environment to protect documents, databases, or emails, a key management scheme must focus the requirements like the Distribution of encryption keys, Key management in a group, Distribute keys to trusted infrastructure, Certification Problems, and Direct Compromise.

Identity-Based Encryption (IBE) is considered an encouraging approach to overcome the above-mentioned problems. It uses one of the user's publicly available details (e.g., email address, date of birth, birth place, pin code/zip code, mobile number, etc.) to use it as a public key, which can be used for encryption. A Private Key Generator (PKG) act as a trusted third party uses an algorithm to generate a private key based on a public key. Through this public key, the receiver of the data can calculate its private key from PKG as and when needed. In addition, it does not require worrying about distributing a public key. Compare to a traditional PKI, IBE reduces the complexity of key management for the administrators and users. Additionally, specialized software is not required to manage the public keys or to read the communication. Shamir [19] proposed the first IBE-based algorithm which uses arbitrary string as a public key. In this algorithm, there are four phases:

- Setup: This phase generates a master key and global system parameters, which are used to generate pair of private and public key.
- Extract: To generate the corresponding public key string of the receiver, the master key is used by the sender.
- Encrypt: This phase does encryption using the public key.
- Decrypt: This phase does decryption using the corresponding private key, which is generated during the extract phase to decrypt messages.

Shamir's original motivation for IBE came from the complex certificate management of public key infrastructure, which is used in the e-mail system. For example, if the sender wants to send mail to someone that he may use the receiver's public key string "receiver@company.com" to encrypt the message, hence, no need to obtain the public key certificate of the receiver. Once the encrypted data is received by the receiver, he may require to contact PKG to generate the corresponding private key. As a part of the process, PKG authenticates the receiver to obtain his private key. This private key is used by the receiver to decrypt the message so he can read the message.

These technical inventions transformed into noticeable advantages by looking at the six requirements of the public key cryptosystem. As IBE architecture meets the requirements of an efficient key management scheme, hence it becomes the acceptable choice for the management of cryptographic keys. Several schemes based on IBE have been proposed [5, 11, 12, 20, 21] since the problem was raised by Shamir. However, there are inherited problems like colluding users, the time required to generate parameters from a corresponding public key, tamper-resistant hardware required for some solutions, etc. The solution provided in [4] is the first practical solution in the field of cryptography. This made the construction of a usable IBE system, which was an open problem. Interestingly, [19] has introduced the related notions of authentication and identity-based signature, having satisfactory solutions in [7, 8].

2 Motivation

To achieve efficiency and scalability in terms of public verifiability, PDP protocol [9, 18] needs public key management and distribution through Public Key Infrastructure (PKI). It incurs substantial overheads at the time of remote data verification, as the verifier requires validating the public key certificates. As per the cloud security threats discussed by [14]; it is very difficult to manage the certificate generation, issuing, renewal, revocation, and storage if there are a million numbers of users. So computation and storage are the major issues for the verifiers. This complicated certificate management can be eliminated using Identity-based public key cryptography [19] with better efficiency. Thus, it will be very meaningful to provide a solution-using IBE.

3 Modeling IBE-RDP

In this section security definition and the IBE system model are given. An IBE-RDP protocol consists of the following three agents as follows: PKG, Client, and Remote Storage Service similar to [22]. Table 1 presents a list of the notations used in this paper.

Table 1 Notation and description

Notation	Description
G_1	Cyclic multiplicative group with order q
Z_q^*	$1, 2, \ldots, q-1$
g_1	Generator G_1
g_2	Generator G_2
H	Cryptographic hash function
n	The block number
s	The sector number
$C_F = F_1, \ldots, F_n$	n blocks
$C_{F_i} = F_{i1}, \ldots, F_{is}$	s sectors per block
(F, T_i)	Tag generated at client side
(F', T_i')	Proof generated by the server

3.1 Proposed Solution (IBE-RDP-I)

In this section, solutions related to IBE-RDP have been devised. We have formulated our solution in two way. One of them provides weaker security which is efficient (IBE-RDP-I) whereas another one provides stronger security that is less efficient (IBE-RDP-II). We have modified the previous work of [16] and proposed solution (IBE-RDP-I) and modified the previous work of [3] and proposed solution (IBE-RDP-II) to provide an efficient solution compared to [1, 6].

1. **Setup** = $(1^s) \rightarrow$ {master-key}
   On input 1^s, select two primes p and q of size k such that $q \mid (p-1)$.
   Select generator $g \in Z_p$ of order q, where g is a subgroup of Z_p^*.
   Select a cryptography hash function: $H : \{0, 1\} \times G \times G \rightarrow Z_q^*$.
   Select $s \in Z_q$ randomly and compute $v = g^{-s} \bmod p$.
   The system parameters params $= (p, q, g, v, H)$ and the master-key is s which is known to the PKG only.

2. **Key Generation** (params) $\rightarrow \{sk_{\text{int}}, pk_{\text{int}}\}$
   Given a public identity ID, choose randomly $t \in Z_q$ and calculate
   $X = g^t \bmod p, \alpha = H(\text{ID}, X, v), Y \equiv t + s\alpha \bmod q$,
   where ID is public identity (α, Y) is private key.

3. **Tag Generation** $(sk_{\text{int}}, pk_{\text{int}}, F_{\text{id}}) \rightarrow T_i$
   Modified to our previous work [17] for each block of data $C_{F_{ij}}$, where $1 \leq i \leq n$, the client selects r_j uniformly at random from Z_q^* and computes:

$$T_i = \left(\left(\prod_{j=1}^{m} r_j^{C_{F_{ij}}} \right) h(i) \right)^{Y} \bmod q$$

Output T_i and send (F, T_i) to the server.

4. **Audit Integrity** $(sk_{\text{int}}, pk_{\text{int}}, \text{chal}) \rightarrow$ resp. Modified to our previous work [17] and Compute

$$T' = \prod_{j=1}^{c} T_{\alpha_j}^{\beta_j}, \quad F' = \sum_{j=1}^{c} C_{F_{ij}} * \beta_j$$

5. **Check Proof** $(pk_{\text{int}}, \text{chal}, \text{resp}) \rightarrow$ {"success", "failure"}
The auditor verifies based on the received resp.

$$e(T, g) = e\left(\prod_{i=1}^{c} h(\alpha_i * (i))^{\beta_i} \prod_{j=1}^{m} r_j^{F'}, R \cdot Y^{H(\text{ID}, R)} \right)$$

6. **Correctness Analysis**

$$\begin{aligned}
e(T', g) &= e\left(\prod_{j=1}^{c} T_{\alpha_j}^{\beta_j}, g \right) \\
&= e\left(\prod_{j=1}^{m} \left(r_j^{C_{F_{ij}}} h(i) \right)^{Y\beta_j}, g \right) \\
&= e\left(\prod_{j=1}^{m} \left(r_j^{C_{F_{ij}}} h(i) \right)^{\beta_j}, g^{Y} \right) \\
&= e\left(\left(\prod_{j=1}^{m} h\left(\alpha_i * (i)\right)^{\beta_j} \right) r_j^{\sum_{j=1}^{m} C_{F_{ij}} \beta_j}, g^{Y} \right) \\
&= e\left(\prod_{i=1}^{c} h\left(\alpha_i * (i)\right)^{\beta_i} \prod_{j=1}^{m} r_j^{F'}, R \cdot Y^{H(\text{ID}, R)} \right)
\end{aligned}$$

3.2 Proposed Solution (IBE-RDP-II)

1. **Setup** $(1^S) \rightarrow$ {master-key}
   Initial parameters are same as IBE-RDP-I.
   Choose $x_i \in Z_q$ randomly and compute $X_i \equiv g^{x_i} \bmod p$ for $1 \leq i \leq n$. The system parameters params $= (p, q, g, X_i, H)$ and the master-key is (x_i) which is known to the PKG only.

2. **Key Generation** (params) $\rightarrow \{sk_{\text{int}}, pk_{\text{int}}\}$
   Given a public identity ID, choose $r_i \in Z_q$ randomly and compute

$$R_i \equiv g^{r_i} \bmod p$$

$$s_i \equiv -r_i - H(R_i \parallel \text{ID}) \cdot x_i \bmod q \quad \text{for } 1 \leq i \leq n.$$

   The user private key is (R_i, s_i).

3. **Tag Generation** $(sk_{\text{int}}, pk_{\text{int}}, F_{\text{id}}) \rightarrow$ Tag
   Let us denote each symbol of C_{F_i} for $(1 \leq i \leq n)$. For each block of data C_{F_i}, where $1 \leq i \leq n$, compute:

$$z_i = g^{s_i} \bmod p$$

$$y_i = s_i + C_{F_i} \cdot x_i \bmod q$$

   Output Tag = (z_i, y_i) and send (C_F, Tag) to the server.

4. **Audit Integrity** $(sk_{\text{int}}, pk_{\text{int}}, \text{Tag}, \text{chal}) \rightarrow$ resp.
   Modified to our previous work [17] and compute

$$Y = \sum_{i \in I} \beta_i y_i \bmod q$$

5. **Check Proof** $(pk_{\text{int}}, \text{chal}, \text{resp}) \rightarrow$ {"success", "failure"}
   The auditor verifies based on the received resp.

$$\delta_1 = \prod_{i \in I} z_i^{\beta_i} \bmod p$$

   and

$$\delta_2 = g^{-Y} \prod_{i=1}^{n} X_i^{\mu_i} \bmod p$$

if $e(\delta_1, g) = e(\delta_2, g)$ then success otherwise fail.

6. **Correctness Analysis**

$$\begin{aligned} e(\delta_2, g) &= e\left(g^{-Y} \prod_{i=1}^{n} X_i^{\mu_i}, g\right) \\ &= e\left(g^{-\sum_{i\in I} \beta_i y_i} \prod_{i=1}^{n} g^{x_i \mu_i}, g\right) \\ &= e\left(g^{-\sum_{i\in I} \beta_i (s_i + C_{F_i} \cdot x_i)} g^{x_i \sum_{i\in I} \beta_i C_{F_i}}, g\right) \\ &= e\left(g^{-\sum_{i\in I} \beta_i s_i - \sum_{i\in I} \beta_i C_{F_i} \cdot x_i} g^{x_i \sum_{i\in I} \beta_i C_{F_i}}, g\right) \\ &= e\left(g^{-\sum_{i\in I} \beta_i s_i}, g\right) \\ &= e\left(\prod_{i\in I} z_i^{\beta i}, g\right) \end{aligned}$$

4 Discussion on Some Design Decisions

We have adopted some of the design decisions based on [23]. The decisions have been taken to satisfy security definition and performance design requirements. First, our tag generation and proof generation protocols are new. In the proposed scheme the client requires to perform $O(mn)$ multiplication operations plus $O(n)$ exponentiation operations, where n and m are blocks and the number of symbols in each block respectively of data blocks. In contrast, the earlier solutions perform $O(mn)$ exponentiation and multiplication operations. In the IBE-RDP-II, for verification $z_i = g_1^{-s_{1i}} g_2^{-s_{2i}} \bmod p$ used, which is the very important signature scheme [23]. On the other hand, $z_i = g_1^{-s_{1i}} \bmod p$ is used in IBE-RDP-I which makes it secure with the cost of efficiency. This is to imply the CDH problem on which the security of our algorithm depends.

Table 2 Comparison of the computational complexity of IBE schemes

Scheme	Key generation	Tag generation	Proof generation and verification
Schnorr	$2T_e + T_m + T_h + T_a$	$(n+m)T_e + nT_m + T_h$	$2T_p + (2c+m)T_e + (2c+m)T_m + cT_h$
IBE-RDP I	$2T_e + T_m + T_h + T_a$	$nT_e + nT_m + T_a$	$2T_p + (c+n+1)T_e + (2c+n+1)T_m$
IBE-RDP II	$4T_e + 4T_m + 2T_h + 2T_a$	$2nT_e + 2nT_m + 2T_a$	$2T_p + (c+n+2)T_e + (4c+n+2)T_m$

5 Implementation and Result Analysis

The scheme of Schorr proposed in [22] has been implemented for the basis of comparison. This is the only scheme proposed in the literature which provides an implementation of PDP using IBE. We have conducted 20 trials for the experimental analysis due to variations in the results among different trials.

The use of tate pairing outperforms over the benchmarks. The computation complexity of setting the parameters for key generation of the above-mentioned schemes are $2T_e + T_m + T_h + T_a$, $2T_e + T_m + T_h + T_a$, and $4T_e + 4T_m + 2T_h + 2T_a$, for Schnorr, IBE-RDP-I and IBE-RDP II respectively. Here T_e = exponentiation time, T_h = hash time, T_m = multiplication time, T_a = addition time. A comparison of the computational complexity of IBE schemes is given in Table 2.

In the tag generation phase, the client needs to do $(n+m)T_e + nT_m + T_h$, $nT_e + nT_m + T_a$, $2nT_e + 2nT_m + 2T_a$ operations in Schnorr scheme, IBE-RDP-I and II, respectively. In the proof generation and proof verification phases, the client and server needs to perform $2T_p + (2c+m)T_e + (2c+m)T_m + cT_h$, $2T_p + (c+n+1)T_e + (2c+n+1)T_m$, $2T_p + (c+n+2)T_e + (4c+n+2)T_m$ for Schnorr scheme, IBE-RDP-I and II, respectively. Our proposed IBE-RDP-I and II protocols are more efficient than the RSA-based RDP schemes [1, 2, 6, 10].

Figure 1 gives the idea for the time required to generate tags at the client side. We have taken sector size $m = 4$. Figure 2 shows proof generation computation cost. Figure 3 shows proof verification computation cost.

6 Conclusion

In this work, two solutions have been provided to solve the issues encountered in the IBE scheme with public verification support and data privacy. The solutions are proposed work based on the concept of bilinear pairing. These solutions provide protection against replay and replace attacks as well as attacks against data privacy by embedding indexes. It is proved that the proposed solutions are secure under a random oracle. Experiment analysis shows that both of the solutions do not incur

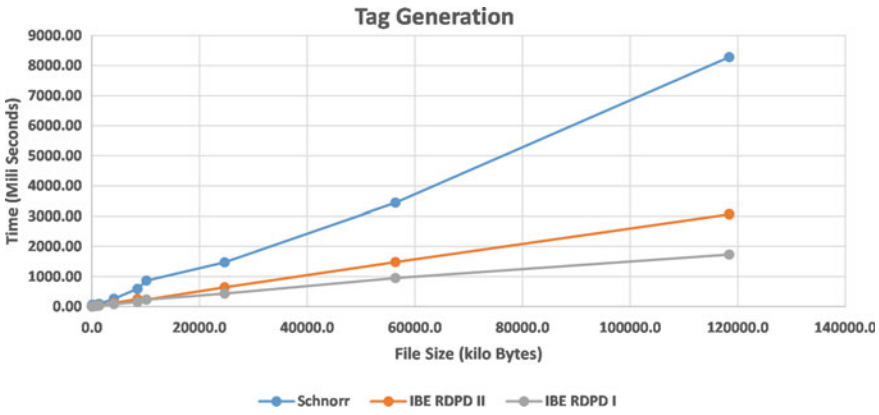


Fig. 1 Time required to generate tags (IBE-RDP)

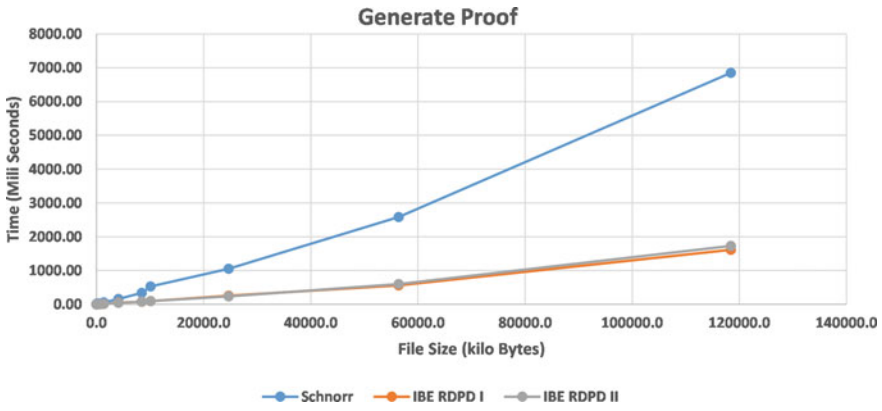


Fig. 2 Time required to generate proof (IBE-RDP)

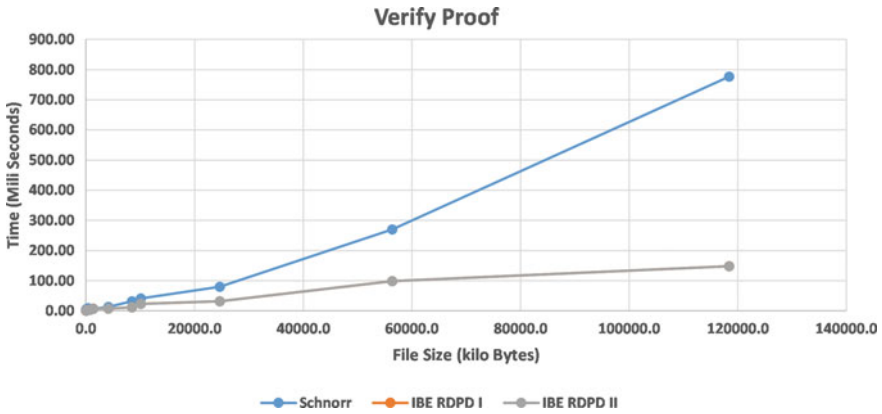


Fig. 3 Time required to verify proof (IBE-RDP)

more overhead in terms of execution time. It has been suggested to use the secure IBE RDP protocol, which works well if the CDH problem is considered to be challenging. The IBE RDP protocol improves previous protocols in transmission and computation, along with the structural benefit of removal of certificate management and verification. Experiments reveal that the provided probabilistic methods ensure sampling a file at the server's storage, making it practically efficient to confirm ownership of enormous data sets. Our tests demonstrate that such approaches likewise place a heavy I/O and computational burden on the server.

Our proposed solutions can be used in our previous work [9, 13, 15, 18] to reduce computational overhead as future work. Our work can be applied to safeguard sensitive data's confidentiality, privacy, and integrity, including student records, financial information, classified government data, medical records, and other Personal Identifiable Information (PII) and financial data. In the healthcare industry, it can protect sensitive patient information while being stored and transferred. Financial institutions can use these techniques to secure bank account details, credit card information, and transaction records to prevent identity theft and financial fraud. Government agencies can secure national security data and PII with IBE-based techniques, while cloud storage providers can ensure that the data they store is protected against unauthorized access or data tampering. Similarly, educational institutions can secure student records, financial aid information, and other sensitive data with these techniques.

References

1. Ateniese G, Burns R, Curtmola R, Herring J, Kissner L, Peterson Z, Song D (2007) Provable data possession at untrusted stores. In: Proceedings of the 14th ACM conference on computer and communications security, pp 598–609
2. Ateniese G, Di Pietro R, Mancini LV, Tsudik G (2008) Scalable and efficient provable data possession. In: Proceedings of the 4th international conference on security and privacy in communication networks, pp 1–10
3. Bellare M, Namprempre C, Neven G (2009) Security proofs for identity-based identification and signature schemes. J Cryptol 22(1):1–61
4. Boneh D, Franklin M (2001) Identity-based encryption from the Weil pairing. In: Annual international cryptology conference. Springer, pp 213–229
5. Desmedt Y, Quisquater JJ (1986) Public-key systems based on the difficulty of tampering (is there a difference between DES and RSA?). In: Conference on the theory and application of cryptographic techniques. Springer, pp 111–117
6. Erway CC, Küpçü A, Papamanthou C, Tamassia R (2015) Dynamic provable data possession. ACM Trans Inf Syst Secur (TISSEC) 17(4):1–29
7. Feige U, Fiat A, Shamir A (1988) Zero-knowledge proofs of identity. J Cryptol 1(2):77–94
8. Fiat A, Shamir A (1986) How to prove yourself: practical solutions to identification and signature problems. In: Conference on the theory and application of cryptographic techniques. Springer, pp 186–194
9. Kadvani S, Patel A, Tilala M, Prajapati P, Shah P (2019) Provable data possession using identity-based encryption. In: Information and communication technology for intelligent systems. Springer, pp 87–94

10. Lim HW (2006) On the application of identity-based cryptography in grid security. Ph.D. thesis, University of London
11. Maurer UM, Yacobi Y (1991) Non-interactive public-key cryptography. In: Workshop on the theory and application of cryptographic techniques. Springer, pp 498–507
12. Menezes AJ, Katz J, Van Oorschot PC, Vanstone SA (1996) Handbook of applied cryptography. CRC Press
13. Prajapati P, Shah P (2014) Efficient cross user data deduplication in remote data storage. In: International conference for convergence for technology-2014. IEEE, pp 1–5
14. Prajapati P, Shah P (2022) A review on secure data deduplication: cloud storage security issue. J King Saud Univ Comput Inf Sci 34(7):3996–4007
15. Prajapati P, Shah P, Ganatra A, Patel S (2017) Efficient cross user client side data deduplication in Hadoop. JCP 12(4):362–370
16. Schnorr CP (1996) Security of 2^t-root identification and signatures. In: Annual international cryptology conference. Springer, pp 143–156
17. Shah P, Ganatra A (2017) Reliable and efficient distribution of multicast session key for deduplicated data in cloud computing. Int J Intell Eng Syst 10(1):48–55
18. Shah P, Prajapati P (2022) Provable data possession using additive homomorphic encryption. J King Saud Univ Comput Inf Sci 34(6, Part B):3448–3453
19. Shamir A (1984) Identity-based cryptosystems and signature schemes. In: Workshop on the theory and application of cryptographic techniques. Springer, pp 47–53
20. Tanaka H (1987) A realization scheme for the identity-based cryptosystem. In: Conference on the theory and application of cryptographic techniques. Springer, pp 340–349
21. Tsujii S, Itoh T (1989) An ID-based cryptosystem based on the discrete logarithm problem. IEEE J Sel Areas Commun 7(4):467–473
22. Wang H, Wu Q, Qin B, Domingo-Ferrer J (2014) Identity-based remote data possession checking in public clouds. IET Inf Secur 8(2):114–121
23. Zheng Q, Xu S (2012) Secure and efficient proof of storage with deduplication. In: Proceedings of the second ACM conference on data and application security and privacy, pp 1–12

Prediction of Stroke from Hand Patterns Obtained Using Spiral/Wave Test Through Densely Connected Convolutional Neural Networks

R. Punitha Lakshmi and V. Vijayalakshmi

Abstract Patients affected from brain stroke often suffer from difficulties in speech or weakness in one half of the body after an acute symptom. Blood vessel blockage causes death of cell. It completely stops the supply of oxygen to all the parts of the body causing patient death. The two most common methods used to detect brain stroke are "Computer Tomography (CT)" and "Magnetic Resonance (MR)." A neurosurgeon with a great deal of experience will be able to apply these methods effectively only in the advanced stages of the problem. The aforementioned imaging techniques cannot provide a clear picture of how the stroke is affecting the brain in any specific area. Having the precise stage of the stroke will help the surgeon prepare for surgery and begin recovery. The region of the affected part decides the severity of the disease. Main objective of the research work is to propose a prediction model by combining hand gestures and bio signals obtained by the patients instead of EMG and EEG obtained from Electronic Health Records. Proposed hybrid model is to be implemented using Deep learning algorithm through which prediction rate and the accuracy can be enhanced. It is used for early detection of stroke and diagnosis of stroke lesion. The proposed technique is used to improve the segmentation of MRI images. Automatic classification with improved techniques can be used to find the severity of stroke without human intervention.

Keywords Stroke · Ischemic stroke · Deep learning models · Convolutional neural networks

1 Introduction—Stroke

The number of strokes reported in India stands at 0.93 million, of which the majority occur in people aged 45 + [1]. Having determined that the problem is vascular, the clinician must next attempt to distinguish whether it is hemorrhage or

R. P. Lakshmi (✉) · V. Vijayalakshmi
Department of ECE, Puducherry Technological University, Puducherry, India
e-mail: punitha278@pec.edu

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_26

ischemia. Overall, ischemic strokes is about 80%–85%. Intracerebral hemorrhage, subarachnoid hemorrhage account for about 10% and 5% of all strokes, respectively [2].

The presence of symptoms with headache or neck stiffness is suggestive of a hemorrhagic process, as is early loss of consciousness in a patient with a suspected supratentorial lesion. Ischemia is more likely when symptoms are consistent with neurologic dysfunction in a single arterial territory or when improvement is rapid or occurs early in the clinical course. Although the distinction between hemorrhage and ischemia is rarely difficult clinically, there are exceptions, and sometimes both occur simultaneously (e.g., hemorrhagic infarction). Computed tomography has revolutionized the physician's ability to distinguish between hemorrhage and infarction in emergencies, virtually eliminating all cases in which uncertainty exists [3]. Having clarified the localization of the process, one should next consider the underlying mechanism.

A. *Categories of Stroke*

The major categories of cerebral ischemic events that relate to the underlying pathophysiology. Prehospital stroke scale explains about the facial drop, arm drift, abnormal speech. The doctor checks all the above features once the patient is received in the hospital.

Figure 1 explains the symptoms in detail. Another useful method of classifying the subtypes of ischemic stroke is to use the categories of thrombotic infarction, in which a locally reduced blood supply is caused by a blockage formed in situ in an artery.

Embolic infarction, which refers to a blockage caused by a piece of material that has become detached from a more proximal site; and lacunar infarction, which is caused by thrombosis of one of the small penetrating branch arteries.

B. *Historical Evaluation of Key Signs and Symptoms*

A detailed history is the most important part of the evaluation of a patient with cerebrovascular disease.

Prehospital stroke Scale

Facial Droop (have patient show teeth or smile):
• Normal—both sides of face move equally
• Abnormal—one side of face does not move as well as the other side
Arm Drift (patient closes eyes and holds both arms straight out for 10 seconds):
• Normal—both arms move the same or both arms do not move at all (other findings, such as pronator drift, may be helpful)
• Abnormal—one arm does not move or one arm drifts down compared with the other
Abnormal Speech
• Normal—patient uses correct words with no slurring
• Abnormal—patient slurs words, uses the wrong words, or is unable to speak
Interpretation: If any 1 of these 3 signs is abnormal, the probability of a stroke is 72%.

Fig. 1 Prehospital stroke scale

Table 1 Modified rank in scale (mRS)

Grades	Symptoms
Grade I	The patient has no significant disability; able to carry out all pre-stroke activities
Grade II	The patient unable to carry out all pre-stroke activities but able to look after self without daily help
Grade III	The patient has moderate disability; able to walk without the assistance of another individual
Grade IV	The patient unable to walk or attend to bodily functions without assistance
Grade V	The patient has severe disability; bedridden, incontinent, requires continuous care
Grade VI	The patient has expired

Initial attention should be directed toward identifying and characterizing:

- The time of onset and possible precipitating events.
- The features of the circumstance of onset, including the patient's activities, the temporal profile of the onset of symptoms.
- The rapidity with which maximal deficit developed (a typical vascular profile).
- The presence of focal or generalized neurologic deficit at the onset, or alterations of level of consciousness at onset.
- The presence of headache, vomiting, or seizure activity (focal or generalized).
- The chronologic course of neurologic symptoms after onset.

Frequently, the patient may not remember the precise details of the early temporal course and other important historical details. In this case, family members are often the best source of the information. Modified Rankin Scale (mRS) is the measure of global disability outcome measure in stroke clinical trials is given in Table 1.

The patient should be asked what is specifically meant by certain words that are used to describe symptoms (for "dizziness," "headache," and "poor vision") because these terms have a wide range of different meanings with different implications for diagnosis and management [4]. This study aims to find an engineering solution that will make it easier for neurologists to assess patients.

2 Literature Survey

Various image-processing methods have been proposed in medical informatics for different requirements. Machine learning emerged from artificial intelligence is considered as an important for identification of disease [5]. The predictive module uses hand gestures collected from spiral/wave test. However, its accuracy and corresponding performance criteria still need to be improved. Other models are more powerful in classifying medical pictures. Disadvantage of these networks are face convergence problems, overfitting problem, and the vanishing gradient problem [6]. Electromyography was acquired in real time, after which key parameters are taken

for model development. Prediction is performed using random forest classification, which yields 90% accuracy, which needs to be improved. Therefore, based on DenseNet, CNN provides options for accurate classification of medical images as it allows features. The developer [7] proposes moving of hands was detected and combined with motion detection algorithm using active appearance model. The limitation of the work, stroke is detected by the sudden change on the wrinkle line.

More number of parameters has to be included for attaining better accuracy. Tuladhar et al. discuss about the automatic segmentation process. CNN is proposed for segmentation of NCCT lesions and to quantify ischemic stroke. The algorithm uses Bayesian optimization for prediction [8].

High segmentation accuracy for large lesions was obtained by using dataset with NCCT Images [9]. However, it has been used in multiple area to get high accuracy. Due to its excellent problem sensitive model fine tuning process it generates a set of particular models. It shows that DenseNet could be used for identifying stroke disease from CT and MRI images. The CNN model can produce better segmentation. Prediction of stroke using the primary symptoms is the objective of the work.

3 Objective of the Study

The objective of the proposed research is to design deep convolutional neural network model for automatic detection of stroke and lesion segmentation and also to evaluate the effectiveness of hybrid-classification approaches for the detection and segmentation of stroke.

The main challenge is to continuously enhance the ways of detecting brain strokes in patients. The research objectives are:

- To predict the type of stroke using earlier symptoms.
- To determine the location and severity of the stroke.
- To propose and analyze a fully-automated technique for segmenting stroke affected areas.
- To perform exceptional classification without affecting crucial features in MRI image.

Main objective of the research work is to propose a prediction model by combining hand gestures and bio signals obtained by the patients instead of EMG and EEG obtained from Electronic Health Records. Proposed hybrid model is to be implemented using Deep learning algorithm through which prediction rate and the accuracy can be enhanced. It is used for early detection of stroke and diagnosis of stroke lesion. The proposed technique is used to improve the segmentation of MRI images. Automatic classification with improved techniques can be used to find the severity of stroke without human intervention.

This paper is systematized as follows: Stroke analysis is discussed in detail in Sect. 1. Selection of methods used in the literature is explained in Sect. 2. Section 3 contracts objective of the study. Components of CNN, Architecture and the algorithm used in CNN are explained in Sect. 4 to provide systematic understanding. Section 5 deals about the research methodology which includes the data used for preprocessing, Sect. 6 analyses the performance evaluation metrics and result. Finally, Sect. 7 concludes the paper.

4 Proposed Work

Convolution is a linear operator, dot-product like correlation, not a matrix multiplication, but can be implemented as a sparse matrix multiplication, to be viewed as an affine transform. A CNN arranges its neurons in three dimensions (width, height, depth). Every layer of a CNN transforms the 3D input volume to a 3D output volume. It can handle volumetric images like bio medical signal and facial differences.

Figure 2 explains the features extracted. CNN has many features, including a hierarchy-based learning system, automatic feature extractors, and multitasking. CNN learns the feature of the patient through a backpropagation algorithm. The objective of the proposed research was to develop hybrid-CNN model for automatic prediction and segmentation of stroke [10].

In the proposed architecture, two datasets are used: a healthy dataset and a spiral and wave image dataset affected by the stroke. These records are from the affected patients. The image datasets are initially located in this architecture, which is stored in Google Drive [11]. Included in this architecture is a training module that allows you to select a specific image.

C. *DenseNET*

DenseNet is a modern CNN visual object recognition architecture that has reached the state of the art with fewer parameters.

With some basic changes, DenseNet is very similar to ResNet [12]. DenseNet-201, the architecture of DenseNet-121 [$5 + (6 + 12 + 24 + 16) \times 2) = 121$] was used in this study. The total number of layers are explained in Table 2.

In general, CNNs compute the output layers (gth) using H_g (.) on previous layer X_{g-1}. The basic functions are given in Fig. 3.

$$X_g = H_g\left(X_{g-1}\right) \tag{1}$$

DenseNets uses a simple model to communicate among layers. The equation is then transformed again into:

$$X_g = H_g\left(\left[X_0, X_1, \ldots X_{g-1},\right]\right) \tag{2}$$

The architecture of the CNN is shown in Fig. 4.

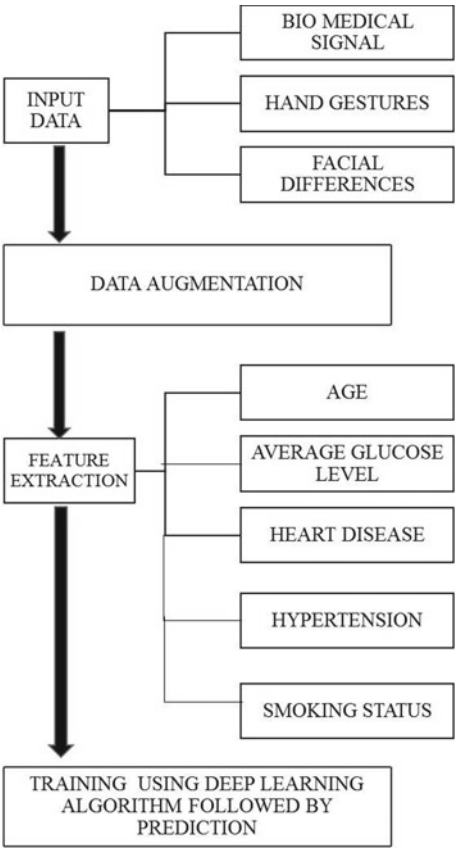


Fig. 2 Proposed work—block diagram

Table 2 Structure of DenseNet—121

DenseNet-121	Total no of layers
Convolutional layers and pooling layers	5
Transition layers	3−(6, 12, 24)
Denseblocks	2−(1 × 1 And 3 × 3 Conv)

D. *Augmentation*

Augmenting technique allows developers to greatly increase the size of dataset [13]. To increase the benefits of training and decrease the use of network regulation, image magnification was used in this study. Figure 5 gives the basic techniques used in the work.

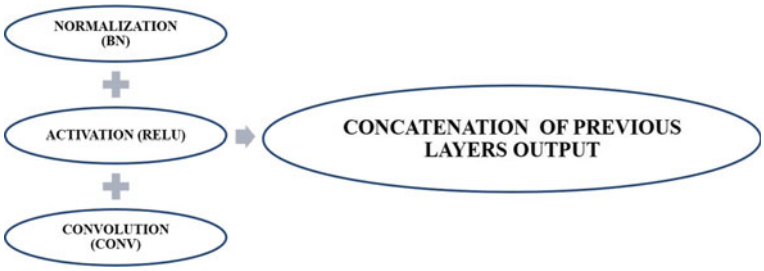


Fig. 3 Concatenation of separate blocks

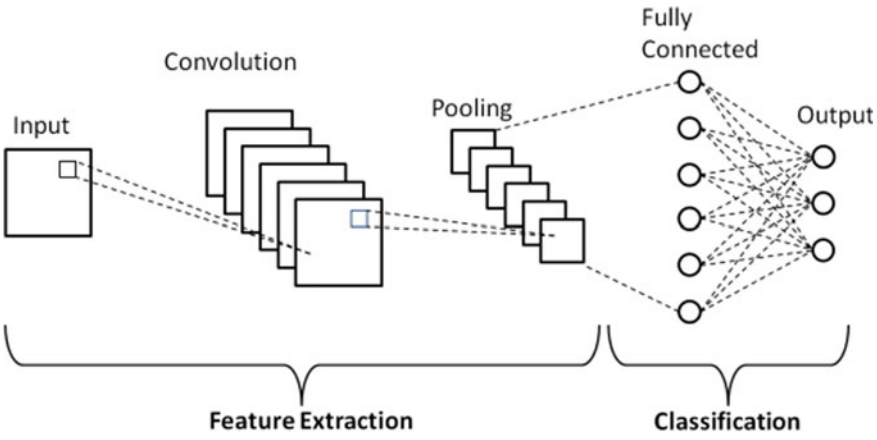


Fig. 4 Convolutional neural network components

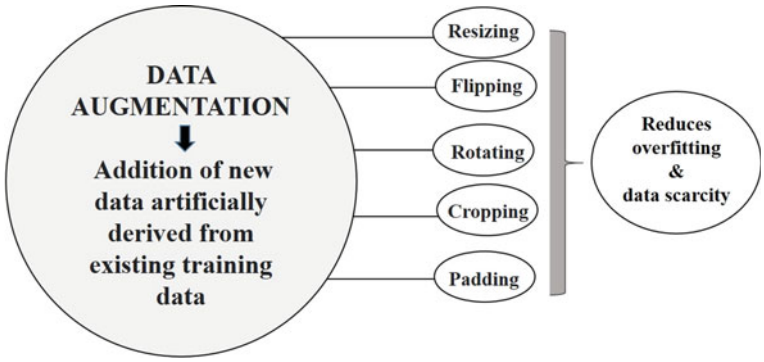


Fig. 5 Data augmentation techniques

Datagen = ImageDataGenerator(horizontal_flip = True,vertical_flip = True,rotation_range = 20,zoom_range = 0.2, width_shift_range = 0.2,height_shift_range = 0.2,shear_range = 0.1,fill_mode = "nearest").

pretrained_model3 = tf.keras.applications.DenseNet201(input_shape = (100,100,3),include_top = False,weights = 'imagenet',pooling = 'avg').

pretrained_model3.trainable = False.

5 Research Methodology

A. *Data*

To detect strokes with a CNN based on the DenseNet architecture, a real patient image dataset from an IGGH in Puducherry was used in this study. Kaggle's publicly available dataset was also used for research development. Silent strokes in various areas of the brain may lead to inflammation that causes dopamine neurons in the basal ganglia to die, leading to Parkinson disease. Based on the literature survey it is found that the drawing speed was slower and the pen pressure lower among Parkinson's patients. This was especially pronounced for patients with a more acute/advanced forms of the disease. The patients are given a spiral dotted sheet. The accuracy of the drawing pattern and the time taken by the patient to draw the circles are notes manually in real time practice. The pattern sheet is given in Fig. 6.

The classes of images were divided into four labels: healthy—spiral, healthy—wave, unhealthy—spiral, unhealthy—wave. Figure 7a and b show some patterns of healthy and unhealthy spiral patterns.

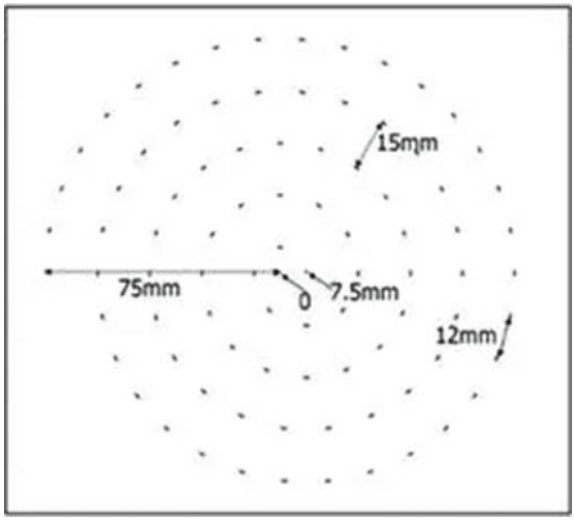


Fig. 6 Spiral/wave test pattern sheet

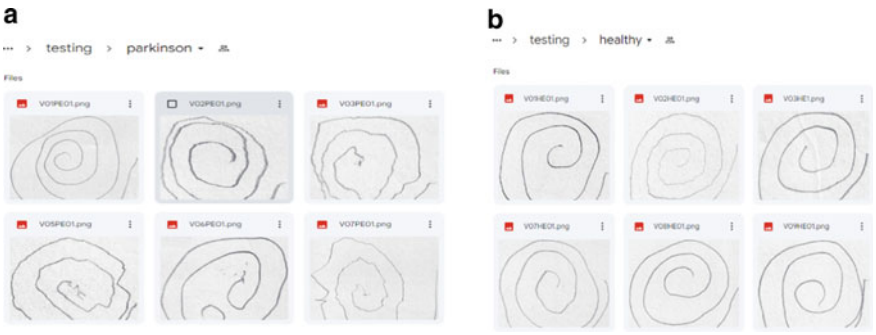


Fig. 7 **a** Spiral dataset—healthy and **b** spiral dataset—unhealthy

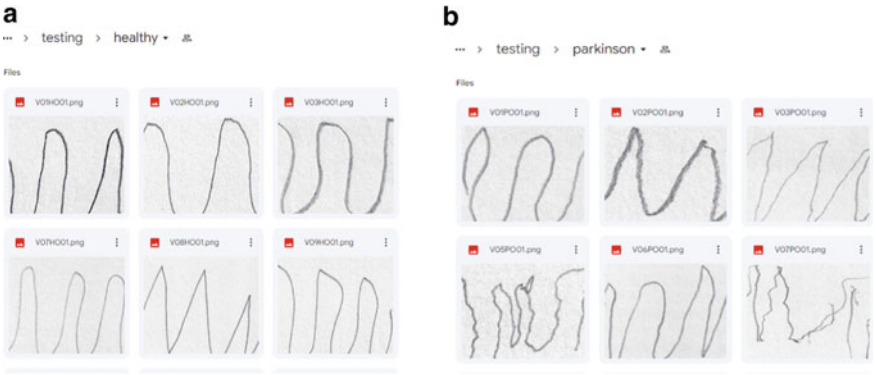


Fig. 8 **a** Wave dataset—healthy and **b** wave dataset—unhealthy

Figure 8a and b show a healthy and an unhealthy wave pattern obtained by wave pattern sheets.

B. *Preprocessing*

The goal of image preprocessing is to suppress unwanted rotations in the image, resize, and normalize the image for further processing. In the previous literature, there are numerous image preprocessing techniques based on the requirement of model building [14]. The techniques used are Handling missing values, Removing duplicate values, Label encoding and Normalization. Split the dataset into train and test set with 80:20 ratio. The required features are generated and features are selected using XG Boost and Principal Component Analysis methods [15].

6 Results and Discussion

The model accuracy and loss plot of the DenseNet CNN are shown in Fig. 9a and b. Performance of medical image understanding algorithms is evaluated using a variety of metrics. The DenseNet-121 has an overall accuracy of 96% with 97% sensitivity (recall), 97% precision, 97% $F1$-score, and 96% root mean square value. The results are tabulated in Table 3.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

$$\text{Accuracy} = \frac{100 - (\text{TP} + \text{TN})\,\%}{\text{Total No Of Images}} \tag{5}$$

Precision, recall, accuracy of the system is given by the Eqs. (3), (4) and (5), respectively.

where

- True Positive (TP)—the no. of correctly detected stroke regions.
- False Negative (FN)—the no. of missed stroke lesions.
- False Positive (FP)—the no. of noise lesions detected as stroke lesions.
- True Negative (TN)—the no. of incorrectly detected stroke regions.

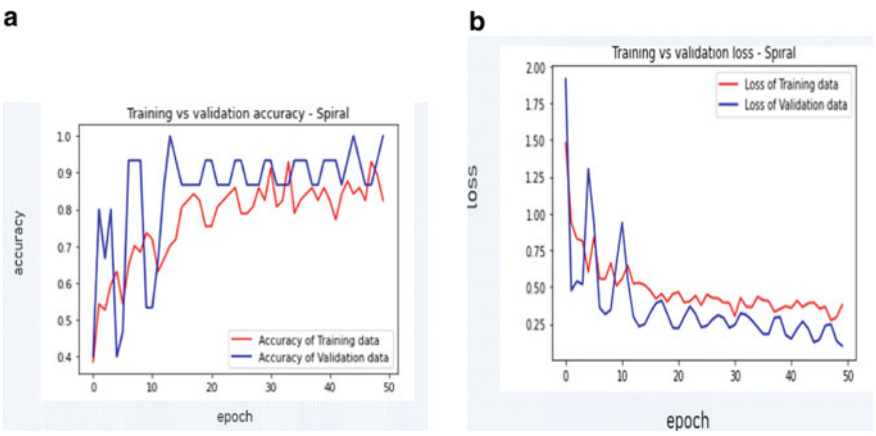


Fig. 9 **a** Training and validation accuracy and **b** training and validation loss

Table 3 Performance metrics

S. no	Measurement	Value
1	Accuracy	0.9655
2	$F1$_score	0.9721
3	Precision	0.9762
4	Recall	0.9722
5	Kappa_score	0.9530
6	MAE	0.0344
7	MSE	0.0334
8	R_square	0.9654

7 Conclusion

In this paper, a detailed review of recent research on convolutional neural networks is presented to highlight the efficiency of CNN. Convolutional neural networks are widely used in both industry and research projects due to their advantages such as local connections, weight sharing, and down-sampling. In conclusion, some important advantages of this study are presented below: (1) The handshake signals obtained from the spiral and wave test provide better performance in stroke detection compared with other bio signals, (2) In summary, the physicians can predict stroke at an earlier stage, and (3) The DenseNet-121 has an overall accuracy of 96% and a sensitivity of 97% in predicting stroke at an earlier stage.

References

1. Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees for brain stroke detection. Mach Learn 63(1):3–42
2. Bhanu Prakash KN, Gupta V, Hu J, Nowinski WL (2008) Automatic processing of diffusion-weighted ischemic stroke images based on divergence measures: slice and hemisphere identification, and stroke region segmentation. Int J Comput Assist Radio Surg 3(6):559–570
3. Gillebert CR, Humphreys GW, Mantini D (2014) Automated delineation of stroke lesions using brain CT images. NeuroImage: Clin 4(10):540–548
4. Mitra J, Bourgeat P, Fripp J, Ghose S, Rose S, Salvado O, Connelly A, Campbell B, Palmer S, Sharma G, Christensen S, Carey L (2014) Stroke laterality bias in the management of acute ischemic stroke. Neuro Image 98(3):324–335
5. Choi Y-A, Park S-J, Jun J-A, Pyo C-S, Cho K-H, Lee H-S, Yu J-H (June 2021) Deep learning-based stroke disease prediction system using real-time bio signals. Int J Sens Imaging Sens 21(13):1–17
6. Yu J, Park S, Kwon S-H, Ho CMB, Pyo C-S, Lee (Sept 2020) AI-based stroke disease prediction system using real-time electromyography signals. Int J Appl Sci MDPI 10:1–19
7. Umirzakova S, Abdusalomov A, Whangbo TK (2019) Fully automatic stroke symptom detection method based on facial features and moving hand differences. Int Symp Multimedia Commun Technol (ISMAC) IEEE 12:1–5

8. Tuladhar A, Schimert S, Rajashekar D, Kniep HC, Fiehler J, Forkert ND (2020) Automatic segmentation of stroke lesions in non-contrast computed tomography datasets with convolutional neural networks. Spec Sect Deep Learn Algorithms Internet Med Things IEEE J 8:94871–94879
9. Kuang H, Menon BK, Qiu W (Oct 2019) Segmenting hemorrhagic and ischemic infarct simultaneously from follow-up non-contrast CT images in patients with acute ischemic stroke. IEEE Trans Med Imaging 7:39842–39847
10. Liu L, Kurga L, Wu FX, Wang J (2020) Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease. Med Image Anal 10179 10(8):152–160
11. Gottapu RD, Dagli CH (2018) DenseNet for anatomical brain segmentation. Proc Comput Sci 140:179–185
12. Lodhi B, Kang J (2019) Multipath-DenseNet: a supervised ensemble architecture of densely connected convolutional networks. IEEE Trans Inf Sci 48:63–72
13. Sajjad M, Khan S, Muhammad K, Wu W, Ullah A, Baik SW (2019) Multi-grade brain tumour classification using deep CNN with extensive data augmentation. J Comput Sci 30:174–182
14. Punitha Lakshmi R, Vijayalakshmi V (Jan 2023) Analysis of brain stroke using deep convolutional neural network models: review. Int J Curr Sci (IJCSPUB) 13:175–184
15. Najafabadi MM, Villanustre F, Khoshgoftaar TM (2015) Deep learning applications and challenges in big data analytics. Big Data 2:1–21 10(2):79–81

Two Decades of Innovative Work Behaviour—A Bibliographic Study

Saswati Chakraborty and Suruchi Pandey

Abstract The present conceptual study examines the existing literature in the field of "Innovative work behaviour" in last two decades across different countries and culture, along with exploring the association other variables in an organized way. This study conducted a contextual analysis of top cited and rated peer-reviewed journal research papers in "Innovation work behaviour" from the period of 2002–2022. The article adopted an established five steps' methodology comprising systematic literature review and bibliometric analysis. The sampling process was performed by employing the PRISMA approach. This research has identified 29 antecedents with six different themes or clusters, 21 mediator, 18 moderators, and key outcomes highlighted that are responsible for "innovative work behaviour". The paper discusses the results of "innovative work behaviour" studies in the last two decades and attempts to find theoretical and managerial implications, followed by the limitations of the study, and outlining directions for future research. Thus, contributed to conceptualize the existing literature by connecting the dots and gaps identified in our evaluation.

Keywords Innovation work behaviour · Systematic literature review · Disruption

1 Introduction

Today, organizations are having an immense demand from their employees to innovative at work to gain competitive advantage, economic growth and increase market share, which thus help the organization to survive by delivering new or improved products or services. Most transformed corporation is now creating a demand for their products within the customers through the means of continuous innovation.

S. Chakraborty (✉)
Symbiosis International (Deemed University), Lavale, Pune, Maharashtra, India
e-mail: saswati.chakra@gmail.com

S. Pandey
Symbiosis Institute of Management Studies, Symbiosis International (Deemed University), Pune, Maharashtra, India
e-mail: suruchi.p@sims.edu

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_27

This makes "Innovation work behaviour" within employees as an essential ingredient for organizational competitive advantage and to sustain changing world economic conditions [1]. Companies are looking for opportunity for becoming a disruptor to be ahead of the curve in the market before they get disrupted by someone else through innovation. There are demand and expectation from intellectual capital to get involve in innovation is gaining focus. Therefore, a developing need to understand those aspects of employee behaviour that encourage employees to participate in innovation work. If we flash back in history, innovation was always part of human evolution and not something new as a concept. It is as immense part in the growth of mankind. Human is being started innovation from the very first day of human existence, by doing something new and inimitable for survival and protect itself and its environment. The concept of "innovative work behaviour" was developed by many researchers in past few decades like West and Farr [2], Farr and Ford [3], Scott and Bruce [4], Janssen [5].

Innovation is the most essential ingredient for organization success and survival in today's world. Innovation at work is a discretionary behaviour which employee may or may not exhibit during the lifecycle of employment; hence, it becomes very important for organization to provide an environment, which enhances "Innovation work behaviour" amongst employees. There are various factors influencing and boosting innovation at work, which may influence direct and indirect. In this paper, author tried to look at one of the key factors and themes prompting innovation.

1.1 Innovative Work Behaviour

According to West and Farr [2], innovation is defined as "intentional introduction and application within a role, group or organization of ideas, processes, products or procedures", whereas Farr and Ford [3] defined "innovative work behaviour" as the individual's behaviour to generate new and useful ideas to improve products, processes, and procedures. "Innovation work behaviour" was also later described by Janssen [5] as "the intentional creation, introduction and application of new ideas within a work role, group or organization, to benefit role performance, the group, or the organization".

"Innovative work behaviour" to be personality-based attribute and defined it individual willingness to be creative at work [6]. Although "innovative work behaviour" benefits the organizations, it is a discretionary effort that may not be directly and formally rewarded argued [7].

Scott and Bruce [4], Janssen [5] defined stages of "innovative work behaviour" as "idea generation, idea promotion, and idea implementation". Idea generation includes search and generation of new ideas. Idea promotion is the second stage, which promotes the benefit from the idea, requires endorse ideas, and finds promoters to influence authority. Idea implementation is the third stage, which calls for effort and action-oriented attitude to implement the ideas. De Jong and Den Hartog [8] extended the stage of idea generation by dividing it into opportunity exploration

and idea generation, based on proposal of Scott and Bruce [4], that idea generation comprises identification of problem statements and generation of new and useful ideas. Later, De Jong and Den Hartog [9] argued that the discovery of opportunities is seen as a behaviour outcome preceding idea generation and proposed four dimensions of innovative behaviour as opportunity exploration, idea generation, idea championing, and idea application.

2 Methodology

The study conducted a solid contextual study on the existing literature published form a period of year 2002–2022 on "Innovative work behaviour".

2.1 Data Collection

Data source of Scopus, Web of Science, JSTOR, ABDC, Science Direct (Elsevier), Taylor & Francis, Directory of Open Access Journals, and Google Scholar are used as the online search engines for the study. These databases are known for having the rich and broadest range of research papers from the best journals in the areas of social sciences, Management, Social Psychology, Creativity, and Innovation Management. The research used Google Scholar to find the publication list on—"Innovative work behaviour". The search found 437 articles published for the period of 2002–2022.

Year-wise break-up of the total number of papers came in the search for "innovative work behaviour" as shown in Fig. 1.

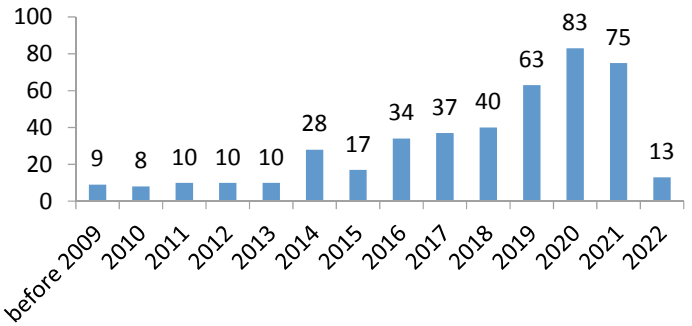


Fig. 1 Number of papers year wise. *Source* Authors

2.2 Selection Criteria

The first inclusion criteria for papers was citation, where highest citations of 2618 to lowest 10 citations were selected. Total of 80 numbers of research papers were initially identified. Now next selection criteria included papers from top ranked journals; thus, 65 papers were selected from initially included list of 80 papers, which comprised research papers from journal ranked between 1 and 87. An exception was made to include six research papers because of their high citation numbers. Therefore, exclusion of only 15 papers from low ranked journals was made and the final list of selected search paper included 65 most cited research papers and from top ranked journals for the study, and the below graph shows the ranking of the journal for selected papers. The process flow of selection of papers for this study is shown in Fig. 2.

3 Findings

The papers were analyzed on, research type, demographic details, methodology, theoretical framework(s), variables (mediator/moderators), research question and objective, discussion, and research findings.

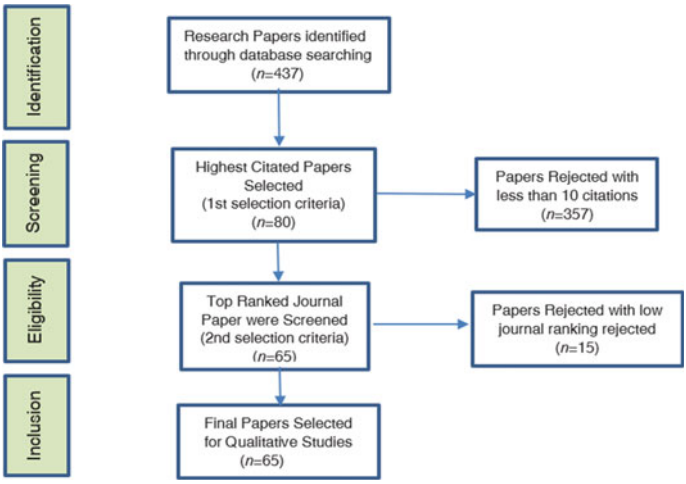


Fig. 2 Selection process shown by flowchart. *Source* Authors

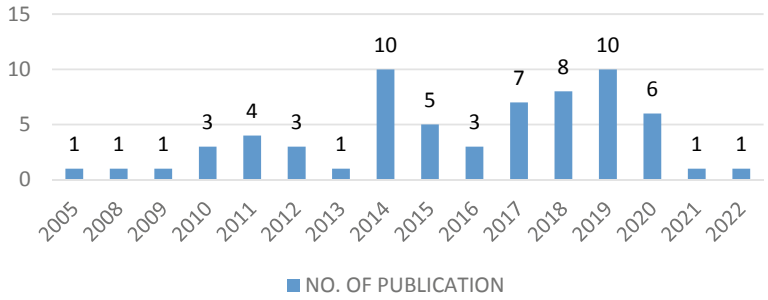


Fig. 3 Year and frequency of publication for the selected articles. *Source* Authors

3.1 Publication Year and Journals

The analysis was done on research done in the period of 2002 and 2022. Figure 3 shows the year-wise number of selected research articles found from 2002 to 2022. Between 2005 and 2013 (8 years), only 14 research articles were published, whereas 52 articles were published within the period of 2014–2020 (6 years), which means that in last decade which is year between 2012 and 2022 shows more focus in innovation research. 2014, 2017, 2018, 2019, and 2020 are the pick years which show maximum research done during those periods.

By looking at the articles contribution per journals as shown in Table 1, "Creativity and Innovation Management", "Journal of Occupational and Organizational Psychology", "The International Journal of Human Resource Management" and "European Journal of Work and Organizational Psychology", shows the maximum contribution towards innovation studies. Thirty-five out of 65 articles are published by below top ten journals, which contribute about 55% of the total selected research papers for the study.

3.2 Research Design, Data Analysis Techniques

Similarly, by looking at type of research in scope for "Innovative work behaviour" has shown that 84% of the studies are quantitative in nature, followed by 8% qualitative, and rest 8% are on scale development and longitudinal types. Within 84% of the quantitative studies, factor analysis and structural equation modelling (49%) and regression analysis (39%) dominated the data analysis techniques.

Table 1 Number of articles per journal

Journal name	Count of articles
Creativity and innovation management	7
European journal of work and organizational psychology	6
Journal of occupational and organizational psychology	4
The international journal of human resource management	4
European journal of innovation management	3
International journal of innovation management	3
Journal of business and psychology	2
Journal of workplace learning	2
Management decision	2
Personnel review	2

Source Authors

3.3 Location of Publication

The review shows the studies on "Innovative work behaviour" grasped worldwide demand and importance as the studies were across various countries as shown in Fig. 4. Maximum research is done mostly within European countries like Germany (18%), the Netherland (14%), Belgium (11%), Italy (11%), and France (5%), which contribute around 58% of the total selected research papers for the study, followed by Asia, which adds to 26% of the total research, mostly from countries like India (11%), China (8%), Pakistan (5%), Bangladesh (2%), and Thailand (2%). Looking from per capita% average view, Belgium and the Netherland are having less per capita but still add to 25%, which indicates more focus on research on "innovative work behaviour". Continents like Australia and Africa show low research on the innovation behavioural studies.

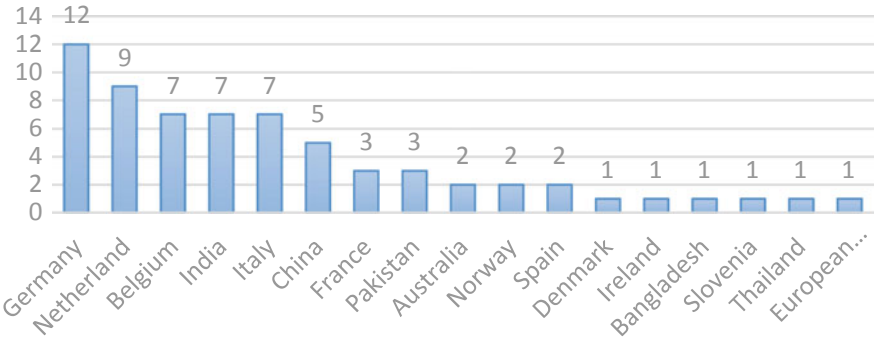


Fig. 4 Location of publication. *Source* Authors

Cites	Authors	Title	Source	Publisher	GSRank
3533	Janssen (2000)	Job demands, perceptions of effort-reward fairness and innovative work behaviour	Journal of Occupational and Organizational Psychology	Wiley Online Library	9
1997	De Jong & Den Hartog, 2010	Measuring innovative work behaviour	Creativity and Innovation Management	Wiley Online Library	1
842	Ramamoorthy, Flood, Slattery, & Sardessai (2005)	Determinants of innovative work behaviour: Development and test of an integrated model	Creativity and Innovation Management	Wiley Online Library	2
789	Agarwal, Datta, Blake-Beard & Bhargava (2012)	Linking LMX, innovative work behaviour and turnover intentions: The mediating role of work engagement	Career Development International	emerald.com	13
571	Agarwal (2014)	Linking justice, trust and innovative work behaviour to work engagement	Personnel Review	emerald.com	6
536	Reuvers et al. (2008)	Transformational leadership and innovative work behaviour: Exploring the relevance of gender differences	Creativity and Innovation Management	Wiley Online Library	5
357	Messmann & Mulder (2012)	Development of a measurement instrument for innovative work behaviour as a dynamic and context-bound construct	Human Resource Development International	Taylor & Francis	3
343	Bos-Nehles, Renkema & Janssen (2017)	HRM and innovative work behaviour: A systematic literature review	Personnel review	emerald.com	7
318	Radaelli, Lettieri, Mura & Spiller (2014)	Knowledge sharing and innovative work behaviour in healthcare: A micro-level investigation of direct and indirect effects	Creativity and Innovation Management	Wiley Online Library	12
302	De Spiegelaere, Van Gyes, De Witte, Niesen & Van Hootegem (2014)	On the relation of job insecurity, job autonomy, innovative work behaviour and the mediating effect of work engagement	Creativity and Innovation Management	Wiley Online Library	20

Fig. 5 Most cited and top ranked journal articles (*Source* Author)

3.4 Citation and Ranking of Articles

An effort was made to highlight the topmost cited articles in the field of "Innovative work behaviour" and articles from top ranked journals as shown in Fig. 5. Thus, provides a testimony of high-quality articles been carefully selected to be part of this research. Few highlights of the finding are, paper by Janssen [5] is cited 3533 times and score to be the most cited articles in the list, whereas De Jong and Den Hartog [9] is from most top ranked journals.

3.5 Antecedents and Consequences

Out of the 65 articles reviewed, the top five categories which are mostly studied as antecedent and consequence relationship influencing "innovative work behaviour" are job attributes (which include job performance, job autonomy, job resources, job demand, job insecurity, meritocracy and pay, performance-related pay) contributing 23% of the total research, followed by leadership (which includes transformational leadership, leader–member exchange, empowering leadership) which contributes

about 18%, organizational attributes (which includes employability, organizational climate, organizational culture, support for innovation, organizational social capital) by 12%, and employee–employer relationship (perceived organizational support, psychological contract, person-organization fit) and employee behaviour (learning orientation, organizational citizenship behaviour, work-related attitude, team learning behaviours, justice perceptions) which contribute 11% each towards the research, as shown in Table 2. This contributes to 74% amongst the total selected articles, within this top three variables are, transformational leadership (9%), leader–member exchange (8%) and Job Performance (4%) are the most examined variable in relation with "innovative work behaviour". But at the same time shows scarcity of studies in these areas of psychological Resources like psychological climate for innovation, motivation, Innovation-specific resources, Intellectual Capital, procedural and interactional justice, work engagement, work–life conflict, work environment due to limited number of research. The finding also highlighted various theoretical applications that were underlined in the research, and it was evident that social exchange theory (25%) is amongst the most adapted theory, followed by "Job Demands Resources model" (13%) and "Leader-member exchange theory" (13%) theories within "innovative work behaviour" studies.

Table 2 Antecedent variable used in the research

Categories/themes	Antecedent variable	Articles count
Job attributes	Job performance, job autonomy, job resources, job demand, job insecurity, meritocracy and pay, performance-related pay	17
Leadership	Transformational leadership, leader–member exchange, empowering leadership	13
Organizational attributes	Employability, organizational climate, organizational culture, support for innovation, organizational social capital	9
Employee–employer relationship	Perceived organizational support, psychological contract, person-organization fit	8
Employee behaviour	Learning orientation, organizational citizenship behaviour, work-related attitude, team learning behaviours (TLBs), justice perceptions	8
Work attributes	Work engagement, work–life conflict, work environment	7
Psychological resources	Psychological climate for innovation, motivation, innovation-specific resources, intellectual capital, procedural and interactional justice	7
HRM	Human resource management system, HRM practices	5

Source Authors

Table 3 Different measurement scale used (*Source* Authors)

Scales of "IWB"	Items and dimensions
Scott and Bruce [4]	Six items; one dimension
Bunce and West [10]	Five items; one dimension
Basu and Green [11]	Four items; one dimension
Janssen [5]	Nine items; three dimensions (Idea generation, promotion, and implementation)
Kleysen and Street [12]	Fourteen items; one dimension
Krause [13]	Eight items; two dimensions (creativity and implementation)
Reuvers [14]	Four items; one dimension
De Jong and Den Hartog [9]	Ten items; four dimensions (Idea Exploration, generation, championing, and implementation)

3.6 Measurement Scales

"Innovative work behaviour" was measured by different scales across the articles, but most used scales were: nine-item scale by Janssen [5] was used in 38% studies; ten-item scale by De Jong and Den Hartog [9] was used in 27% research, and 14% times six-item scale by Scott and Bruce [4] was used, as shown in Table 3.

3.7 Moderating and Mediating Variables and Key Outcomes

Precisely 26% of the studied research papers used moderating variables, 40% of the times mediating variables were used, 14% articles had control variables, and 20% articles with no moderation and mediation variables. Moderating effect of perceptions of fairness in pay at work on the relationship between "innovative work behaviour" and Job demands was highlighted [5]. According to Orth and Volmer [15], "innovative work behaviour" is enhanced by creative self-efficacy. Devloo et al. [16] highlighted, intrinsic motivation influences both the one-way relationship between job satisfaction and "innovative work behaviour" as well as the reciprocal interaction between the two. Agarwal [7] probe, "work engagement" plays a promoting role in the beneficial relationships between "leader-member exchange", "innovative work behaviour", and perceived organizational support. Age, gender, organizational climate, education level, and a few other control variables were examined. Figure 6 presents the conceptual model underlining, moderators, mediators, independent variables and highlights the key outcome of the interaction between the variables showing the possible antecedent and consequences for "innovative work behaviour". The model provides a framework for consolidating the various body of research on "innovative work behaviour". The independent factors examined with "innovative work behaviour" and their antecedent variables are shown in Column

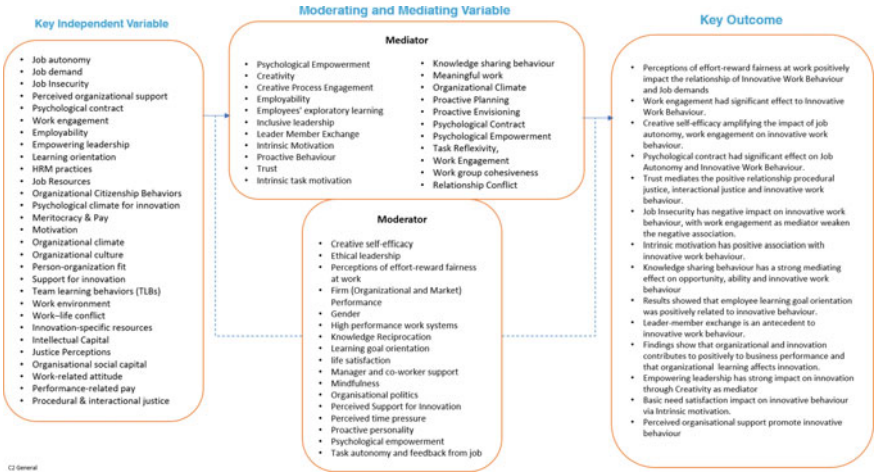


Fig. 6 Conceptual model of key independent, mediating and moderating, and outcome variables. (*Source* Authors)

1. The major outcome is displayed in Column 3, while mediating and moderating variables are represented in Column 2. The outcomes between the variables can be both be positive and negative.

4 Implications

The research demonstrated both practical and theoretical consequences. From lens of theoretical perspective, the articles' studies make a significant influence towards the "innovative work behaviour" topic and contribute to the active body of knowledge. Also, various factors like leadership impact, job performance, job autonomy, job resources and demand had little attention in past, which were the prime focus within the research in last 20 years, which thus provides a strong foundation and analysis on innovation and work-related behaviours. Moreover, from the practical implications' side, various researches suggested that the managers' leadership style plays an important part in creating an encouraging environment for "innovative work behaviour" amongst employees [17, 18]. The studies provide guidelines for manager like effective communication, regular one to one, receiving and giving performance feedback, career coaching, and skill enhancements as few useful practices in enhancing employee's innovative behaviours. At organization level, management can train and provide support to leaders to create conducive organizational culture that amplifies behaviours triggering innovative behaviour [19]. Leadership style and HRM practices are influenced by a supervisor's assessment of an organization's performance (high or low) [20]. Supervisors shape the environment of learning and skill development for team members [21]. Investing in employees' career potential might be a critical

success factor for attaining innovation in a knowledge-based industry. Additionally, De Spiegelaere et al. [22] aid in analyzing organizational performance from more varied perspectives and provide insight into possibilities, challenges, and solutions that actually exist in real-world settings for fostering "innovative work behaviour" amongst employees. Employees receiving continuous feedback over their work and performance will feel engaged to take extra-job initiatives leading towards workplace innovation [23]. Suggestion for HR managers to focus on organizational design in the job content and provide more autonomy at work to promote employee's innovation behaviour. Innovation establishes direct relationship with creativity, and organization can incentivize such behaviour through recognition or benefits [24]. Moreover, proactive goal setting and planning instrument by line managers is also an important factor for provoking innovative performance. Organizations to promote effective involvement in innovation should have clear goal define and a plan for skills' development via training programs for employees [25]. The organizational leaders suggested to create jobs content in a way that will meet both the employee's intrinsic and extrinsic demands of an employee, such as pay hike, promotion, performance dialogue, to find purpose in their work and increasing their level of job satisfaction [18].

5 Future Research Directions

The study highlights the dearth of research on employee behavioural studies, such as organizational citizenship behaviour and leader–member exchange, employee creativity, and "innovative work behaviour" in Asian and African contexts. In addition to being able to interpret phenomena in terms of the meanings and use innovative techniques to gain insight, qualitative research methods should be encouraged because there is a strong preference for quantitative measures over qualitative in terms of research design which creates future research opportunities [26]. According to the study, sectors like the hospitality industry, public services, and service industries received less attention in the past as highlighted [27]. The impact of innovative behaviour on the performance of individual has been suggested as a future research area [17]. Additionally, future study opportunities in the field of transformational leadership and to look at how leadership affects creativity and "innovative work behaviour" were emphasized [18, 19]. They also suggested opportunity for future research for understanding effect of creative self-efficacy, its relevant towards creativity, and innovative behaviour. Future researchers to understand the moderating effect of nationality and culture on innovation could also be a scope for future study [28].

6 Conclusions

The literature review pooled together the existing research work done on "innovative work behaviour" from last 20 years of study. To synthesize the review, there was an attempt to explore the concept of "innovative work behaviour". This review enhances our overall knowledge on "innovative work behaviour" and elaborates the relationships between other variables with innovative, highlight the mediator and moderators used in the study, research analysis country wise, theories used, and outcomes. The review proposed a model representing the antecedent and consequences of innovative work behaviour, which can set the foundation for further research opportunities. The review highlights the gaps in the literature analysis, which provides an important research opportunity for practitioners. Researchers can use this review to learn about the many measuring tools that are available, how to use them, and what situations they can be used in. Our sequential analysis provides evidence for the significance of "innovative work behaviour" and the variables influencing it. Thus, contribute to gaining understanding of current opportunities, limitations, and solutions for boosting innovative behaviour at work.

7 Limitations

The paper reviewed was sorted based on the study done on "Innovative Work Behaviour" in last two decades, similarly but many other variables which could have an impact on "Innovative Work Behaviour" are not covered under this study. Despite of using robust methodology for the review, the limitation of search criteria was only those research papers with top ranked journal and highest number of citations; thus, many other relevant research articles could not be part of this study.

8 Suggestions for Future Research

The literature review embarks a journey to understand the existing literature for "innovative work behaviour" in last two decades. In most of the results and finding, it has showed that research of "innovative work behaviour" is not popular within Asian and African countries; thus, a need to examine the linkages in a cross-national and cultural context is required [29]. As transformational disruption is the new phenomena in global economic world, thus, there is a demand of study disruptive innovation process which can be future topics for research [30]. Empirical research is recommended to conduct on other industries, amongst employees having knowledge based or professional in nature [31]. Need for continuous innovation is the biggest want from the organization around the world, hence studying innovation at work is an important topic to explore [32]. Need of more conceptual foundation studies is

identified as a future scope [33, 34]. Also studying the effect of pandemic, disruption, and new ways of working impact on innovative behaviour is next opportunity of future research [35].

References

1. Dervitsiotis KN (2010) Developing full-spectrum innovation capability for survival and success in the global economy. Total Qual Manag 21:159–170
2. West MA, Farr JL (1989) Innovation at work: psychological perspectives. Soc Behav 4(1):15–30
3. Farr JL, Ford CM (1990) Individual innovation. In: West MA, Farr JL (eds) Innovation and creativity at work: psychological and organizational strategies. John Wiley & Sons, Chichester, England, pp 63–80
4. Scott S, Bruce R (1994) Creating innovative behavior among R&D professionals: the moderating effect of leadership on the relationship between problem-solving style and innovation. In: Proceedings of 1994 IEEE international engineering management conference—IEMC, vol 94. IEEE, pp 48–55
5. Janssen O (2000) Job demands, perceptions of effort-reward fairness, and innovative work behavior. J Occup Organ Psychol 73:287–302
6. Hurt HT, Joseph K, Cook CD (1977) Scales for the measurement of innovativeness. Hum Commun Res 4(1):58–65
7. Agarwal U (2014) Linking justice, trust and innovative work behaviour to work engagement. Pers Rev 43(1):41–73
8. De Jong J, Den Hartog D (2007) How leaders influence employees' innovative behaviour. Eur J Innov Manag 10(1):41–64
9. De Jong J, Den Hartog D (2010) Measuring innovative work behaviour. Creativity Innov Manage 19(1):23–36
10. Bunce D, West MA (1995) Self perceptions and perceptions of group climate as predictors of individual innovation at work. Appl Psychol 44(3):199–215
11. Basu R, Green SG (1997) Leader-member exchange and transformational leadership: an empirical examination of innovative behaviors in leader-member dyads. J Appl Soc Psychol 27(6):477–499
12. Kleysen RF, Street CT (2001) Toward a multi-dimensional measure of individual innovative behavior. J Intellect Cap 2(3):284–296
13. Krause DE (2004) Influence-based leadership as a determinant of the inclination to innovate and of innovation-related behaviors: an empirical investigation. Leadersh Q 15(1):79–102
14. Reuvers M, Van Engen ML, Vinkenburg CJ, Wilson-Evered E (2008) Transformational leadership and innovative work behaviour: exploring the relevance of gender differences. Creativity Innov Manage 17(3):227–244
15. Orth M, Volmer J (2017) Daily within-person effects of job autonomy and work engagement on innovative behaviour: the cross-level moderating role of creative self-efficacy. Eur J Work Organ Psy 26(4):601–612
16. Devloo T, Anseel F, De Beuckelaer A, Salanova M (2015) Keep the fire burning: reciprocal gains of basic need satisfaction, intrinsic motivation and innovative work behaviour. Eur J Work Organ Psy 24(4):491–504
17. Purwanto A, Purba JT, Bernarto I, Sijabat R (2021) Effect of transformational leadership, job satisfaction, and organizational commitments on organizational citizenship behavior. INOBIS J Inovasi Bisnis dan Manajemen Indonesia 9(1):61–69
18. Pradhan S, Jena LK (2019) Does meaningful work explains the relationship between transformational leadership and innovative work behaviour? Vikalpa: J Decis Makers 44(1):30–40

19. Slåtten T, Svensson G, Sværi S (2011) Empowering leadership and the influence of a humorous work climate on service employees' creativity and innovative behaviour in frontline service jobs. Int J Qual Serv Sci 3(3):267–284
20. Stoffers JM, Van der Heijden BI, Notelaers GL (2014) Towards a moderated mediation model of innovative work behaviour enhancement. J Organ Chang Manag 27(4):642–659
21. Cooke WN, Meyer DG (2007) The effects of workplace climates on employee performance capacities: a knowledge-based perspective. Retrieved from Massachusetts Institute of Technology, Web site at http://web.mit.edu/sis07/www/cooke.pdf
22. Stoffers JMM, Van der Heijden BIJM (2009) Towards an HRM model predicting organisational performance by enhancing innovative work behaviour: a study among dutch SMEs in the province of Limburg. Bus Leadersh Rev 4:1–13
23. De Spiegelaere S, Van Gyes G, De Witte H, Niesen W, Van Hootegem G (2014) On the relation of job insecurity, job autonomy, innovative work behaviour and the mediating effect of work engagement. Creativity Innov Manage 23(3):318–330
24. Radaelli G, Lettieri E, Mura M, Spiller N (2014) Knowledge sharing and innovative work behaviour in healthcare: a micro-level investigation of direct and indirect effects. Creativity Innov Manage 23(4):400–414
25. Montani F, Odoardi C, Battistelli A (2015) Envisioning, planning and innovating: a closer investigation of proactive goal generation, innovative work behaviour and boundary conditions. J Bus Psychol 30:415–433
26. Halcomb E (2016) Understanding the importance of collecting qualitative data creatively. Nurse Res 23(3):6–7
27. Hornsby JS, Kuratko DF (2003) Human resource management in US small businesses: A replication and extension. J Dev Entrep 8(1):73
28. Yi L, Uddin MA, Das AK, Mahmood M, Sohel SM (2019) Do transformational leaders engage employees in sustainable innovative work behaviour? Perspective from a developing country. Sustainability 11(9):2485
29. Ramamoorthy N, Flood PC, Slattery T, Sardessai R (2005) Determinants of innovative work behaviour: development and test of an integrated model. Creativity Innov Manage 14(2):142–150
30. Markides CC (2012) How disruptive will innovations from emerging markets be? MIT Sloan Manag Rev 54(1):23
31. Bhatnagar J (2013) Mediator analysis in the management of innovation in Indian knowledge workers: The role of perceived supervisor support, psychological contract, reward and recognition and turnover intention. Int J Hum Resour Manage 25(10):1395–1416
32. Agarwal UA, Datta S, Blake-Beard S, Bhargava S (2012) Linking LMX, innovative work behaviour and turnover intentions. Career Dev Int 17(3):208–230
33. Bäckström I, Bengtsson LG (2018) Employee involvement in firm innovation–a mapping study of research on employee innovation. In: Academy of Management proceedings, vol 2018(no 1). Briarcliff Manor, NY 10510, p 16710
34. Bos-Nehles A, Renkema M, Janssen M (2017) HRM and innovative work behaviour: a systematic literature review. Pers Rev 46(7):1228–1253
35. Messmann G, Mulder RH (2012) Development of a measurement instrument for innovative work behaviour as a dynamic and context-bound construct. Hum Resour Dev Int 15(1):43–59

Comparison of Collision Avoidance and Path Planning Algorithms for Unmanned Surface Vehicle

K. A. Nethravathi, Srikanth Vadada, Divam Trivedi, and Abhilash Preetham Jayanthi

Abstract Artificial intelligence is a flourishing technology for autonomous surface vehicles, with approaches like artificial potential fields, evolutionary algorithms, fast marching methods, and others becoming more prominent for challenges like course planning and collision avoidance. This paper does a comparative analysis of current autonomous surface vehicle path planning and collision avoidance algorithms. The Rapidly Exploring Random Tree (RRT) method and its variants, the A* algorithm and its variants, and the Dynamic A* (D*) algorithm and its variants are among the path planning and collision avoidance algorithms described. Their efficiency is assessed using predetermined levels of environmental complexity, and a comparison is carried out. The observations show that if the environment is known before run time, the A* algorithm and its variants work efficiently; however, if the environment is dynamically changing, the D* method and its variants are a viable option. The simulations of the algorithms were carried out on Python IDE.

Keywords Collision avoidance · Path planning · Unmanned surface vehicle

1 Introduction

An unmanned surface vehicle (USV) [1] is a ship that navigates in an unmanned manner in a water environment. It includes a self-driving collision avoidance capacity that makes it suited for maritime tasks and safe navigation. The cornerstone for accomplishing nautical activities and also a reflection of its intelligence is an autonomous obstacle avoidance ability that operates on multiple algorithms. Oceanography, remote sensing, weapons delivery, force multipliers, surveying, anti-submarine warfare, surface warfare, environmental monitoring, electronic warfare,

K. A. Nethravathi (✉) · D. Trivedi · A. P. Jayanthi
R.V. College of Engineering, Bangalore, India
e-mail: nethravathika@rvce.edu.in

S. Vadada
LRDE, DRDO, Bangalore, India
e-mail: vadadasrikanth@lrde.drdo.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_28

and maritime interdiction operations' support are just some of the missions that USVs are used for in civil, military, and research applications. One of the most difficult tasks for USVs is to avoid (static and dynamic) impediments in their route. The research evaluates and contrasts several obstacle avoidance algorithms, as well as their benefits and drawbacks. Obstacle avoidance is both the foundation and one of the most pressing concerns in autonomous mobile robot navigation. To navigate through a dynamic environment, a robot needs to be equipped with a collision avoidance system that can deal with unknown hurdles and obstructions. It can be applied to a variety of machines, including automobiles, robots, and unmanned aerial vehicles. Collision avoidance is divided into two stages: Obstacle Detection and Collision Avoidance. For obstacle detection, different algorithms use different types of sensors. To communicate and avoid obstacles, information from a sensor is processed, and the controller sends a signal to the end effector. Modern techniques enable the robot to detour obstacles using quantitative measurement of the dimensions of obstacles using cameras and sensors, as opposed to primitive algorithms that used to stop robots to prevent collision. Various algorithms have been constructed, and simulation results are presented in this paper. The best method for avoiding obstacles is determined by the environment, equipment availability, and user requirements. Finally, the paper summarizes all the algorithms such as RRT Algorithm and its variants [1], A* algorithm [2] and its variants, and D* algorithm [3, 4] and its variants [5] by noting their benefits and drawbacks, which can be used to choose the best collision avoidance and path planning algorithm based on the needs and resources of the user.

1.1 Objectives

The main purpose of the work is to simulate various collision avoidance and path planning algorithms, to create different scenarios of operation in order of increasing complexity for the algorithms and to compare the performance of the selected algorithms in all the scenarios so that the best possible algorithm for the purpose of Unmanned Surface Vehicles could be selected.

2 Path Planning and Collision Avoidance Algorithms

Since autonomous cars must traverse their surroundings safely, effective collision avoidance needs to be a priority. The goal may be realized by using algorithms that assess several courses, identify potential collisions, and select the best route for safe travel. However, the employment of these collision avoidance algorithms must adhere to international regulations for preventing collisions at sea (COLREGS). The next subsections address two kinds of route planning and collision avoidance algorithms that are considered: search-based and sampling-based algorithms.

2.1 Sampling-Based Algorithm

Collision checking units in sampling-based algorithms offer information on the feasibility of various trajectories and sample a group of locations from the obstacle-free route while generating a roadmap for probable viable trajectories for route planning. The main concept behind a search-based algorithm is to traverse space by taking samples from it while treating the obstacle space as completely unknown (i.e. it will explore the state space without explicitly constructing barriers). Instead of being calculated directly, the full space is sampled all over. These methods provide a solution that is probabilistically complete. The Rapidly Exploring Random Tree (RRT) method is one of the sampling-based algorithms discussed.

2.1.1 Rapidly Exploring Random Tree (RRT) Algorithm

Robot path planning problem is a main concern these days in the field of automation. As in earlier path planning algorithms, the algorithms were only able to do either navigation or path planning while avoiding all the possible collisions. But this problem is addressed by the introduction of the RRT algorithm which was popularized by Dr. Karaman and Dr. Frazzoli. It is a modern optimized algorithm which is able to do both path planning as well as avoiding all the obstacles detected in the path of the robot. A RRT is a sampling-based motion planning method which aims to create a tree which fills the entire space incrementally by pushing the search towards the probable empty spaces. With each iteration, it tries to push the potential new nodes in the generated tree away from the nodes which were already generated from the previously carried out iterations. The points in the graph are quite randomly generated and are then connected to the nearest available node. At each iteration, when a vertex is created, it is made sure that it lies outside of the detected obstacle. While connecting to the next vertex also, it is made sure that it will avoid probable obstacles. The algorithm ends when the limit for the iterations hits or when the node is generated for the goal while simultaneously providing the optimized path for the problem.

2.1.2 Rapidly Exploring Random Tree* (RRT*) Algorithm

RRT* works in a similar manner as that of RRT but introduces three new procedures that are the near function, the wire procedure, and the parent function. These three procedures restructure the tree and give rise to the sub-optimality of the algorithm. The near function returns a set of nodes in the nearest neighbourhood by forming a sphere of radius with radius r and is focused at x. This ensures optimal convergence. The parent function helps to choose a parent node from the set of nodes returned by the near function while keeping in mind the lowest cost to reach the new node. If the lower cost path is found between the available nodes, the rewire procedure deletes

an edge from the graph between two nodes and inserts that node to the graph. The parent procedure chooses the parent from a sphere of radius for a newest node and the rewire procedure restructures the entire tree to find less costly paths keeping in consideration the new node. The parent and rewiring functions progressively improve the route cost of RRT* as the number of nodes added to the tree grows, resulting in asymptotic optimality.

2.2 Search-Based Algorithm

When a set of user requirements must be met with limited resources and minimal interdependencies between the given requirements, search-based methods are used to optimize the goal, find the best possible path, and when a set of user requirements must be met with limited resources available and minimum interdependencies between the given requirements, search-based methods are used.

2.2.1 A* Algorithm

Path finding is basically any method which involves the problem of obtaining a path between two points *A* and *B*. One of the most popularly used path finding algorithms is A* algorithm. A* algorithm is an improvised version of other previous algorithms such as Dijkstra's algorithm. A* algorithm has had numerous variations which were developed over the years, generally seeking to improvise various aspects of the algorithm, like minimizing memory requirements or speeding up calculations in dynamic environments or to improve efficiency. A* improvises the previous versions by adding a heuristic value *h(v)* which calculates the remaining area between the examined node and the goal node. The next node n is thus chosen by the minimum value of *f(n)* defined by Eq. 1. As long as the heuristic function is admissible that is it does not overestimate the distance to the target from the node, A* will find the shortest possible distance. Consider a square grid with numerous impediments on its route, with only a beginning cell and a target cell provided. The A* algorithm comes to the rescue if a rapid descent from the beginning cell to the target cell is necessary (if admissible). At each stage, the A* Algorithm selects a node based on a value - *'f'*, which is the sum of two other factors, namely *'g'* and *'h'*. It selects the node/cell with the lowest *'f'* at each stage and processes that node or cell. A* calculates the lowest distance by adopting the following Eq. 1.

$$f(x) = g(x) + h(x), \tag{1}$$

where '*g(x)*' and *'h(x)'* are defined as: $g(x)$ = the cost of moving from the starting point to a certain square on the grid, using the path that was generated to get there, and $h(x)$ = the expected cost of moving from that particular grid square to the final destination.

2.2.2 Real-Time Adaptive A* (RTAA*) Algorithm

In a real-time setting, robots must move smoothly and hence must seek for their path in real time. The RTAA* algorithm is capable of accomplishing this task because it picks its local search spaces in a fine-grained manner by updating the values of all the states in its local search spaces in a short amount of time. It improves the standard A* algorithm by computing and assigning an $f(x)$ value to each expanding state throughout the A* search.

2.2.3 Bi-directional A* (BI-A*) Algorithm

A bi-directional search algorithm conducts two searches at the same time. To find the target node, one search is started from the starting location, termed forward-search, while the other is started from the target node, called backward search. Bi-directional search splits a single search graph into two separate sub graphs, one starting from a beginning point and the other from a target point. When these two separated graphs intersect, the search comes to an end.

2.2.4 Lifelong Planning A* (LLPA*) Algorithm

Lifelong Planning A* algorithm is a modified version of A* that has been evolved over time and includes an adaptability factor that allows it to adjust to changes in the grid without having to recalculate the entire grid. It accomplishes this by updating, throughout the current search, the distance values from the starting point from the previous search and correcting them as needed. LLPA* also employs a heuristic (if applicable), which calculates the shortest path between any two nodes and the objective goal.

2.2.5 Dynamic A* (D*) Algorithm

D* was chosen because it is comparable to A*, except that it is dynamic that the arc cost parameters may change during the issue solving process. If the robot motion is properly coupled to the algorithm, D* provides ideal trajectories. The purpose of path planning is to get the robot from one location to another in the outside environment while avoiding all obstacles. Path planning's goal is to get the robot from one location in the outside environment to another while avoiding all obstacles and lowering the cost metric. Any issue environment can be described as a collection of states, each with its own cost, that represent robot locations connected by directional arcs. The robot starts in one state and progresses through the arcs to the goal state, *G*. With the exception of *G*, every state *M* has a back-pointer to the next state *N*, which is denoted by $b(M) = N$. D* employs back-pointers to represent paths to the goal. The arc cost function *c(M, N)* returns a positive value for the cost of travelling an arc

from state *N* to state *M*. If N does not have an arc to *M, c(M, N)* is undefined. If *c(M, U)* or *c(N, M)* is defined, the states *M* and *N* in the space are neighbours. The D* algorithm, like the A* algorithm, keeps an OPEN list of states. The OPEN list is used to communicate changes in the arc cost function as well as to compute path costs to states in the space.

2.2.6 Dynamic A* (D*) Lite Algorithm

D* Lite is a new re-planning algorithm that uses the same planning philosophy as D* but a different algorithm. D* Lite is shorter than D* since it compares priorities based on only one criterion, making it easier to remember them. It also removes the need for nested if-statements with complex conditions that can take up to three lines apiece, making the output easier to understand. These features also make it simple to extend it, as well as to use inadmissible heuristics and various criteria to boost efficiency. When the edge costs of a graph change as the robot travels closer to the end vertex, D* Lite calculates the shortest paths between the robot's current vertex and the final vertex on a regular basis. It makes no assumptions about how edge costs change, whether they grow or decline, whether they change close to or distant from the robot's present vertex, whether they change due to the environment or simply because the robot's initial estimates were changed. D* Lite can be used to tackle the problem of unfamiliar terrain goal-directed navigation. The landscape is designed to be an eight-node graph. At beginning, its edges have a single cost. The robot changes to infinite when it realises that it cannot traverse.

3 Simulation Results

To compare various algorithms, a similar environment should be provided. For the implementation of the discussed algorithms, a similar environment has been created to compare the time and space complexity. Additionally, different levels of complexity are also created to test the efficiency of all these algorithms.

3.1 Simulation of RRT and RRT Algorithms*

From the experimental results, it can be inferred that although quite robust and efficient, RRT and its variants take a lot of time and memory to reach the desired goal. As observed, they work efficiently in simple environments, but once the complexities of the environment are increased, the entire path planning process becomes costly. So, the implementation was carried out in only level 1 and level 2 as it took a lot of run time for next consecutive levels. The implementation of the algorithms in level 2 complexity is shown in Figs. 1 and 2.

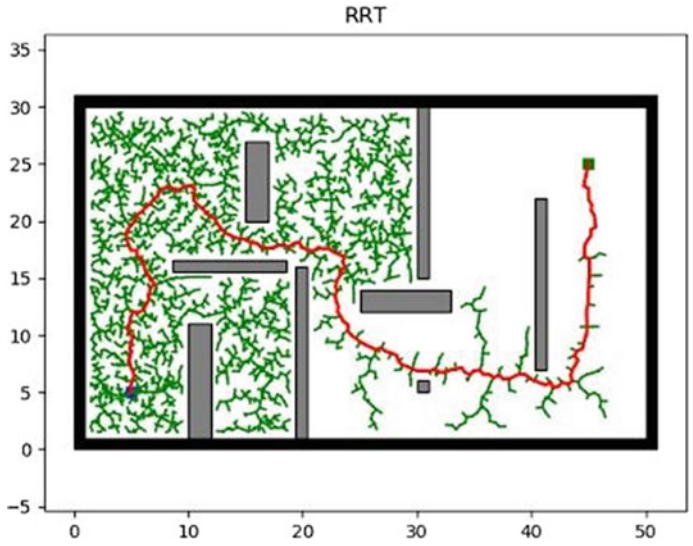


Fig. 1 Level 2 complexity environment for RRT

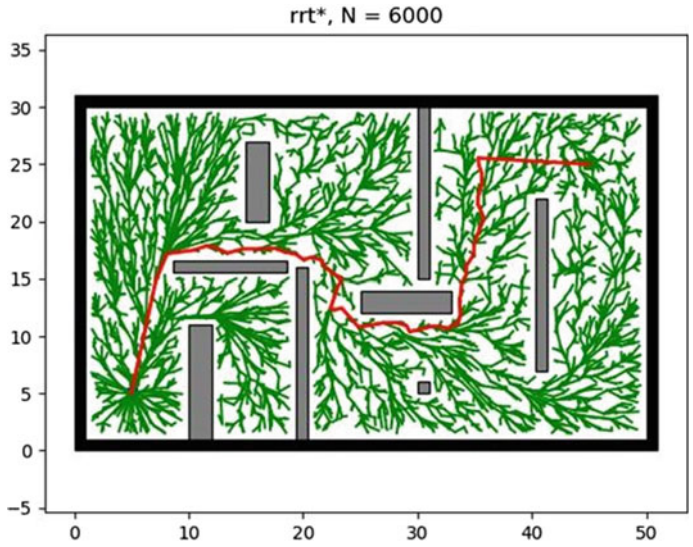


Fig. 2 Level 2 complexity environment for RRT*

The functionality of the RRT algorithm was verified on level 2 complexity of the environment as shown in Fig. 1. The total simulation time taken to reach the destination from the source is about 10.02 s, while the total distance of the selected optimal path found is 92.78 m.

Table 1 Simulation time (in secs) of each scenario

Scenarios	RRT	RRT*	A*	BI-A*
L1	1.89	106.90	0.03	0.06
L2	10.02	99.43	0.03	0.05
L3	Longer	Longer	0.13	0.07
L4	272.45	322.01	0.01	0.04
L5	Failed	Failed	0.07	0.14

Table 2 Simulation time (in secs) of each scenario

Scenarios	LLP*	RTA-A*	D*	D*LITE
L1	0.30	0.06	0.41	0.35
L2	0.33	0.07	0.28	0.27
L3	0.18	0.22	0.23	0.40
L4	0.34	0.12	0.22	0.30
L5	0.35	0.14	0.29	0.36

Table 3 Simulation distance (in metres) of each scenario

Scenarios	RRT	RRT*	A*	BI-A*
L1	86.9	58.3	50.5	50.5
L2	92.7	69.2	51.1	51.1
L3	Failed	Failed	115.5	116.8
L4	Failed	Failed	276.7	277.5
L5	Failed	Failed	310.1	310.17

The functionality of the RRT* algorithm was verified on level 2 complexity of the environment as shown in Fig. 2. The total simulation time taken to reach the destination from the source is about 99.4 s while the total distance of the selected optimal path found is about 69.28 m.

As it is mentioned earlier that RRT and its variants took very long run time, so the results were just included for the two levels of environment complexity as tabulated in Tables 1, 2, 3 and 4 also.

3.2 Simulation of A Algorithm and Its Variants*

The working of the A* algorithm and its variants Bidirectional-A*Algorithm Real-time Adaptive A* Algorithm, Lifelong Planning A* Algorithm were verified on five different levels of complexity of the environment. The simulation results for Level 5 are shown in Figs. 3, 4, 5, and 6.

Table 4 Simulation distance (in metres) of each scenario

Scenarios	LLPA*	RTA-A*	D*	D*LITE
L1	50.5	28.5	50.5	50.5
L2	51.1	26.7	51.1	51.1
L3	111	33.2	115.3	115.3
L4	101	30.4	276.8	101.4
L5	310.1	82.8	310.2	310.1

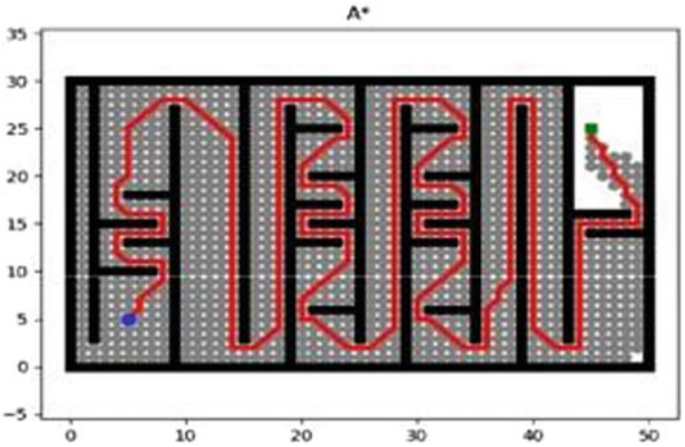


Fig. 3 Level 5 complexity environment for A*

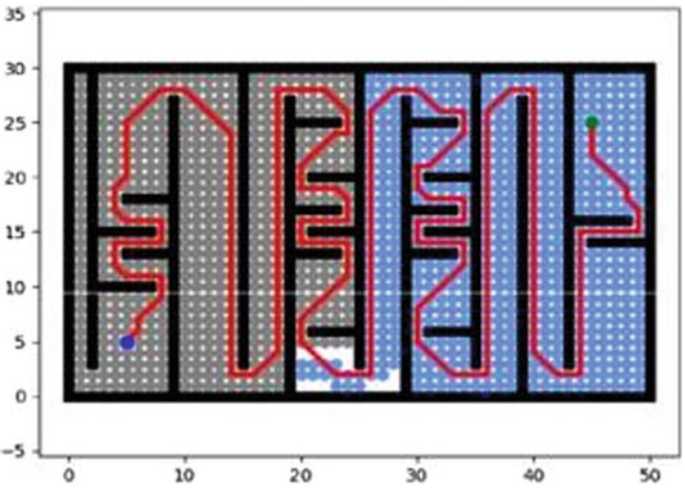


Fig. 4 Level 5 complexity environment for BI-A*

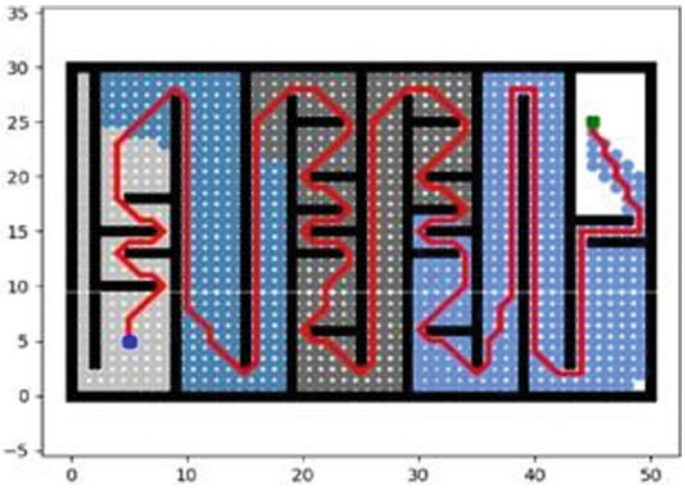


Fig. 5 Level 5 complexity environment for RTA-A*

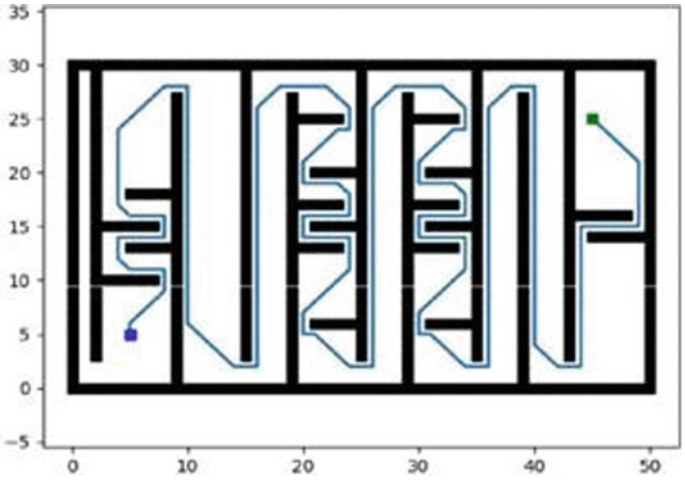


Fig. 6 Level 5 complexity environment for LLP-A*

The functionality of the A* algorithm was verified on level 5 complexity of the environment as shown in Fig. 3. The total simulation time taken to reach the destination from the source is 0.07 s, while the total distance of the selected optimal path found is 310.1 m.

The functionality of the BI-A* algorithm was verified on level 5 complexity of the environment as shown in Fig. 4. The total simulation time taken to reach the destination from the source is 0.14 s, while the total distance of the selected optimal path found is 310.15 m.

The functionality of the RTAA* algorithm was verified on level 5 complexity of the environment as shown in Fig. 5. The total simulation time taken to reach the destination from the source is 0.14 s, while the total distance of the selected optimal path found is 82.89 m.

The functionality of the LLPA* algorithm was verified on level 5 complexity of the environment as shown in Fig. 6. The total simulation time taken to reach the destination from the source is 0.35 s, while the total distance of the selected optimal path found is 310.12 m.

3.3 Simulation of D and D* Lite Algorithms*

The working of the D* and D* Lite algorithms was verified on five different levels of complexity of the environment. The simulation results for level 5 are shown in Figs. 7 and 8, respectively.

The functionality of the D* algorithm was verified on level 5 complexity of the environment as shown in Fig. 7. The total simulation time taken to reach the destination from the source is 0.29 s, while the total distance of the selected optimal path found is 310.12 m.

The functionality of the D* Lite algorithm was verified on level 5 complexity of the environment as shown in Fig. 8. The total simulation time taken to reach the destination from the source is 0.36 s, while the total distance of the selected optimal path found is 310.12 m.

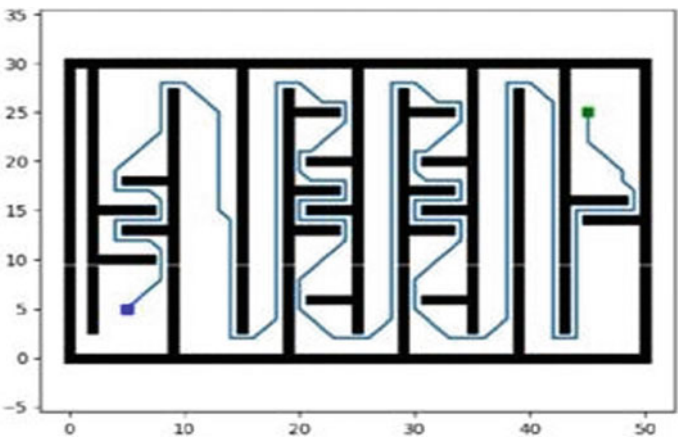


Fig. 7 Level 5 complexity environment for D*

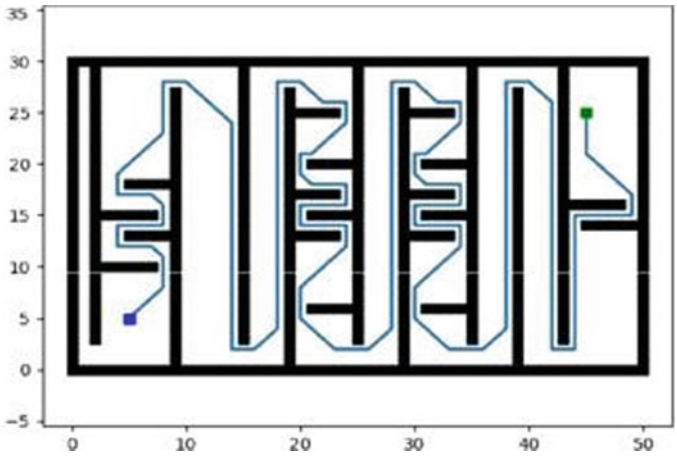


Fig. 8 Level 5 complexity environment for D* LITE

4 Analytical Results

For different levels of complexity in terms of obstacles, based on the same environment, all the discussed algorithms were tested for their efficiencies and the corresponding time taken to reach from the start to the goal was also calculated. The following section discusses the results inferred from the carried-out simulation of the algorithms.

Tables 1 and 2 enlist the simulation results for different compared algorithms in terms of time taken from the source node to the goal. Similarly, the simulation results for the compared algorithms in terms of distance of the path covered from the start point to the destination is shown in Tables 3 and 4.

As the results indicate, RRT algorithm and RRT* algorithm failed after running on the second complex environment level. These algorithms usually take a lot of run time because they rapidly keep exploring the space long after the solution is found also. This costs additional memory and hence takes a lot of time to conclude to an optimal solution. The other algorithms, however, found the solution in all the levels of the environment and the corresponding results were tabulated.

5 Graphical Results

The graphical representation of time taken by each of the algorithms at different levels of environment complexity is shown in Fig. 9.

As the graph indicates, D* algorithm and A* algorithm are extremely efficient in terms of the time taken to reach from the start to goal, which means that both of them have very less time complexity in comparison to all the other algorithms. The

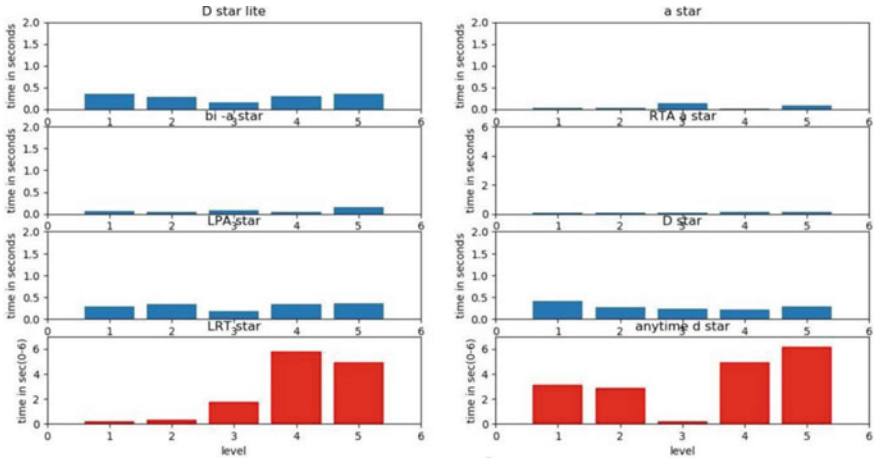


Fig. 9 Time complexity graph (time vs. level)

graphical representation of the length of the path formed by each of the algorithms at different levels of environment complexity is shown in Fig. 10.

The graphical results both in terms of time and distance conclude that D* Algorithm is the best suitable choice algorithm for the path planning and collision avoidance purpose. From the graph, it can be clearly inferred that RTA-A* algorithm and LRTA* algorithm are extremely efficient in terms of the path taken to reach from the start to goal, which means that both of them have very less space complexity in comparison to all the other algorithms.

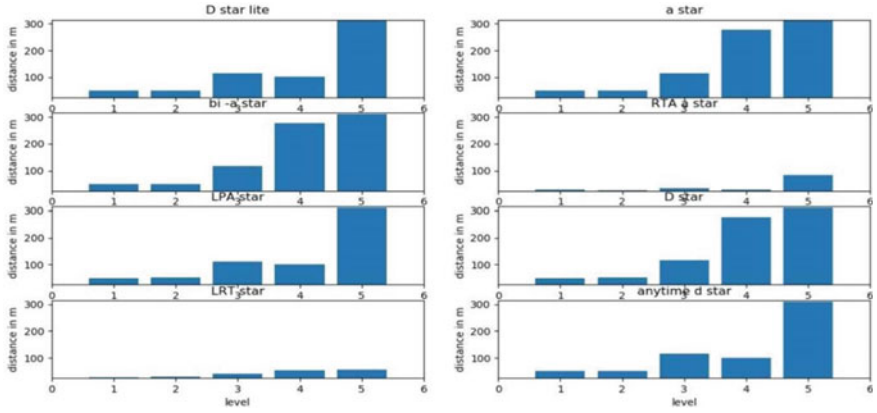


Fig. 10 Space complexity graph (distance vs. level)

6 Conclusion

After the comparison of the results for four different environments for all the algorithms discussed, a few algorithms were omitted because of the high time and space complexity exhibited by them. For the last environment simulation, D*, D* Lite, A*, Bidirectional A*, RTA*, and LPA* were considered and the rest of the algorithms were discarded because of their non-feasibility. From the simulation results, both D* and A* and their variants perform well in complex environments. However, these algorithms also have certain constraints. In the case of dynamically changing environment D* algorithm is best suited as it is less complex to run. However, if the environment is particularly known, A* and D* can be considered for efficient results. Overall, from the work done, it is concluded that A* and its variants would be the best choice for the purpose of collision avoidance for the Unmanned Surface Vehicles (USVs).

7 Future Scope

The work carried out here regarding the choice of collision avoidance and path planning algorithms for USVs was entirely on a predefined environment set. The efficiency of these algorithms in practical dynamic environment is uncertain. Furthermore, the entire study is based on the collision avoidance behaviour of the USVs. Therefore, possible advancement in the study could be the consideration of the dynamically changing environment and other aspects of the USVs also. Further, the entire analysis of the algorithms and the USVs behaviour was done using simulation processes on a software. The behaviour of the algorithms on a particular hardware and dynamically changing environment remains unpredictable. The future scope of this work is the further implementation and testing of these algorithms on hardware level.

References

1. Vagale A, Oucheikh R, Bye R, Osen O, Fossen T (2021) Path planning and collision avoidance for autonomous surface vehicles I: a review. J Marine Sci Technol
2. Cui S-G, Wang H, Yang L (2012) A simulation study of a-star algorithm for robot path planning, pp 506–509
3. Xu C, Xu Z, Xia M (2021) Obstacle avoidance in a three-dimensional dynamic environment based on fuzzy dynamic windows. Appl Sci 11(2):504. https://www.mdpi.com/2076-3417/11/2/504. ISSN-20763417. https://doi.org/10.3390/app11020504

4. Stentz A (1994) Optimal and efficient path planning for partially-known environments. In: Proceedings of the 1994 IEEE international conference on robotics and automation
5. Koenig S, Likhachev M (2002) D*lite. Am Assoc Artif Intell 476–483

A Review on Fingerprint Features and Algorithms

Deep Solanki and Nishant Doshi

Abstract In this paper, we will evaluate the fingerprint retrieval methods and techniques. Earlier algorithms used features of level-1 and features of level-2 fingerprint for indexing and retrieval, but after the advent of various features of level-3 such as pore detection, many algorithms have been developed which use pore-based level-3 features. Algorithms based on level-1 and level-2 features have proven their effectiveness for a long period of time, however, there are certain drawbacks. Earlier retrieval algorithms had high calculation consumption and became less effective while dealing with partial images of the fingerprint. At the same time, fingerprint images with high resolution using level-3 features have shown an ability to overcome earlier drawbacks by a significant margin. Subsequently, many algorithms using level-3 pore-based features were developed with improved efficiency, accuracy, and effectiveness than conventional fingerprint retrieval algorithms. In this review, we will evaluate existing fingerprint detection and retrieval techniques and some recent research works using the most latest technology and algorithms.

Keywords Fingerprint · Fingerprint images · Level-1 features · Level-2 features · Level-3 features · Ridge · Orientation field · Minutiae · Pores · Fingerprint features · Fingerprint retrieval

1 Introduction

Fingerprint-based verification and identification are widely used and accepted as a method to authenticate a person. Over the period of decades, many methods have been developed to efficiently identify similarities between fingerprints. In the late nineteenth century, the author [1] developed the first classification technique

D. Solanki (✉) · N. Doshi
Pandit Deendayal Energy University, Gandhinagar, Gujarat, India
e-mail: deepsolanki5799@gmail.com

N. Doshi
e-mail: nishant.doshi@sot.pdpu.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_29

for studying fingerprints. After that, authors [2] while working for British-India's police department, developed refined classification in more classes which still is being used by various agencies across the world [3]. In the last few decades, law-enforcement departments, corporations, and civilian sectors have benefited tremendously from fingerprint identification through an automatic process [4]. Authors of [5–10] proposed fingerprint detection and retrieval based on level-1 and level-2 features consisting of ridge orientation, pattern-based, and minutiae-based algorithms. However, fingerprint encoding is largely done consisting of unordered minutiae sets where the number of minutiae is also different for each encoding. These inconsistencies make fingerprint retrieval more time-consuming [11]. Minutiae are extracted using a number of processes such as fingerprint segmentation, enhancement, thinning, and binarization. Traditional techniques are still inefficient and inaccurate for vast-scale real-life applications where fingerprints are of low-quality or partial in nature [12].

For the purpose of retrieval of fingerprint images, features of level-1 such as singular point and ridge orientation and features of level-2 such as minutiae features were used. Now with the advent of algorithms based on high-resolution pore-based level-3 features, the effectiveness, accuracy, and efficiency of the fingerprint retrieval process are significantly improved. This algorithm becomes even more superior to their previous version while comparing partial fingerprint images. Recently, pore-based retrieval algorithms have shown the ability to provide high-accuracy fingerprint verification. These algorithms use level-3 pore features and are more faster, more effective, and more accurate than earlier minutiae-based, pattern-based, and ridge-based algorithms [13–15]. Authors of [16–20] have proposed methods of fingerprint verification using pore features of fingerprints. Their analysis claims to provide better results with high accuracy and a faster retrieval rate.

However, pore-based retrieval has to deal with computation-consumption problems as the amount of pores in the image of the fingerprint is a lot higher than the number of minutiae-based features. Another pore-based verification algorithm was developed to get rid of this problem which used cluster-based indexing for the retrieval of query images in [17, 18]. However, the disadvantage of that algorithm is that it is dependent highly on the precision of clusters used for indexing. To overcome these challenges authors of [16] proposed a new method using the UNSW algorithm and random walker graph algorithm.

1.1 Organization of the Paper

In Sect. 1, we gave a brief idea about fingerprints and detection and retrieval algorithms. In Sect. 2, we have given the literature survey of the existing algorithms and methods used for the retrieval of fingerprint images. In Sect. 3, we have given the analysis of the algorithms and methods. Section 4 consists of the conclusion and then references are at the end.

2 Literature Survey

The subsequent is a framework of the literature review. The entire literature is divided into three parts, first part gives brief info about the features used for fingerprint retrieval algorithms, the second part gives information about related works, and the third part is about a detailed evaluation of pore-based algorithm based on indexing of fingerprint pores of paper [16].

2.1 Features Used for Fingerprint Retrieval

This section contains brief information about various features which are being used for the purpose of fingerprint scanning, detection, retrieval, and verification (Fig. 1).

Ridge orientation field:

Ridge orientation field is a map of direction of ridges present in the image of a fingerprint [22]. Authors [23] presented an algorithm based on local ridge orientation and frequency information. It is used as one of the feature for recognition of fingerprint. Techniques of image processing like gradient-based or wavelet-based methods are applied to estimate the orientation field. Orientation field is used to differentiate fingerprints with similar minutiae pattern but different ridge orientation, which improves the accuracy of matching process.

Minutiae:

These are unique features of the fingerprint consisting primarily of two features—ridge ending and ridge bifurcation, where ridge ending is the end of a ridge in fingerprint and ridge bifurcation is a part where the ridge divides into two or more ridge branches [3]. Every fingerprint has these unique points known as minutiae which are used for fingerprint recognition algorithms. Some of the minutiae features are ridge ending, ridge bifurcation, ridge orientation, and ridge curvature. Algorithms [5–10] are mainly based on minutiae features.

SIFT features:

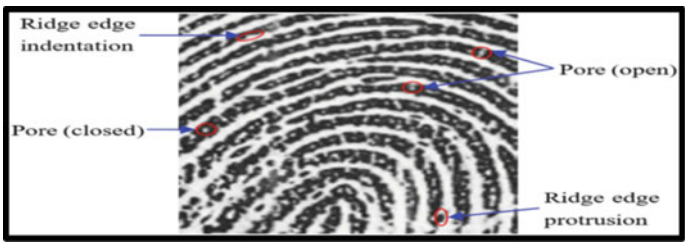


Fig. 1 Basic fingerprint features [21]

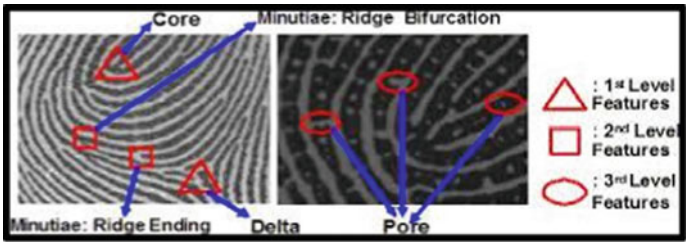


Fig. 2 Level 1, 2 and 3 features of fingerprint [3]

Scale-invariant feature transform (SIFT) feature is used in applications of image recognition. Authors in [24] have used SIFT for fingerprint indexing. It has two main components: keypoints and descriptors, keypoints are unique features in an image which is detected by SIFT and are invariant to scale and rotation and descriptors maintain detailed description about appearance of those keypoints in the image.

Pattern-based features:

These algorithms focus on the overall pattern of the fingerprint, rather than individual minutiae points. Authors [25] proposed an algorithm using structure and pattern of fingerprints.

Pore features:

Pores present in the human fingerprint are unique to each fingerprint. One of the first attempts to use pores as the distinct feature for fingerprint identification was proposed by [26]. Fingerprint recognition based on pores is still a developing area. High-resolution images and sensors are required to use pores for fingerprint recognition and retrieval. These are way more efficient and faster than other features. They are also effective for partial fingerprint images. In recent times researchers of paper [16–20] have proposed algorithms that use pores for fingerprint retrieval (Fig. 2).

2.2 Related Works

This section mentions different methods and algorithms which are being used for fingerprint retrieval.

Fingerprint indexing:

Features that are used by different algorithms for the purpose of indexing are ridge orientation field [3, 23], SIFT features [24], minutiae [5–10] and pores [16–20, 26]. Out of all other algorithms, minutiae-based algorithms are considered the most reliable which can be classified as indexing algorithms based on geometric features

and minutiae descriptors. Also, an algorithm using pores and minutiae with high-resolution fingerprint indexing which is based on Delaunay triangulation with its drawback and solution of a small world algorithm is also present.

Minutiae-Based algorithms:

Minutiae-based algorithms are most widely used for fingerprint detection. Different researchers have developed different mechanisms to utilize minutiae for fingerprint detection. Fingerprint detection and retrieval algorithms [5–10] developed by researchers primarily use minutiae features. In the work of paper [5], authors proposed fingerprint indexing using minutiae quadruplets. Authors in paper [7] proposed combining traditional minutiae matching with their own developed density map and results were better than fingerprint matching algorithms who used single minutiae scheme. In paper [8], authors use combination of multiple algorithms to extract minutiae from binary images that were skeletonized. Authors of paper [10] proposed a method based on adaptive tracing of fingerprint ridges with gray-level and they used images consisting piece-wise lines of various sizes.

Pore-Based algorithms:

Algorithms using pore-based features are new and still in the developing stage. Still, these new algorithms are outperforming conventional algorithms in terms of time consumption, accuracy, and efficiency. Recently developed algorithms such as [16–20] are based on pore features. For pore identification [20, 27–30] have proposed algorithms for pore extraction.

One of the most first attempts to use pore features was proposed in the paper [26]. They developed a feature-matching algorithm using pores and ridge contours to match fingerprint images with 1000 DPI quality. Authors of [17] proposed pore extraction from high-resolution images using a method in which pores are filtered in a dynamic way. In paper [18], authors proposed an algorithm that uses Delaunay triangulation for pore extraction. Authors of [19] proposed to first identify the correspondence of pores on local features and then applies RANSAC in the refinement stage.

Navigable small world (NSW) algorithms:

k-NN algorithm is widely used and popular for creating proximity graphs but *k*-NN gives a weak performance in clusters since it only uses local connections. Navigable small world proposed in [31] provides the solution for this problem. To solve the issues in the *k*-NN graph, NSW uses long-range edges in the network. Using the same idea, authors in [32] proposed hierarchical navigable small world (HNSW) which creates the structure of multiple layers.

2.3 Pore-Based Algorithm

This section contains a detailed evaluation of one of the pore-based algorithms to understand how an algorithm can utilize level-3 pore features for fingerprint retrieval. In the proposed method by the authors of the paper [16], two main steps of the algorithm are described.

Pore enrollment and indexing:

Here fingerprints are scanned to identify pores from the image for the purpose of enrollment and searching. DAMP algorithm [27] and DAISY [33] descriptor are used for identifying, extracting, and creating a local descriptor for each pore in the fingerprint image. Then a binary code is generated using the algorithm of iterative quantization (ITQ) [34]. The same process used for enrollment of the image can be used for query image to generate its binary codes. The binary codes of each pore guide to multiple nearest neighbors in a given indexing space. To find nearest neighbors randomly generated, long-term edges are used in a hierarchical navigable small world (HNSW) which cannot provide effective high speed while dealing with large databases. And for that purpose uniform navigable small word (UNSW) is preferred over HNSW. UNSW uses spatial distribution for the creation of the top layer instead of creating it randomly. Further, search-layer algorithm is used to identify the nearest neighbors of the query pore in all layers.

Nearest Neighbor Refinement:

Every query pore and its nearest neighbors are in one-to-many correspondence. Although many refinement algorithms already exist still, they are not feasible due to geometric constraints. This paper proposes a random walk algorithm (RANSAC) for the refinement of the pores correspondence. After the refinement process, we use the remaining pore correspondences to compute the resemblance between enrolled fingerprint images and the query fingerprint image. Final resemblances are sorted, are in descending order, and finally, top resemblances are retrieval results for the query image.

Experiment Analysis and Results Method of Research [16]:

This part shows data used for experiments and performance of different methods.

Data Sets and Experiment Setting:

Two databases DBI (partial fingerprints) and DBII (full-size fingerprints) are used to determine the results and performance. These databases contain 1480 images of 140 fingers with 1200 dpi resolution. Many are used for the enrollment of fingerprints, while the remaining ones are used as query images to carry out experiments.

Indexing Performance of UNSW:

It depicts that while dealing with a high number of pores, it takes a low percentage of pores in the top layer which gives better results than the HNSW algorithm. According

to their results, HNSW performs better when number of enrolled pores are less, but with the increment of number of enrolled pores speed of UNSW is better than HNSW.

Refinement Result of the Graph Comparison Algorithm:

To eliminate the false correspondences, a random walk algorithm is used which uses the following equation to derive iterative difference: $\Delta x(n) = \Sigma(x(n+1) - x(n))$. Then for graph refinement, random walk algorithm with affinity preservation is used and finally graph comparison is applied to find final retrieval of the query image. They have demonstrated an image where, the image in the middle is a query image while the other are eight enrolled images found after completion of the retrieval algorithm.

Indexing Performance of the Retrieval Algorithm:

In the analysis part, they have given final comparison of their proposed method and other pore-based retrieval algorithms. From the results they derive that proposed method outperforms earlier cluster-based algorithms. They also derive time cost of HNSW is 5.3 ms on billion image features, while time cost of UNSW is 4.8 ms on billion image features which shows better efficiency of UNSW compared with HNSW.

3 Analysis

In this section, we will compare which algorithms are supporting which features for fingerprint retrieval. We will also analyze database details of pore-based algorithms and compare them with each other.

From the Table 1, we can observe that geometric and minutiae feature-based algorithms are mostly using combination of singular points, ridge orientation, and minutiae, while pore-based algorithms use combination of minutiae and pore features. Pore-based algorithms also provide partial fingerprint detection effectively which is very challenging for other algorithms.

Table 1 Different features used by different algorithms

Algorithm type	Papers	Singular points	Minutiae	Ridge orientation	Pore features	Partial fingerprint detection
Geometric feature-based algorithms	[23, 25]	yes	yes	yes	no	no
Minutiae-based algorithms	[5–10]	yes	yes	yes	no	no
Pore-based algorithms	[16–20]	no	yes	no	yes	yes

Table 2 Details of database used by pore-based algorithms

Pore-based algorithm	Database	Number of fingerprint images in dataset	Number of fingers used for fingerprints	Size of image (height × width) (pixels)	Image resolution
[16]	DBI	1480	140	320 × 240	1200 dpi
	DBII	1480	140	640 × 480	1200 dpi
[17]	DBI	1480	148	320 × 240	1200 dpi
	DBII	1480	148	640 × 480	1200 dpi
	IITI-HRFP	1984	248	320 × 240	1000 dpi
	IITI-HRF	1984	248	1000 × 1000	1000 dpi
[18]	DBII	1480	148	640 × 480	1200 dpi
[19]	Their own new database	1480	148	320 × 240	1200 dpi
[20]	DB touch-based	Not available	Not available	320 × 240	1200 dpi
	DB touchless	44	22	250 × 250	1000 dpi
	DB latent	36	22	320 × 240	1200 dpi

In Table 2, details of databases used by pore-based algorithms are given. Here, we can see that fingerprint images are 1000 dpi and 1200 dpi high-resolution images, which are normally required to scan and extract tiny pores. DBI are partial fingerprint images, while DBII are full-scale fingerprint images. Along with some DBI and DBII, other databases are developed by respective authors by their own efforts.

The limitation of our study is that the study may not provide specific guidance for implementing fingerprint detection and retrieval algorithms, as it mainly focuses on providing a high-level overview of different features and algorithms. Also, our work is primarily focused on the technical aspects of the fingerprint features and algorithms, ethical and privacy concerns related to fingerprint detection are not discussed in this study. However, the study is beneficial in understanding the advantages and disadvantages of different algorithms based on their features. It explores the effectiveness of different level-1, level-2, and level-3 features for fingerprint retrieval and also reviews recent research works and the latest technology used in fingerprint detection and retrieval.

4 Conclusion

We evaluated various fingerprint algorithms and fingerprint features used by those fingerprint algorithms. Automated fingerprint detection has benefited civilian and commercial sectors along with law-enforcement. Minutiae-based algorithms are most widely used for fingerprint detection. Earlier level-1 and level-2 features-based

state-of-the-art algorithms have remained useful for many years, but new pore-based algorithms are providing even better results and solutions such as partial fingerprint detection. We also analyzed fingerprint features used by different algorithms and database details used by some pore-based algorithms. Pore-based algorithms are fast, accurate, and efficient. However, pore-based algorithms cannot provide accurate results when fingerprint images are of low quality which makes pore identification difficult. Fingerprint detection using pore features is still a new and developing area where there is a huge scope for progress and research. We hope our work provides useful information about fingerprint features and algorithms that can be helpful for future studies and research purposes.

Author Contributions We express our gratitude to the reviewers and organizers of the conference for their insightful feedback. Deep Solanki is identified as the first author of the paper, and both Deep Solanki and Dr. Nishant Doshi are the corresponding authors.

References

1. Francis Galton (1892) Finger prints. London McMilan
2. Henry ER (1900) Classification and uses of fingerprints. London Routledge
3. Zaeri N (2011) Minutiae-based fingerprint extraction and recognition
4. Maltoni D, Maio D, Jain AK, Feng J (2022) Handbook of Fingerprint Recognition, Cham. Springer, Switzerland
5. Iloanusi O, Gyaourova A, Ross A (Jun 2011) Indexing fingerprints using minutiae quadruplets. In: Proceedings CVPR workshops, pp 127–133
6. Ross A, Mukherjee R (Apr 2007) Augmenting ridge curves with minutiae triplets for fingerprint indexing. In: Proceedings SPIE, vol 6539, Art. no. 65390C
7. Wan D, Zhou J (2006) Fingerprint recognition using model-based density map. IEEE Trans Image Process 15(6):1690–1696
8. Farina A, Kovács-Vajna ZM, Leone A (1999) Fingerprint minutiae extraction from skeletonized binary images. Pattern Recognit 32(5):877–889
9. Yang X, Feng J, Zhou J (2014) Localized dictionaries based orientation field estimation for latent fingerprints. IEEE Trans Pattern Anal Mach Intell 36(5):955–969
10. Jiang X, Yau WY, Ser W (May 2001) Detecting the fingerprint minutiae by adaptive tracing the gray-level ridge. Pattern recognition 34(5):999–1013
11. Duan Y, Feng J, Lu J, Zhou J (2022) Estimating fingerprint pose via dense voting
12. Feng Y, Kumar A (2023) Detecting locally, patching globally: an end-to-end framework for high speed and accurate detection of fingerprint minutiae
13. Xu Y, Lu G, Lu Y, Liu F, Zhang D (2019) Fingerprint pore comparison using local features and spatial relations. IEEE Trans Circuits Syst Video Technol 29(10):2927–2940
14. Liu F, Zhao Q, Zhang D (2011) A novel hierarchical fingerprint matching approach. Pattern Recognit 44(8):1604–1613
15. Liu F, Zhao Q, Zhang L, Zhang D (Aug 2010) Fingerprint pore matching based on sparse representation. In: Proceedings 20th international conference pattern recognition, pp 1630–1633
16. Xu Y, Lu Y, Chen F, Lu G, Zhang D (2022) High resolution fingerprint retrieval based on pore indexing and graph comparison
17. Anand V, Kanhangad V (2020) Pore-based indexing for fingerprints acquired using high-resolution sensors. Pattern Anal Appl 23(1):429–441
18. Anand V, Kanhangad V (Feb 2017) Pore based indexing for high-resolution fingerprints. In: Proceedings IEEE international conference identity, security behavior analysis (ISBA), pp 1–6

19. Zhao Q, Zhang L, Zhang D, Luo N (Jul 2009) Direct pore matching for fingerprint recognition. In: Proceedings 3rd international conference ICB. Alghero, Italy, pp 597–606
20. Labati RD, Genovese A, Muñoz E, Piuri V, Scotti F (2018) A novel pore extraction method for heterogeneous fingerprint images using convolutional neural networks. Pattern Recognit Lett 113:58–66
21. Liu F, Zhao Q, Zhang D (2020) advanced fingerprint recognition: from 3D shape to ridge detail
22. El Bovik (2009) The essential guide to image processing, chapter-23. Jain A, Pankanti S Fingerprint Recognition
23. Hong L, Wan Y, Jain A (1998) Fingerprint image enhancement: algorithm and performance evaluation. IEEE Trans Pattern Anal Mach Intell 20(8):777–789
24. Gupta P, Tiwari K, Arora G (2019) Fingerprint indexing schemes—a survey. Neurocomputing 335:352–365
25. Candela GT, Chellappa R (Apr 1993) Comparative performance of classification methods for fingerprints. NIST Technical Report NISTIR 5163
26. Jain A, Chen Y, Demirkus M (2006) Pores and ridges: fingerprint matching using level 3 features. In: Proceedings 18th international conference pattern recognition, pp 477–480
27. Zhao Q, Zhang D, Zhang L, Luo N (2010) Adaptive fingerprint pore modeling and extraction. Pattern Recognit 43(8):2833–2844
28. de Paula Lemes R, Segundo MP, Bellon ORP, Silva L (Aug 2014) Dynamic pore filtering for keypoint detection applied to newborn authentication. In: Proceedings of 22nd international conference pattern recognition, pp 1698–1703
29. Wang H, Yang X, Ma L, Liang R (2017) Fingerprint pore extraction using U-Net based fully convolutional network. In: Proceedings of Chinese conference biometric recognition Shenzhen. Springer, China, pp 279–287
30. Jang H-U, Kim D, Mun S-M, Choi S, Lee H-K (2017) DeepPore: fingerprint pore extraction using deep convolutional neural networks. IEEE Signal Process Lett 24(12):1808–1812
31. Malkov Y, Ponomarenko A, Logvinov A, Krylov V (2014) Approximate nearest neighbor algorithm based on navigable small world graphs. Inf Syst 45:61–68
32. Malkov YA, Yashunin DA (2020) Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. IEEE Trans Pattern Anal Mach Intell 42(4):824–836
33. Tola E, Lepetit V, Fua P (2010) DAISY: an efficient dense descriptor applied to wide-baseline stereo. IEEE Trans Pattern Anal Mach Intell 32(5):815–830
34. Gong Y, Lazebnik S, Gordo A, Perronnin F (2013) Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval. IEEE Trans Pattern Anal Mach Intell 35(12):2916–2929

Survey of Explainable AI Techniques: A Case Study of Healthcare

Tanaya Patil and Sandhya Arora

Abstract With its unparalleled predictive capabilities and decision-making power, artificial intelligence (AI) has transformed countless disciplines. Nevertheless, as AI technology advances in complexity and ubiquity, so does the requirement for transparency and interpretability increase commensurately. The opaqueness of AI models makes them difficult to comprehend, which leads to mounting concerns regarding their application. To address this issue effectively, Explainable Artificial Intelligence (XAI) has emerged as an attempt to demystify how machine learning model decisions are made possible by offering comprehensible explanations for opacities caused by variations in input features or noise in data collection processes using methods such as LIME or SHAP especially vital in domains such as health care where ethical considerations outweigh those faced in other areas. This study provides an overview of XAI techniques, a case study that talks about the importance of using techniques such as LIME and SHAP in a healthcare organization to predict whether or not the patient is likely to get a heart disease or not, why and how the outcomes have been obtained. In the concluding section, the paper elaborates on the challenges facing XAI and outlines prospective future directions for research in this domain.

Keywords XAI · Transparency · Interpretability · LIME · SHAP

1 Introduction

Artificial Intelligence (AI) has made considerable progress over recent years [1], significantly contributing to numerous fields such as healthcare, finance, and criminal justice. Nevertheless, doubts linger on the correctness and accountability aspects applicable to AI models that are lacking in transparency or interpretability resulting

T. Patil (✉) · S. Arora
Cummins College of Engineering for Women, Pune, India
e-mail: tanaya.patil@cumminscollege.in

S. Arora
e-mail: sandhya.arora@cumminscollege.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_30

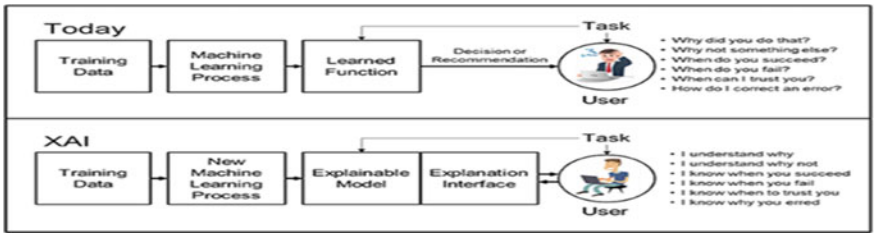


Fig. 1 Explainable AI (XAI) versus conventional AI

from advancements made thus far. It is apparent that solutions must be found for these issues; hence Explainable Artificial Intelligence (XAI) emerged aiming at enhancing translucency amidst developing meaningful cognition options for implementing improved machine learning technologies.

In the illustration above, XAI solves the "wh" problem that is missing in traditional AI in a new way. Since it is also important to describe how to achieve the answer (that is, the answer to the "WH" question), XAI has a wide range of applications, including medical care, defence, law and order. Both colleges and industries focus on XAI research. There is a lot to be done in order to completely utilize XAI. In recent years, various applications have successfully used artificial intelligence. The model of artificial intelligence is becoming more and more complicated. Their decision-making process is usually opaque, which makes it difficult to understand their arguments. The lack of transparency and explanatory in the AI model poses a lot of concerns. It is incompetent to explain the results of AI decision thereby causing prejudice and unfairness of individuals and society (Fig. 1).

The field of Explainable AI (XAI) is expanding its horizons by exploring how artificial intelligence models arrive at decisions and the reasoning behind them. XAI enhances transparency and interpretability in AI models, empowering users to comprehend the decision-making process, identify probable biases, and enhance the system's trustworthiness. The advantages provided by XAI span various domains [2].

2 Literature Review on XAI Techniques

XAI technology aims to provide AI model transparency and interpretability, and understand how people make decisions. These technologies [3] can be divided into two main groups: post-HOC and transparent methods (Fig. 2).

Post-hoc methods are the methods that involve the analysis of an already trained AI model to recognize the factors and features that the model took into consideration while making its decisions.

Model agnostic techniques do not require any comprehension of the inner functions of the inspected model and can be applied to any model. These methods include

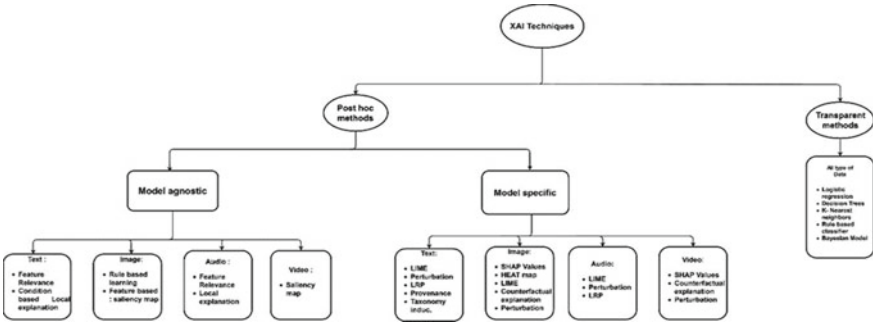


Fig. 2 XAI techniques

feature relevance and condition-based local explanation for text, rule-based learning, and feature-based: saliency map for image, feature relevance and local explanation for audio and saliency map for audio.

- *Feature relevance and condition-based local explanation for text*: Identify the most important words or phrases in a text input that influenced the model's decision.
- *Rule-based learning for text*: Extract rules from the model that explain the decision-making process in human-readable statements.
- *Feature-based saliency map for image*: Highlight the regions of an input image that had the greatest influence on the model's decision by calculating gradients of the model's output with respect to the input image pixels.
- *Feature relevance and local explanation for audio*: Identify the most important features or input variables in an audio signal that influenced the model's decision.
- *Saliency map for video*: Highlight the most important regions in a video that influenced the model's decision. This can be done by applying the same approach used for images to each frame of the video and combining the results.

On the other hand, as the name suggests, model-specific methods require knowledge of the inner functions of the model and should be tailor-made by the specific model. These methods include LIME, Perturbation, LRP, Provenance, Taxonomy induc for text, SHAP values, Heat map, LIME, Counterfactual explanation and Perturbation for Image, LIME, Perturbation and LRP for Audio, SHAP values, Counterfactual explanation, and Perturbation for Video.

- *Local Interpretable Model-Agnostic Explanations (LIME)*: Explains the behaviour of an AI model at the local level by approximating the model with a simpler model that is easier to interpret.
- *Perturbation*: Involves changing the input data in specific ways to determine how the model's output changes. This can be used to identify the most important features in the input data.
- *Layer-wise Relevance Propagation (LRP)*: Uses backward propagation to calculate the relevance of each input feature to the model's output.

- *Provenance*: Tracks the origin and processing history of data to provide insight into the model's decision-making process.
- *Taxonomy induction for text*: Organizes input text into a hierarchy of categories to help explain the model's decision.
- *SHAP values*: Calculates the contribution of each feature to the model's output using a game-theoretic approach.
- *Heat map*: Displays the regions of an image that were most important to the model's decision by assigning each pixel a weight based on its contribution to the model's output.
- *Counterfactual explanation*: Generates new input data that would have resulted in a different model output. This can be used to help identify the most important features in the input data.
- *LIME for audio*: Similar approach as LIME for text to explain the behaviour of an AI model at the local level.
- *Perturbation for audio*: Changes in input data in specific ways to determine how the model's output changes.
- *SHAP values, Counterfactual explanation, and Perturbation for video*: Similar to the approaches used for images, but applied to each frame of the video and combined to explain the model's decision over time [4] (Table 1).

3 Case Study in Healthcare Organization

Healthcare centres uses advanced machine learning model to predict the patient's results on the basis of their medical reports. The implemented algorithms are based on a massive dataset of patient records and has exhibited impressive results in terms of accuracy, surpassing traditional methodologies by leaps. In this scenario, utilizing XAI is crucial due to the machine learning model's "black box" nature, leaving doctors unable to comprehend the reasoning behind the predictions. This lack of transparency makes it difficult for doctors to trust the model and use it to inform patient care decisions. However, implementing XAI techniques can increase transparency and explainability thereby allowing for better understanding and trust among medical professionals, ultimately enhancing patient care. The organization employed various XAI methods in enhancing the transparency and interpretability of the machine learning model. These comprised of LIME, SHAP and feature relevance. Feature relevance analysis was used to find out the most important features contributing to the predictions made by the model. SHAP tech was used to give an explanation of the model's predictions for independent patients. LIME was used to provide explanations that could be comprehended by humans for the model's predictions.

Table 1 Literature review for XAI Techniques

Author	Technique	Result	Research gap/discussion
Ribeiro et al. [5]	LIME for explaining predictions of models such as decision trees, SVMs, neural networks	Accurate results with easy interpretability and clarifications. Demonstration for wrong predictions	Not for complex models, limited to domains, the quality of explanation is not good
Lapuschkin et al. [6]	LRP technique as model in a toolbox with visualization of relevance scores	Results affirms the efficacy and flexibility of the LRP toolbox	Limited scalability, interpretability, and generalizability and support for certain types of neural network architectures only
Janzing et al. [7]	Uses do-calculus and counterfactual reasoning technique by examining pathways and controlling for confounding variables	Method estimates the effect of each feature on the output, finding relevant features and reducing errors	Assumes a simplified causal model, features are independent and affect the output through a single causal path. No talk about empirical evaluations on real-world datasets
Kuo et al. [8]	Modified CNN for predictions by adding an interpretable layer	Based on feedforward design. Given approach is better as model can identify features that contribute to the prediction thereby achieving good performance	The relevance scores found out may be difficult to understand by non-experts. Not clear how well it will give perform on new and varied datasets and custom layers
Adugna et al. [9]	The performance of two machine learning classifiers-RF and SVM is compared for land cover mapping using coarse resolution satellite images	SVM performed better in terms of producer's accuracy for some classes. However, the study suggests that RF is a more suitable classifier for land cover	Analyses a limited sample size (FY-3C satellite). Explanation for "why" the difference in performance between classifiers occurs isn't addressed. Detecting if these differences are statistically significant or merely happenstance becomes a daunting task
Akita et al. [10]	Bayesian approximation technique is used to compute the feature importance scores of the input variables for a given GCN model	Identifying important features and detecting errors in the GCN model	Computationally high, Interpretability of feature importance scores is challenging. Sensitivity to prior specification of parameters could be highly affected by the distribution choices

(continued)

Table 1 (continued)

Author	Technique	Result	Research gap/discussion
Ribeiro et al. [11]	Anchors is used for generating high-precision explanations for predictions of black box models by identifying a minimal set of features	When performed on several datasets and black box models Anchors results in better precision and interpretability than other XAI techniques	Limited to binary classification, may not be adequate for textual or image-based data. Work sensitive to the choice of the threshold which can greatly impact the quality of the explanation provided
Lundberg and Lee [12]	SHAP for interpreting the predictions of any model. Authors use LIME in combination with SHAP values to provide local explanations for black box models	SHAP values-provide accurate feature importance scores for predictions, LIME-used to generate simple, intuitive explanations for individual predictions	Lacks applications to different types of models. Need for developing more efficient computation methods. Should explore the integration of SHAP values with other XAI techniques, for comprehensive understanding of model behaviour
Ancona et al. [13]	Proposed algorithm SHAP for feature importance scores in deep neural networks. Provides a unified framework for varied feature attribution methods	Outperforms existing methods in terms of accuracy and computational efficiency. Theoretical guarantees for the correctness and convergence of the algorithm	Algorithm may not be scalable to models such as recurrent neural networks and transformers. Although the approach is successful in certain models like tree-based and neural networks, the same cannot be said for other models
Rahab et al. [14]	Uses decision trees for sentiment analysis of Arabic social media content	Including decision trees improved the accuracy. Highlights the potential of decision trees as an understandable and effective model	Constrained by the quantity and calibre of data available. The intricacies of natural language may stump the algorithm as it depends on pre-existing strictures
Binder et al. [15]	Layer-wise relevance propagation (LRP), DNN	Visual representation of internal DNNs layers, accurate and detailed explanation of predictions	Fixed hyperparameters used in model, not the generalized model, performance may decrease, explainability is also not good

(continued)

Table 1 (continued)

Author	Technique	Result	Research gap/discussion
Bach et al. [16]	Focus deep neural networks (DNNs) and LRP on image data	LRP provides explanations to better understand the behaviour of DNNs	Challenge in scaling up to handle large data, complex attribution maps, lack of quality
Sundararajan et al. [17]	Uses integrated gradients attribution methods, LRP and Grad-CAM	Accuracy and interpretability is good on several datasets	Not applicable to other attribution methods, predictions are sensitive to changes in the input features

4 Implementation of Xai Techniques

A machine learning model which is trained on a heart disease dataset is able to flawlessly predict whether or not a person is in danger of suffering from a heart disease based on factors such as age, gender, blood pressure, cholesterol levels, and other health metrics. With the help of LIME and SHAP techniques, we can identify which factors are most dominant in the model's decision-making mechanism and how much importance they hold in relation to one another. With the help of this information, we can gain a better understanding of the hidden factors that give in to the risk and thereby take measures to reduce the risk of heart disease (Fig. 3).

4.1 Proposed XAI Techniques

LIME and SHAP techniques have been implemented on heart disease dataset to understand the decision-making process of these models.

(1) **SHAP**

By using SHAP to interpret machine learning models, we can gain invaluable insights on heart disease risk by attributing a score to each feature indicating its contribution to the prediction. This data can be used to identify biases within the model and make informed medical decisions.

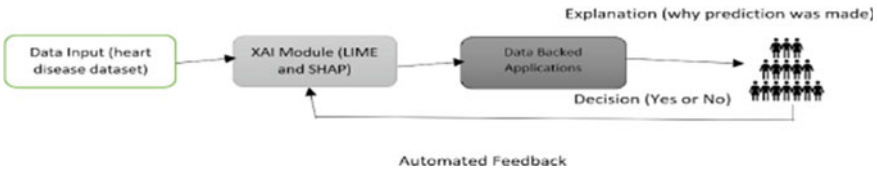


Fig. 3 Proposed model flow

Figure 4a shows the list of important features—from most significant to the least significant one (top to bottom). Thalassemia is the one with the most predictive power according to the model and the number of major blood vessels (bloodVessels) ranks second in predictive power. In Fig. 4b, each dot represents the feature value for an individual data instance. Colours tell whether the feature has a high value (red) or a low one (blue).The X-axis shows the positive or negative contribution to the predicted output. When we apply those analyses to the features, it was observed that majority of the high values (red dots) have a positive contribution to the predicted output, i.e. if the thalassemia quantity for an individual data instance is high, its chance of having a positive outcome (being diagnosed with a heart disease) increases strongly. The higher the number of major blood vessels involved for the patient; he/she is most likely to be diagnosed with a heart disease. Since age and blood pressure features seem to be confusing for the analysis, we can use the dependence plot to have more granular information. Figure 4c is a dependency plot for age and 4D is a plot for blood pressure from which we can clearly see that patients below 45 years old have a low risk of being diagnosed with any heart disease whereas patients having a blood pressure of more than 120 are likely to be diagnosed with a heart disease.

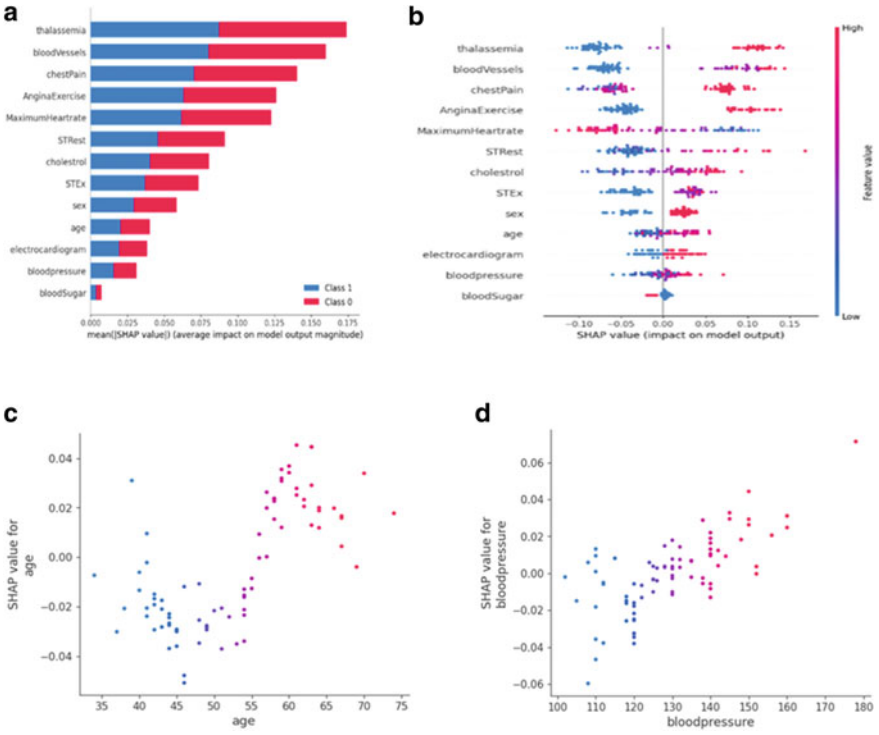


Fig. 4 Results of SHAP on performance on heart disease dataset

(2) **LIME**

The LIME framework interprets machine learning models by creating a simplified alternative that highlights the critical predictors for predicting a specific patient's illness. To enhance decision-making, we can apply LIME to diagnose heart disease in patients and identify how the machine learning model generated the forecast. These insights can provide better medical care and optimize the model's overall performance.

LIME provides additional assistance by helping doctors prioritize risk factors in a patient's treatment plan and highlighting areas where the model may be less accurate. With its unpredictably diverse vocabulary and varied writing style, LIME is a powerful tool for medical professionals and researchers alike to better understand heart disease datasets. From Fig. 5 we observe that the model predicts a particular patient to not suffer from heart disease with 92% of confidence. The prediction is explained by the fact that thalassemia (a blood disorder) is higher than 1. The value indicates that the blood flow is normal. In addition, STRest (exercise-induced ST-depression in relation to rest) is less than 1.8. These two factors contributed significantly to the prediction made by the model. In the model, the patient's blood pressure and cholesterol were the only two values that indicated that a heart disorder was possible with a probability of 8%. On the right, we can see the real value of the features for the patient. According to the value for chest pain, the patient has experienced non-anginal chest pain, which does not result from reduced blood flow to the heart. This indicates that you will not suffer a heart attack. Considering another example, the cholesterol value for a normal person is less than 130 mg/dl for low-density lipoprotein and above 40 mg/Dl for high-protein lipoprotein. However, the cholesterol level for this patient 239 mg/dl. This result indicates that the person has abnormal cholesterol levels which could lead to heart disease in the near future.

With the use of XAI techniques, the health centre could make the machine learning algorithm more transparent and explainable. This helped doctors derive how the model made the predictions about a person suffering from a heart disease or not, which in turn facilitated their trust in the model created.

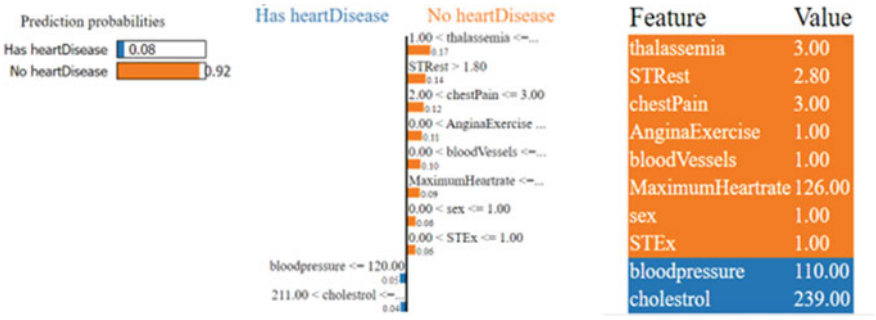


Fig. 5 Result of prediction using LIME framework

5 Challenges and Future Directions

The advancements in XAI are undeniable, yet it has numerous daunting challenges. One such challenge is the absence of a standardized evaluation framework for XAI methods. Balancing between Interpretability and Performance is a challenge. XAI techniques have the potential to shed light on decision-making processes. However, it is crucial to take into account the particular context in which decisions are being made as what works in one scenario might not apply in another. Many XAI techniques are created without considering the user's requirements, leading to technical-focused solutions that may not meet user needs. In order to make the techniques more efficient and effective in resource-limited environments like mobile devices and IoT devices, there must be an emphasis on developing lightweight and efficient XAI techniques that can be deployed without compromising performance. XAI approaches frequently necessitates a substantial number of computational resources, rendering them infeasible in situations with limited resources. The use of XAI techniques has raised significant concerns about privacy and security as it often requires access to vast amounts of sensitive data.

Despite the challenges, XAI has a promising future, and several research directions are being pursued to address these challenges. One research direction is the development of hybrid models that combine the strengths of rule-based, model-based, and post-hoc methods. Hybrid models, with their ability to furnish vastly intricate and explanations about the decision-making process, hold the potential of building robust and trustworthy AI systems. Developing XAI techniques to comprehend the decision-making process of complex deep neural networks is a mighty challenge. These networks are omnipresent in various domains, but the insignificance of their internal workings clarifies their effectiveness. Nonetheless, another research route is to unravel these mysteries using unpredictable XAI techniques. The progress of XAI technology capable of managing dynamic and evolving data is another area of study in XAI. Data in various fields like healthcare and finance is continuously changing and developing. Crucial are XAI methods that can adapt to altering inputs and deliver real-time elucidations of the decision-making procedure.

6 Conclusion

The paper offers two valuable contributions. Firstly, it provides a thorough and wide-ranging exploration of XAI techniques and their applications across domains. This survey will be a valuable resource for researchers, practitioners, and policymakers in understanding the vast potential of XAI. Furthermore, as a testament to the significance of eXplainable Artificial Intelligence in enhancing the transparency and interpretability of AI systems, I have meticulously showcased its implementation through a comprehensive case study that depicts whether a patient is likely to get heart disease or not. The paper's implications are of utmost importance. The ideas and views

presented in the research can facilitate the development of trustworthy, accountable, and fair AI systems. Additionally, the paper focuses on XAI techniques that can assist in the creation of interpretable, transparent, and explainable AI models. This would improve our comprehension of AI systems and advance the decision-making procedure.

References

1. Cockburn IM, Henderson R, Stern S (2018) The impact of artificial intelligence on innovation. In: A chapter in the economics of artificial intelligence: an Agenda. National Bureau of Economic Research, Inc, pp 115–146
2. Palacio S, Lucieri A, Munir M, Hees J, Ahmed S, Dengel A (2021) XAI handbook: towards a unified framework for explainable AI. In: Proceedings of the 2021 IEEE/CVF international conference on computer vision workshops, 21/05/14
3. Schwalbe G, Finzel B (May 2021) A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. Comput Sci Data Min Knowl Discovery
4. Veldhuis MS, Ariëns S, Ypma RJF, Abeel T, Benschop CCG (2022 Jan) Explainable artificial intelligence in forensics: realistic explanations for number of contributor predictions of DNA profiles. Forensic Sci Int Genet 56:102632. https://doi.org/10.1016/j.fsigen.2021.102632
5. Ribeiro MT, Singh S, Guestrin C (2016) Why should i trust you?": explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on KDD '16
6. Lapuschkin S, Binder A, Montavon G, Muller KR, Samek W (2016) The LRP toolbox for artificial neural networks. J Mach Learn Res (JMLR) 17(114)
7. Janzing D, Minorics L, Blöbaum P (2019) Feature relevance quantification in explainable AI: a causal problem. In: International conference on artificial intelligence and statistics, 29 Oct 2019
8. Kuo J, Zhang M, Li S, Duan J, Chen Y (2019) Interpretable convolutional neural networks via feedforward design. J Vis Commun Image Represent 60:346–359
9. Adugna T, Xu W, Fan J (Jan 22) Comparison of random forest and support vector machine classifiers for regional land cover mapping using coarse resolution FY-3C images. Remote Sens
10. Akita H, Nakago K, Komatsu T, Sugawara Y, Maeda SI, Baba Y, Kashima H (2018) BayesGrad: explaining predictions of graph convolutional networks. In: Neural information processing conference, pp 81–92, 17 Nov 2018
11. Ribeiro MT, Singh S, Guestrin C (2018) Anchors: high-precision model-agnostic explanations. AAAI Press
12. Lundberg S, Lee SI (Dec 2017) A unified approach to interpreting model predictions. In: Conference: NIPS
13. Ancona M, Öztireli C, Gross M (2018) Towards better understanding of gradient-based attribution methods for deep neural networks. In: ICLR 2018 conference blind submission
14. Rahab H, Hichem H, Laouid A (June 2019) Improving Arabic sentiment analysis in social media using decision trees. ICCAIS. IEEE Publisher

15. Binder A, Bach S, Montavon G, Muller KR, Samek W (2016) Layer-wise relevance propagation for deep neural network architectures. IEEE ISSPIT, Conference: ICISA 2016. Lecture Notes in EE 376
16. Bach S, Binder A, Montavon G, Klauschen F, Müller KR, Samek W (July 2015) On pixel-wise explanations for non-linear classifier decisions. Public Library of Science
17. Sundararajan M, Taly A, Yan Q (2017) Axiomatic attribution for deep networks. In: 34th international conference on machine learning, PMLR 70:3319–3328

DDoS Attack Prediction System Using Machine Learning Algorithms

P. Kumar and S. Vinodh Kumar

Abstract Distributed denial-of-service (DDoS) assaults have come to be a severe threat to computer networks and systems' confidentiality and integrity, which are essential resources in today's world. The DDoS assault is the main organization-based assault in the field of PC security that influences the objective server's web traffic. DDoS attacks use a variety of devices to flood a network or server with traffic and prevent authorized users from using the service. DDoS attacks may be difficult to detect prior to implementing any mitigation strategies. These attacks make use of restrictions that apply to every arrangement asset, like the framework of the authorized organization's website. The most recent dataset must be used for analysis in order to identify this state of DDoS attacks. Various DDoS attacks are being identified and evaluated for their effectiveness in this section. Any client accessing network services frequently face this serious threat.

Keywords KNN · Decision tree · Multi-layer Perceptron (MLP) · Logistic regression · DDoS

1 Introduction

DDoS attacks are the most common type of cyberattack used by attackers. Because it occurs within dispersed systems, this attack targets numerous systems and a much broader network. Knowledge loss and inadequate counseling pose a serious threat to the network's users the only one. At the beginning of the 2000s, a DDoS assaults was launched against majority of e-commerce websites, including Amazon and other major ones. One of the largest DDoS attacks, containing 2.56 TB of data, was stopped by Google in 2017. Google Cloud Armor security guidelines are now available to assist customers in protecting their data from DDoS attacks and other web-based

P. Kumar (✉) · S. V. Kumar
Rajalakshmi Engineering College, Chennai, India
e-mail: kumar@rajalakshmi.edu.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_31

threats. DDoS attack is a type of cyberattack in which a large number of compromised devices (botnets) flood a target system with traffic or requests, making it unavailable to its intended users. DDoS attacks can cause significant damage to organizations, resulting in loss of revenue, reputation, and customer trust.

To mitigate the impact of DDoS attacks, it is important to detect and mitigate them as soon as possible. Machine learning has emerged as a promising approach to predict and mitigate DDoS attacks. Machine learning algorithms can be trained on large datasets of network traffic to identify patterns and anomalies that indicate a potential DDoS attack. One of the most common types of attacks is still volume-based attacks. Using the open association move speed, this assault sends an enormous volume of data. An amplification that asks to send a lot of traffic to the victim's services causes congestion and slows down responses as a result of this. State exhaustion attacks, also called protocol assaults, attack the victim's system's network layer. This has an effect on the ability of table spaces like firewalls and load balancers to direct requests to the systems of the intended users. The casualty's framework is compromised as such for use. When an attack occurs at the application layer, a large number of requests are sent to the victim's website or service before they are all answered. It may be necessary to request database access or download large datasets. The framework's activity will be dialed back and may try and be quickly secured in the event that 1,000,000 solicitations are made all the while on the objective.

It is impossible to pinpoint the precise IP address from which DDoS attacks originated due to their fluid nature. Its source can be mimicked, and it utilizes a few frameworks (or hubs) from different areas on the web. Sometimes, the assault's source is compromised. This kind of attack is extremely complicated because it comes from multiple sources and requires human intervention. In order to obtain the results, the system supplies the dataset for these well-known protocols to the trained model. In this project, we employ four distinct methodologies for precise prediction of outcomes. We wish to prepare the model utilizing four particular techniques: choice tree, multi-facet perceptron (MLP), and calculated relapse. Every convention or assault can utilize each of the four methodologies; hence, the clearest one is picked for the best results. A DDoS attack prediction system is a machine learning-based solution that analyzes network traffic data to detect patterns and anomalies that may indicate a DDoS attack which is about to occur. The system uses advanced algorithms to learn from historical data and identify indicators of an imminent attack, such as a sudden increase in traffic from multiple sources. The DDoS attack prediction system can be deployed at various points in the network, such as the edge, the core, or specific application servers. It monitors incoming traffic and uses machine learning algorithms to analyze and classify the traffic flow, looking for anomalous patterns that may indicate an attack is underway. Once an attack is detected, the system can take a variety of actions to mitigate the impact of the attack, such as redirecting traffic, filtering out malicious traffic, or blocking the attacking IP addresses. The system can be trained using a variety of supervised and unsupervised machine learning algorithms, such as deep neural networks, decision trees, and clustering algorithms. The system can also incorporate various features such as traffic volume, traffic types, traffic sources, and network topology to improve its accuracy

in detecting and predicting DDoS attacks. Overall, a DDoS attack prediction system using machine learning is a proactive approach to network security that can help organizations to prevent DDoS attacks before they occur, reducing the impact on network performance and ensuring business continuity.

2 Literature Review

First proposed by Beitollahi and Deconinck, the CAPTCHA (Completely Automated Open Turing Test to Differentiate Computers from Humans) method [1] is currently the best defense against spoofing. Distributed transport denial at the software level is offensive. With this method, request–response authentication is performed when someone requests to connect to the server. The main purpose of this connection is to ensure that all responses come from real people and not from automated gimmicks trying to harm the server. Many websites today use CAPTCHAs to protect their servers against program-level operator denial attacks during registration and login. As reported by Ferguson et al., routers no longer use the network intrusion filter [2] to receive packets that they cannot handle. Spoofing reactivations in packages are prevented from entering the community through entry filtering. The interfaces connect the firewalls that make up the community to the intranet and the network. Firewalls can prevent an attacker from impersonating a group in the peer community by dropping all packets with the community's internal sending address and applying an intrusion filter to the network interface. Outbound filtering is a form of filtering used for packets sent from communities accessed by the backend. During outbound filtering, the firewall discards all outgoing addressable packets from the neighboring community. These policies can help prevent DDOS attacks using IP address spoofing against your organization. A rate-limiting mechanism used to limit the number of packets required for a distributed denial-of-service attack is proposed by Paxson et al. To ensure that visitors to criminal sites are not affected, this method is the most effective in slowing down dangerous package charges. Our main goal is to limit fees for devices that no longer send IP addresses. Each combined subset of website visitors has certain attributes. To limit the charge, the router will notify the upstream router whenever it knows it is ahead. If the combo pack meets the payment limit, it cannot be crossed. Otherwise, these packets will be lost because the message is transmitted multiple times to the upstream router. There are other ways to solve the fee cap problem. These policies detect bandwidth attacks by examining the asymmetry between all packets entering and leaving the community. Distributed denial-of-service attacks target servers that are determined not to respond to packets sent to them. Zang et al. [3] proposed a software puzzle from which DDOS attacks can be eliminated. If the consumer resolves the software problem, the server can more effectively respond to the consumer's operator request. It can be created flexibly. The carrier required by the customer can be provided by solving puzzles. This reduces the risk of DDoS attacks by allowing users to solve puzzles in the most efficient way. A new set of hacking rules developed by Priyadharshini and Kuppusamy

has expanded the choice of network clients to attack various device resources [4]. Competitors also try to attract more visitors to their websites than others. Therefore, their goal is to attack the operators of other websites. They frequently visit selected websites and degrade network server media frequently. The software permanently stores permanent tables. This is achieved through enhanced client state and IP address support. When an IP address connects for the first time, it is considered a valid person. Customers 2, 3, and 4 are listed as possible. Mark a specific IP address as the fifth attack. Only five turns are counted in scoring. The application will check if the person wants to increase the time or not. After that, you will no longer have access to the operator providing this website. This IP address cannot be used to access the operator.

An identifier-location separation approach was proposed by Hongbin Luo et al. [5]. The organization hubs are addressed in this manner by the distinguishing proof and area namespaces. This strategy uses a mapping service. Typically, attackers begin by sending more traffic to a system's zombie machine by forwarding packets. I use this identification and location-based strategy, which only gives users a service once they find a place. DDoS attacks and illegal attacks that send packets to a specific computer will therefore be reduced.

Previous DDoS attack detection techniques can be divided into three categories: flow dissymmetry, sudden traffic change, and source IP address-based distribution. Since DDoS attacks would clearly deviate from normal traffic volume flows, the most common techniques [2, 5–7] rely on rapid traffic changes. However, the attacker can send attack packets at various times with various types of flows, packet sizes, and rates that will follow the statistical patterns of regular flow and avoid detection by coordinating a variety of attacking sources. Additionally, the traffic offset is unable to distinguish between normal network congestion flow and DDoS attack flow. The approaches are based on flow is symmetry [8, 9]. However, the normal flow's entry and exit rates differ in other situations, such as audio and video flows. In addition, attackers can simulate normal flow or send an equal number of SYN and FIN/RST packets to send attack packets from multiple attack sources without reducing their attack's effectiveness. The techniques used the dispersion of attacking sources. However, these methods are in sensitive to attacks with a sufficient volume of attack traffic and source IP addresses and are easily disturbed by regular flows. The research above demonstrates that utilizing only one aspect of attack flows is insufficient for accurate and efficient detection. Methods based on numerous aspects of the DDoS attack [10, 11] flow make it easier to detect DDoS attacks. Because of this, the systems that are currently in use are unable to distinguish between legitimate and malicious traffic that originates from the same sources. Finding every bad IP is extremely difficult. A novel model has been proposed in order to categorize and predict DDoS attacks using the existing dataset.

3 Proposed Model

The first step in this process is selecting the dataset to be used. We took kDDcup dataset which contains many types of attacks namely multihop.' 'normal.' 'saint.' 'satan.' 'snmpgetattack."snmpguess."teardrop.' 'udpstorm. There are 42 segments and right around 30,000 lines in this dataset. The selection of your language and tools is the next step. We use Python and a number of free source libraries, such as pandas, numpy, and sklearn, to build this project. After that, a preprocessing approach is utilized to deal with undesirable dataset data.

This preprocessing step removes conflicting and disparate information before filling in the gaps. The information is then normalized over the range −1.0–1.0. Displaying and extracting data are the fourth step. Encryption is used to convert character data to numeric data. The model has separate training and testing datasets. We design and train models with multiplier scaling and hyperparameters to improve model performance. Results are produced when the model reaches an optimal state. His main contribution is to create ideal data usage models and optimize those models. Four algorithms were used in this article. The first three are logistic regression, multi-class perceptron's, decision trees, and *K*-nearest neighbors. A description of the proposed design is shown in Fig. 1. There is a system architecture model. Therefore, all zero values were removed from the image dataset used for training. Datasets are prepared, information is validated, and models are prepared to predict outcomes.

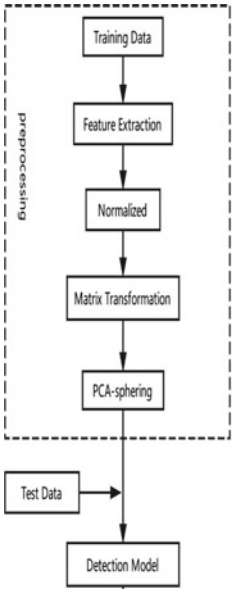


Fig. 1 Architecture model of proposed framework

3.1 Data Collection

The first step in this process is selecting the dataset to be used. In order to make the data set easier to handle, dataset selection is used. The smaller dataset can then be used for data mining and model or classifier training. Condensed subsets of all data points can be selected in a variety of ways. We utilized the kddcup dataset, which contains a variety of attacks such as multi hop, normal, saint, satan, snmpgetattack, snmpguess, teardrop, and udpstorm. Also included are snmpguess and snmpguess. With over 30,000 rows and 42 columns, this dataset has a number of properties. The first step is to gather information about network traffic, such as payloads and headers for packets. Switches, routers, and firewalls are examples of network hardware that can be used to collect this data.

3.2 Feature Extraction

Relevant attributes that depict traffic patterns are derived from the gathered data. Protocol distribution, packet size, packet rate, and traffic volume are all possible examples of these characteristics.

3.3 Data Preprocessing

The retrieved features undergo preprocessing to eliminate any noise or irrelevant data. This could make use of data normalization, data scaling, and data encoding. In this preprocessing step, we take out conflicting and fragmented information prior to filling in the spaces. The information is then standardized to fall between a scope of − 1.0 and 1.0.

3.4 Model Training

The preprocessed data are used to train the machine learning model. The ability to identify data patterns that indicate a DDoS attack is added to the model. A machine learning algorithm that is simple to use is the KNN (k-closest neighbors) method. It can be used for regression and classification problems, but classification problems get the most use. Network traffic patterns that resemble well-known DDoS attacks can be identified with the help of KNN. The interior of the selection tree represents the strength of the record set, the branches represent the selection guidelines, and the center of each leaf represents the outcome. This supervised learning approach can be used to solve class and regression problems. Decision trees, a supervised learning

approach, can be used to solve regression and class problems, but this approach is often preferred. The main attributes for separating between DDoS attacks and legitimate traffic can be found through choice trees. The multilayer perceptron method, a type of neural network, can be used to find connections between linear and nonlinear data. MLP can be used to learn the intricate nonlinear correlations between DDoS attacks and the properties of network traffic. The likelihood of a target variable can be estimated using logistic regression, a supervised learning classification technique. The relationship between a variable and one or more independent variables can be established in statistical software by estimating probabilities with the help of a logistic equation. To determine whether network traffic is a DDoS attack or not, you can use logistic regression.

The decision tree algorithm is a popular machine learning technique that is used for both classification and regression problems. It works by constructing a tree-like structure where each internal node represents a decision rule based on the input features, and each leaf node represents a class label or a predicted value. The goal is to create a decision tree that can accurately predict the outcome of new unseen data.

The decision tree algorithm works as follows:

i. Starting from the root node, choose the best feature that splits the data into the cleanest possible subsets. Purity is measured by impurity functions such as entropy, Gini index, or classification error. The element with the smallest impurity is chosen as the root node.
ii. For each subset generated by the split, the best feature selection process is repeated until the stopping criterion is met. The stopping criterion can be maximum tree depth, minimum number of samples per leaf node, or minimum impurity reduction.
iii. After the tree is built, new samples are predicted by traversing the tree from the root node to the end node according to a decision rule based on the input feature values. Class labels or predicted values associated with leaf nodes are then returned as final predictions.

Decision tree algorithms have several advantages, such as interpretability, ease of use, and the ability to handle both categorical and continuous features. However, it also has some limitations, such as its tendency to match noisy data, its sensitivity to small changes in the training data, and its inability to capture complex relationships between features strength.

3.5 Model Evaluation

A testing dataset is then used to evaluate the trained model's performance. Performance metrics include recall, precision, accuracy, and the F1-score, to name a few.

3.6 Model Testing and Validation

Testing the trained model on a set of validation data to ensure that it accurately predicts DDoS attacks while minimizing false positives and false negatives.

3.7 Deployment

The model can be used to predict DDoS attacks and continuously monitor network traffic once it has been trained and tested. On the off chance that an assault is expected, the framework can make preventive moves, like impeding or restricting traffic from the culpable gadgets. Compared to traditional rule-based approaches, machine learning-based systems for predicting DDoS attacks have a number of advantages. They can perceive already unidentified assaults and conform to moving assault designs. They need a lot of training data and precise parameter adjustments to perform at their best.

4 Results and Discussions

It is essential to identify attacks that cause cloud service outages in a highly distributed environment like the cloud, where DDOS attacks have become more easily detected. To identify such attacks, assault location datasets are every now and again prepared and tried utilizing AI models. Using information gain and multivariate analysis, the project's research objective is to create a feature selection ensemble for a machine learning model. The important kddcup dataset benefited greatly from the consideration of the exploratory dataset. Python and Jupyter Notebooks were frequently utilized for data wrangling work. We divided the dataset into the dependent class and, as a result, the independent class in the second step. Also, the dataset was normalized for the calculation.

4.1 Precision

The proportion of positive predictions that are actually correct. (When the model predicts default: how often is correct?). The equation for determining the precision value is shown in Eq. 1.

$$\text{Precision} = \text{True Positives}/(\text{TruePositives} + \text{FalsePositives}) \quad (1)$$

For positive observations, accuracy is the ratio of correctly predicted positive observations to all predicted positive observations. This measure is designed to determine how many passengers actually survived when they should have survived. A low false positive rate and high accuracy go hand in hand.

4.2 Recall

The proportion of observed positive expectations has realized. The ratio of correctly predicted positive observations to all true observations in a class (the percentage of true defaults that the model correctly predicts) is what is meant by the term "recall". This relationship is shown in Eq. 2.

$$\text{Recall} = \text{True Positives}/(\text{TruePositives} + \text{FalseNegatives}) \tag{2}$$

4.3 F1-Score

The weighted average of precision and recall is the $F1$-score. Therefore, this estimate accounts for both false positives and false negatives. This is often advantageous over precision, but $F1$ is not as intuitive as precision, especially when class distributions are not equal. Accuracy produces the best results when false positives and false negatives cost about the same. Accuracy and completeness are preferred when false positive and false negative costs differ significantly. Equation 3 describes the formula for determining the $F1$-score.

$$\begin{aligned} F1 \text{ - score} =& 2 * \text{True Positives}/2 * \text{True Positives} \\ &+ \text{False Positives} + \text{False Negatives}) \end{aligned} \tag{3}$$

4.4 Performance of LR, KNN, MLP, and Decision Tree Algorithms on Network Traffic

The results demonstrated in Table 1 tell that MLP and Decision Tree algorithms outperformed LR and KNN algorithms in terms of accuracy and $F1$-scores.

Table 1 Report table of network traffic

	LR (%)	KNN (%)	MLP (%)	Decision tree (%)
Precision	98.5	97.4	99.4	99.2
Recall	97.8	97.4	99.3	98.1
$F1$-score	98.6	97.5	99.4	99.2
Accuracy	98.8	98.4	99.8	99.6

Table 2 Report table of packet header

	LR (%)	KNN (%)	MLP (%)	Decision tree (%)
Precision	95.71	94.78	96.95	97.92
Recall	94.4	94.51	96.42	97.52
$F1$-score	95.65	94.73	96.74	97.85
Accuracy	96.41	95.34	97.55	98.22

4.5 Performance of LR, KNN, MLP, and Decision Tree Algorithms on Packet Header

The results in Table 2 indicate that the decision tree algorithm has the highest F1-score, precision, and recall.

4.6 Performance of LR, KNN, MLP, and Decision Tree Algorithms on Network Flow

The results indicated in Table 3 that the MLP algorithm had the highest $F1$-score, precision, and recall. When using network traffic data to anticipate DDoS attacks, these reports claim that MLP and Decision Tree algorithms typically perform better than LR and KNN algorithms. However, the exact performance of any method may alter based on the features used and network traffic conditions.

Table 3 Report table of network flow

	LR (%)	KNN (%)	MLP (%)	Decision tree (%)
Precision	96.41	93.56	97.67	96.62
Recall	96.45	93.47	97.45	96.57
$F1$-score	96.62	93.21	97.24	96.10
Accuracy	96.21	94.19	98.41	97.01

5 Conclusion

Using different gadget mastering strategies, we propose a method to predict DDoS attacks in this study. First, we chose the Kaggle KDD Cup dataset. It includes multi hop, standard, saint, devil, snmpgetattack, snmpguess, teardrop, and udpstorm attack types. To train the dataset, we used supervised gadget learning strategies such as logistic regression, MLP, *K*-Nearest Neighbors, and Choice Trees. About 25% of the records were used to validate the dataset and the last 75% were used for training. *K*-nearest neighbor and decision tree, which had the very best accuracy out of all of the four gadget mastering strategies, had been considered following the schooling of the datasets. The destiny overall performance and effectiveness of those structures may be better in severe ways. In various applications, for example, picture acknowledgment and regular language handling, convolutional neural networks (CNNs) and recurrent neural networks (RNNs), two instances of profound learning draw near, have exhibited exceptional execution. DDoS assault prediction might also additionally advantage from the usage of deep mastering strategies with inside the destiny to growth performance and accuracy. In addition, we intend to expand a hybrid approach with inside the destiny primarily based totally on the ones algorithms so that it will grow the price of accuracy with extra precision.

References

1. Beitollahi H, Deconinck G (2012) Analyzing well-known counter measures against distributed denial of service attacks in computer communications. Elsevier 35(11):1312–1332
2. Ferguson et al (1998) Network ingress filtering: defeating denial of service attacks which employ IP source address spoofing. Tech Rep, Internet Soc
3. Zhang L, Yu S, Wu D, Watlers P (2011) A survey on latest botnet attack and defense. In: Proceedings of 10th international conference on trust, security and privacy in computing and communications. IEEE, pp 53–60
4. Priyadharshini, Kuppusamy (2012) Int J Eng Res Appl (IJERA) 2(3):2263–2267. ISSN:2248–9622. www.ijera.com
5. Luo H, Chen Z, Li J, Vasilakos AV (2017) Preventing distributed denial-of-service flooding attacks with dynamic path identifiers. IEEE Trans Inf Forensics Secur 12(8):1801–1815. https://doi.org/10.1109/TIFS.2017.2688414
6. Rontti T, Juuso AM, Takanen A, Codenomicon Ltd (2012) Preventing DoS attacks in NGN networks with proactive specification-based fuzzing. IEEE Commun Magaz September
7. Paxson V, Bellovin M, Sally F, Mahajan R (2002) Controlling high bandwidth aggregates in the network a technical report
8. Eddy W (2007) TCPSYN flooding attacks and common mitigations. RFC4987
9. Wu Y, Zhao Z, Bao F, Deng RH (2015) Software puzzle: a counter measure to resource-inflated denial-of-service attacks. IEEE Trans Inf Forensics Secur 10(1)
10. Geva M, Herzberg A, Gev Y (2014) Bandwidth distributed denial of service: attacks and defenses. IEEE Comput Reliab Soc
11. Hassanand IU, Kaur A (2018) Prevention and detection of DDoS attack on WSN, pp 245–249

An Efficient Approach for Compression of Endoscopic Video Using MapReduce Technique

Suvarna Nandyal and Heena Kouser Gogi

Abstract Endoscopic video storage is a major issue today in cloud-based health centre. The EHR must include full-length endoscopic surgeries for diagnosis and research. This paper presents MapReduce-based compression. Thus, MapReduce solves the storage issue. The first AI-based solution is employed as an intelligent video splitter to form the key value as "Map" stages to filter the endoscopic video into a group of frames based on redundancy. These outputs are passed to "reduce" to merge them into a single output. After mapping and reducing endoscopic video frames, lossless compression is applied with PSNR 30–40 dB, SSI 0.7–0.8, Bitrate 32.17, MSE 2.1.

Keywords Hadoop · MapReduce · HEVC · PSNR · MS-SSIM · Bitrate · Compression ratio

1 Introduction

Video cloud services for hospitals are just one example of the rapidly expanding market for e-health and telemedicine-related products and services. National healthcare policy is gradually recognising telemedicine services, which use videoconferencing or other ICT technologies to enable distant treatment and consultations, as valid treatment choices. Countries including Germany, France, and Poland no longer mandate in-person patient–doctor meetings as of 2018 [1]. Even while similar offerings have been on the market for some time in the United States, European companies that want a piece of the pie will need to provide solutions that comply with European legislation and account for the wide variety of healthcare systems in Europe. The

Patent ID: 202041040000A.

S. Nandyal (✉)
Department of CSE, PDACE, Kalaburgi, Karnataka, India
e-mail: suvarna_nandyal@yahoo.com

H. K. Gogi
PDACE, Kalaburgi, Karnataka, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_32

telemedicine industry is growing [2]. As video technology advances, the amount of space required to store medical footage grows. Now that full HD systems are the standard and 4K and 8K resolution systems are on the way, this is more important than ever. High-quality surgical stereoscopic video is based on two streams of 1.5 Gbit/s each, while radiological images from a CT scan, MRI, or PET scan might take up several hundred megabytes per test. In 2009, 2.5 PB were needed to store mammograms in the United States, and by 2010, 30% of all photos kept globally were medical images [3]. Medical records are kept for extended periods in many nations. Medical photographs of patients must be stored for 20 years in Poland and France. It stores photos and videos. In 2015, France created 55 TB of picture data. Two hours of footage can be recorded on a full HD endoscopic camera used in endoscopic surgery. These figures highlight the difficulty of storing historical records. So, for regular use by health specialists, medical images and videos should be encoded (compressed) if doing so does not compromise the therapeutic quality of the material. It is vital that the compression phase of the encoder does not result in significant loss (degradation) or additional processing, such as watermarking, whether the data is being transmitted or stored. As a result, high-quality video is essential for medical professionals. Select codes that require less room for filing while still protecting patient safety. Compressing medical data is essential now. Consider the hospital's storage requirements. On a daily basis, endoscopic procedures account for 3% of the 32 operations performed in hospitals [4]. To meet the demands of developing applications like UHD TV in 4K and 8K resolutions, 360° video, and higher quality formats like HDR, HFR, and WCG, a new generation of video codecs is being created. Next-generation video codecs include VoD applications that use AV1. Open and royalty-free AV1 codec. In 2017, studies found that AV1 reduces bitrate by 17–22% compared with H.265/HEVC [5]. ITU-T VCEG and ISO/IEC MPEG formed the Joint Video Exploration Team on Future Video Coding (JVET) in October 2015 to determine if a new codec with compression capabilities beyond HEVC is needed. HD, UHD, HDR, and 360° films have been tested, and advanced compression algorithms have been investigated utilising a shared software platform [4]. JVET hopes to finish a new codec standard by 2020 (ITU-T may call it H.266). The open-source x265 encoder (which inherited its core techniques from x264 H.264/AVC) and commercial HEVC/H.265 encoders are available. The analysis states that only strict compression should be considered when dealing with the massive volumes of data that would be generated by recording medical operations. Therefore, it is essential that the quality for medical applications be considered during the standardisation of new robust compression codecs. Multiple journals have reported their findings on compressing medical videos for long-term storage. However, it's important to note that studies on multimedia data compression in medicine are typically conducted in highly specialised areas of medicine.

Looking at the opportunities and challenges that big data brings to gastroenterology. Although the proposed technique primarily focus on video compression, the storage implications of other data types utilised in the development of gastrointestinal healthcare, such as electronic medical records, patient registries, and patient-generated data and videos [6].

2 Related Work

The growing dimensions of data quantities generated by different medical imaging modalities and the increasing prevalence of medical imaging in clinical practice necessitate data compression for the distribution, storage, and management of digital medical image datasets. Since the advent of new technologies like the cloud, telemedicine, DICOM, and picture archiving and communication systems (PACS), data compression has become crucial. According to research by Ponlatha and Sabeenian [7], advances in medical imaging have hastened the accumulation of photographic evidence. Large and efficient storage systems are needed to house the enormous medical datasets produced by positron emission tomography, magnetic resonance imaging, computed tomography (CT), and single-photon emission tomography. Therefore, compression is an essential part of medical imaging, as it shortens the time it takes to send massive image studies like high-resolution three-dimensional CT datasets and reduces the amount of space they take up in storage. Video compression algorithms have inter-frame redundancies; however, these are not used in current medical picture compression standards. New DICOM WG-4 compression JPEG2000 is an add-on to JPEG. The compression of medical images allows for a lower bit rate, more space in storage, and a faster video quality. The improved quantisation matrix used by JPEG2000 is called a genetically designed GQM. Input blocks are used to quantise feature-based matrices. Here, compression is driven by the application. It is necessary to have a fitness function. Medical applications rely heavily on visual perception [7]. Senthilkumaran and Suguna [8] put forth a proposal for lossless image compression. The system demonstrates that the improved backpropagation neural network technique is superior to Huffman Coding for lossless image compression by comparing X-ray images on the basis of compression ratio, transmission time, and compression performance. Analysing experimental data side by side [8]. In a 2018 study, Nasralla et al. examined the quality of video encoded with HEVC and H.264/AVC. The JCT-VC group compares their findings among intra, low-delay, and random-access settings. In order to provide an accurate comparison, H.264/AVC is set-up to roughly mimic the set-up settings of HEVC. Using 4G network simulations, compare the system resource utilisation in user authentication, message, data secrecy, and compression when transmitting HEVC and H.264 encoded videos. Raja [9] suggests compression and authentication for multimedia data due to its large storage needs. Multimedia cannot function without data compression. Cloud-based meta-data businesses can reap the benefits of compression technologies in a dependable manner. The purpose of DICOM is to lay a new groundwork for multi-scale transforms and encoding methods. Medical picture compression in the cloud with public-key encryption is a step towards increased safety. Some examples of transformations are the bandelet, contourlet, wavelet, and ridgelet. Methods like the Rivest–Shamir–Adleman (RSA) and the wavelet difference reduction (WDR) and the adaptively scanned WDR and the tree zero tree algorithm and the transmission efficiency algorithm are included. Better compression than H.264/AVC means less bandwidth is needed to transmit a HD video over an LTE/LTE network [9]. Since

there is such a large quantity of medical photos to save and transmit, a powerful compression method is necessary, as pointed out by Ayoobkhan et al. [10]. Compression methods used in medical imaging must preserve image and video quality. The integrity of medical images is safeguarded via predictive image coding. Restoring medical picture quality, predictive image coding does so by encoding without losing the diagnostically relevant region (DIR). The image is segmented using a graph-based segmentation method, which separates DIR and non-DIR regions. Two feed-forward neural networks (FF-NNs) are used in the prediction method, one for compression and one for decompression [10]. Medical image data mining, virtual reality, medical visualisations, and neuroimaging are only some of the areas that Deserno et al. [11] investigate in terms of the challenges posed by datasets ranging from a few hundred kilobits to several terabytes. As data volumes increase, improvements in image processing and visualisation will be required. GPU-based parallelisation and scalable approaches have been created. It is a synopsis of their paper. Such approaches are currently able to deal with data at the terabyte level, but the petabyte threshold is rapidly approaching. Consequently, studies including medical image processing are essential. These days, people are more interested in telemedicine, which uses the Internet to link people in far-flung areas with doctors who specialise in their condition. Medical image transmission via telemedicine necessitates large amounts of data storage and network throughput. Compressing medical photos is necessary for both archiving and sending [11]. A sparse representation medical picture compression method based on the geometry of the underlying image structure is provided by Juliet et al. [12]. Regular variations in greyscale are shown by the geometric flow of an image. Important coefficients can be lowered in the decomposition of wavelets via geometric regularisation. When medical images are compressed and an optimal storage and communication solution is developed, vital image data is preserved and processing errors are avoided. MAE, UIQ, and PSNR all have a role in getting there [12]. Scholars and professionals alike are interested in cloud video big data analytics, according to a study by Alam et al. [13]. An efficient video big data analytics platform is necessary because of developments in technology and in business. The current literature does not go far enough in providing an architecture for cloud-based video big data analytics, which would address issues with and opportunities for handling and analysing such data [13]. Let's check out the HEVC next. For better compression efficiency than its predecessor, H.264/AVC, HEVC (high-efficiency video coding, or H.265) was developed as a video compression standard. HEVC is able to accomplish its goal of compressing video without sacrificing quality by making use of a number of cutting-edge techniques [14].

3 Implementation of an Efficient Approach for Compression of Endoscopic Video Using MapReduce Technique

Cloud-based health centres struggle with massive medical data collections. Large volumes provide a problem for storage and communication. The resolution, frame rate, compression method, and duration of an endoscopic video determine its storage space. Endoscopic videos are usually 720p or 1080p and 30 fps or less. The endoscopic video may last a few minutes or several hours, depending on the technique. The following formula estimates the storage space needed for an endoscopic video recorded at 1080p, 30 fps, and H.264 or H.265 compression. Storage space = (bit rate × duration)/8, where bit rate is the data per second needed to encode the video and duration is its length in seconds. For example, if the video's bit rate is 5 Mbps and its runtime is 60 min (or 3600 s), the storage space needed is: (2.25 GB) = (5 Mbps × 3600 s)/8. Thus, a 60 min endoscopic film compressed with H.264 or H.265 at 1080p and 30 fps would take 2.25 GB of storage. This is merely an estimate and may vary depending on video characteristics. MapReduce-based compression is implemented in MATLAB and video compression solution worked well in experiments. Figure 1 shows endoscopic video before compression using MapReduce. It generates GOP from endoscopic video. Hadoop uses MapReduce. The MapReduce parallel paradigm first converts the frame to turn one set of "Key, the value"-formatted input data into a distinct set of output data using predetermined mapping rules.

3.1 Proposed Intelligent Video Splitter

The suggested architecture splits input endoscopic videos into 15 frame-blocks and saves them in HDFS across many devices for testing. The video splitter may compare

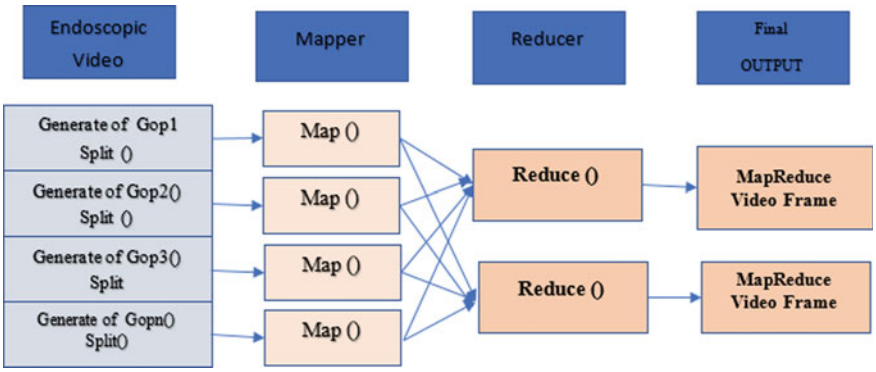


Fig. 1 Typical MapReduce computing framework

data from sources other than the input stream before splitting. Splitter must decide how much work to send to each framework hub and whether to increase or decrease instances. Figure 2 shows the clever video splitter workflow. Small blocks can be transmitted easily. When the size is large, block size is crucial. If fragment frame size exceeds 15, the threshold value of 8–9 has been chosen. However, for images with higher complexity or more significant diagnostic information, a higher threshold value may be required to preserve the necessary details.

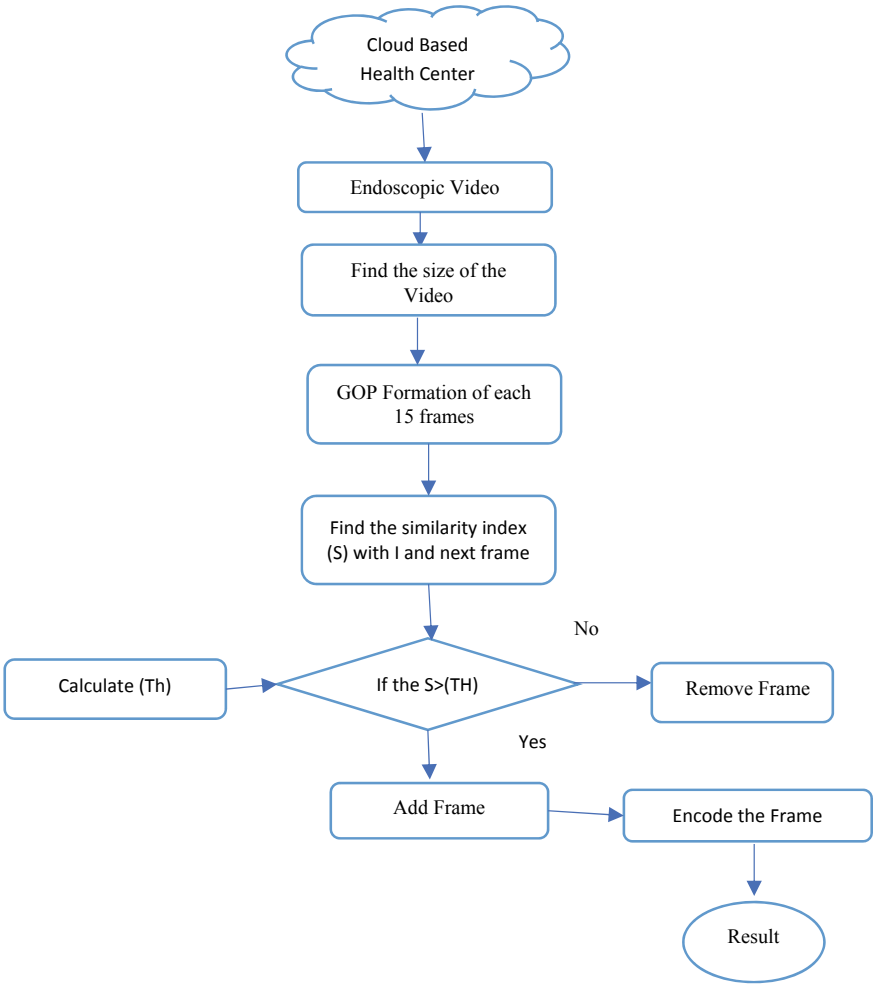


Fig. 2 Flowchart of AI-based video splitter

3.2 Implementation of Hadoop-Based MapReduce Video Encoding Technique

The proposed system takes the input as endoscopic video as shown in Fig. 3 which is of HDFs. It is a distributed file system that uses commodity hardware to store and manage huge datasets in a distributed setting. HDFS is a component of the Apache Hadoop project, which provides a framework for distributed data storage and processing. The steps are follows.

MapReduce compression with an artificial intelligence video splitter:

(1) Input Data Splitting: For parallel processing, the input video is divided into smaller chunks of data. Typically, an AI video splitter is used, which can identify scene changes and split the stream accordingly. Each block of data is handled by

The main steps of Hadoop's MapReduce-based parallel video encoding

Video Map Operation

get started with a Map (String key, String value) // key: Video Title, Value: Video Title and body Call
Cut the video into 15-frame chunks (Nf) with this intelligent video splitter (key, value).
for *each block in* **Nf**
transcode Intermediate (s, "format");
end for
end

Video Reduce Operation

Reduce (String key, Iterator values) starting. //a list of Transcoded Blocks TNf Video v1 when the key is a cut;
for each v1 in TN blocks that have been transcoded.
Discharge (AsMergedVideo(v1)).;
Put it in a Video Cloud Archive
end for
end

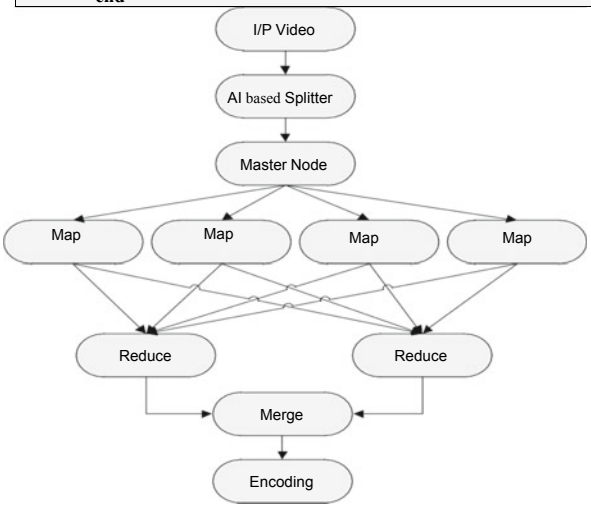


Fig. 3 Architecture of HDFs: MapReduce video transcoding

a map function, which compresses the data. This can be accomplished through the use of various compression methods like as H.264 or HEVC.

(1.1) Call Intelligent Video Splitter (key, value)
for each block in Nf

(2) Map Function: Map Operation (Key, Value) // key: Video Title, Value: Video Title and body Call.
(3) Reduce Function: The grouped data is subsequently processed by a reduction function, which compresses the data even further. This can be accomplished by using extra compression algorithms.
begin to Reduce (String key, Iterator values) // key: a cut, values: a list of Transcoded Blocks TNf
(4) Output Data: Finally, the compressed data is written to a storage system, such as a hard disc or a cloud storage service.

3.3 Experiments and Results

This section describes the MapReduce methodology used in the experimentation setting. In this part, the proposed technique represents the outcomes of the experiments with the suggested Hadoop-based Endoscopic video compression. The proposed system's primary focus is on successfully compressing input videos by transforming them into a series of GOP frames. Information regarding how the endoscopic footage was mapped and compressed is shown in Fig. 5. The effectiveness of a MapReduce task processing a huge video can be greatly affected by the number of mappers and reducers utilised. More mappers and reducers are needed for processing larger videos because each mapper and reducer must process a smaller portion of the video. Figure 5 shows the first original frame of endoscopic image after applying the MapReduce to endoscopic video before compression and HEVC final inter frame. Figure 4 shows the final compressed progresses of an endoscopic video of a typical endoscopic video. For experiment purpose, five endoscopic videos are considered for compression and parameter for study purpose as given in Table 1. Figure 6 shows the graph of inter code encoding bitrate. Typical 1080p video at 30 frames per second, a bit rate of 8 Mbps may be considered a good balance between video quality and file size. However, for 4K video at 60 frames per second, a bit rate of 50 Mbps or higher may be needed to achieve good video quality for experiment achieved that is an average 32.47 Mbps, bitrate may vary depending on the specific requirements of application and video. In general, a PSNR value of 30 dB or higher is considered to be of high quality, while a value of 20–30 dB is considered to be of medium quality, and a value below 20 dB is considered to be of low quality, Fig. 7 shows the sample 5 video, which achieves a PSNR of average value 44.38 dB, which is ideal for present application. Lower MSE is better, as it indicates a smaller difference between the original and compressed video frames and in Fig. 8 shows the graphs for MSE. The ideal SSI for compression of video depends on the specific requirements and the type

of video content being compressed, the achieved average value of 0.83 is shown in Fig. 9. A compression ratio of 0.45 indicates that the compressed data is approximately 45% of the original data size. In other words, the data has been reduced to approximately 45% of its original size as shown in Fig. 10. However, it's important to note that no one parameter can be used to justify the measure of video quality, and should be used in conjunction with other measures, such as peak signal-to-noise ratio (PSNR), mean squared error (MSE), SSIM and subjective testing to determine the optimal compression settings for particular video application.

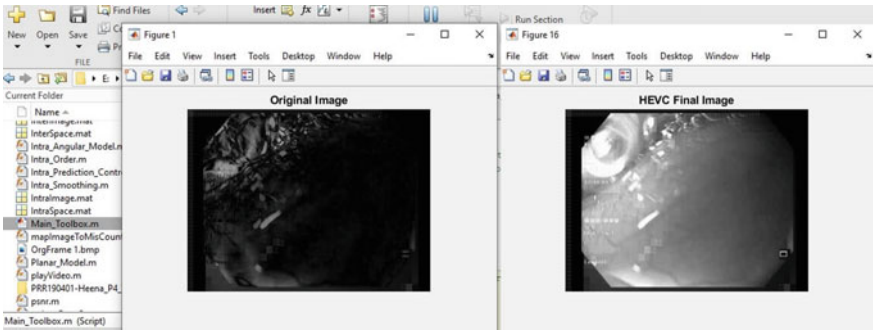


Fig. 4 Original image and final inter frame

```
*       MAPREDUCE PROGRESS       *
**********************************
Map    0% Reduce    0%
Map   10% Reduce    0%
Map   20% Reduce    0%
Map   30% Reduce    0%
Map   40% Reduce    0%
Map   50% Reduce    0%
Map   60% Reduce    0%
Map   70% Reduce    0%
Map   80% Reduce    0%
Map   90% Reduce    0%
Map  100% Reduce    0%
Map  100% Reduce   10%
Map  100% Reduce   20%
Map  100% Reduce   30%
Map  100% Reduce   40%
Map  100% Reduce   50%
Map  100% Reduce   60%
Map  100% Reduce   70%
Map  100% Reduce   80%
Map  100% Reduce   90%
Map  100% Reduce  100%
```

Fig. 5 MapReduce progress of endoscopic video

Table 1 Results of five endoscopic videos

	Bit rate	PSNR (dB)	MSE	SIM	C-R
1	31.91	37.07	2.02	0.70	0.45
2	32.03	35.72	2.75	0.72	0.45
3	32.97	35.44	2.71	0.86	0.44
4	32.23	35.79	2.94	0.84	0.45
5	32.49	36.95	2.07	0.730	0.45

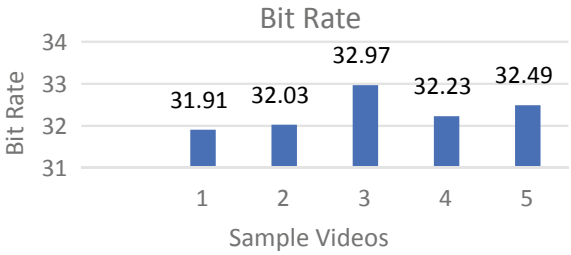


Fig. 6 Inter encoded bit rate of five endoscopic videos

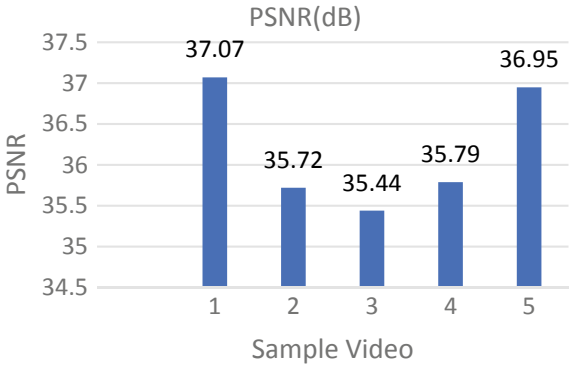


Fig. 7 PSNR value of five endoscopic videos

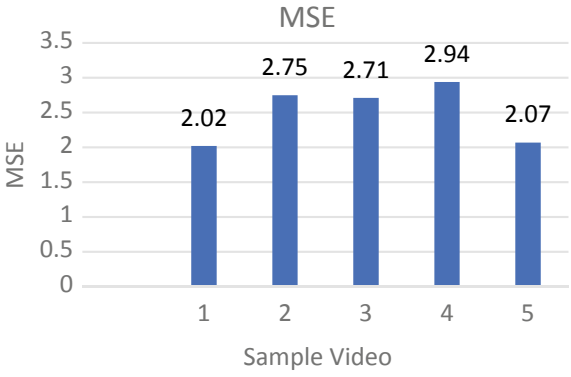


Fig. 8 MSE value of five endoscopic videos

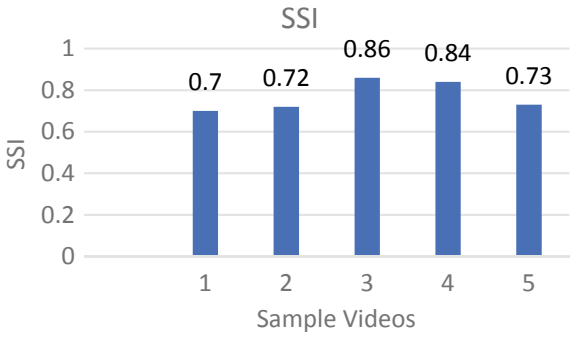


Fig. 9 SSIM of five endoscopic videos

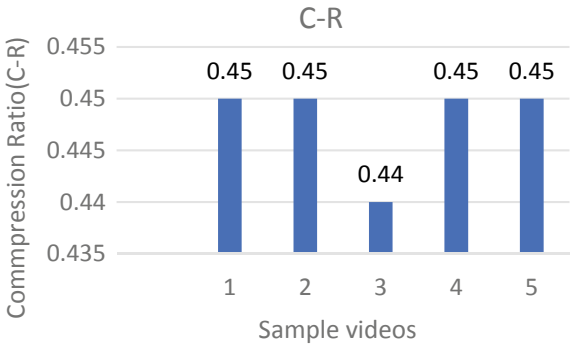


Fig. 10 Compression ratio of five endoscopic videos

3.3.1 Scalability and Resource Usage of MapReduce in Compression

MapReduce is designed to expand and allow the compression burden to be distributed across numerous processors. This means the compress can be done in vast volumes of data in a reasonable amount of time. The proposed system supports the scalability. The experiment is intended to evaluate the scalability of a created application and proposed architecture. The experiment uses endoscopic video and assesses the system's performance when encoding the video using 1, 3, 6, and 9 Worker Video Encoders (module which perform compression), each with a chunk size of 15 frames. To achieve balance, the workload is spread evenly between computing nodes. A master video node in this configuration generates seven encode job messages, which are subsequently handled by the worker encoders. Each worker encoder starts an algorithm with the appropriate parameters for the encoder. The data generated from this experiment can be used to evaluate the system's performance and scalability and make appropriate modifications. It can assist in determining the best configuration for processing films with varying numbers of worker encoders and chunk sizes.

4 Conclusion

The proposed MapReduce-based compression technique solves the storage problem of endoscopic videos in a large-scale private cloud as health-based cloud centre which uses an AI-based approach to split the video frames into groups based on redundancy, which are then processed using the MapReduce framework to produce a single output. The output frames are later compressed using a lossless compression technique to reduce the storage requirement, while maintaining identical quality and bitrate of the original video. The compression results achieved a PSNR of 30–35 dB, SSIM of 0.7–0.8, MSE of 2.1, a bit rate of 32.17 and a compression ratio of 45%. The proposed technique has shown promising results and can be used to efficiently compress and store endoscopic videos for further diagnosis and research.

References

1. Telemedizin in Deutschland. Fernbehandlung läuft schleppend. Krankenkassen-Zentrale. URL: https://www.krankenkassenzentrale.de/magazin/telemedizin-in-deutschland-fernbehandlung-laeuft-schleppend-97525. Accessed 28 Aug 2019
2. EHealth Action Plan 2012–2020. Innovative healthcare for the 21st century. European Commission. URL: https://ec.europa.eu/health/sites/health/files/ehealth/docs/com_2012_736_n.pdf. Accessed 28 Aug 2019
3. Hey T, Trefethen A (2003) The data deluge: an e-science perspective. In: Berman F, Fox G, Hey T (eds) Grid computing: making the global infrastructure a reality. Wiley, Hoboken, NJ, pp 809–824
4. Panayides AS, Pattichis MS, Loizou CP, Pantziaris M, Constantinides AG, Pattichis CS (2015) An effective ultrasound video communication system using Despeckle filtering and HEVC. IEEE J Biomed Health Inform 19:668–676. PMID: 24951708
5. Han J, Mukherjee D, Chen C, Parker S (2021) A technical overview of AV1. Proc IEEE 109(9)
6. Catlow J, Bray B, Morris E, Rutter M (2022) Power of big data to improve patient care in gastroenterology. https://doi.org/10.1136/flgastro-2019-101239
7. Ponlatha S, Sabeenian RS (2013) Comparison of video compression standards. Int J Comput Electr Eng 5(6)
8. Senthilkumaran N, Suguna J (2011) Neural network technique for lossless image compression using X-ray images. Int J Comput Electr Eng 3(2). ISSN: 1793-8163
9. Raja SP (2019) Joint medical image compression–encryption in the cloud using multiscale transform-based image compression encoding techniques. Sādhanā 44(2):28
10. Ayoobkhan MUA, Chikkannan E, Ramakrishnan K (2018) Feed-forward neural network-based predictive image coding for medical image compression. Arab J Sci Eng 43(8):4239–4247
11. Deserno TM, Kuhlen TW, Aach T (2011) Challenges of medical image processing. Comput Sci Res Dev 26:5–13
12. Juliet S, Rajsingh EB, Ezra K (2015) A novel image compression method for medical images using geometrical regularity of image structure. Signal Image Video Process 9(7):1691–1703
13. Alam A, Ullah I, Lee Y-K (2020) Video big data analytics in the cloud: a reference architecture, survey, opportunities, and open research issues. arXiv:2011.07807v1 [cs.DC]
14. Nandyal S, Gogi HK (2023) An efficient compression of endoscopic videos using DPCM encoding. J Data Acquis Process 38(1). ISSN: 1004-9037. https://doi.org/10.5281/zenodo.7766101

Dr. Suvarna Nandyal received the B.E. degree in Computer Science and Engineering, Gulbarga University, Karnataka and M.Tech. degree in Computer Science and Engineering, Visvesvaraya Technological University, Belagavi, India, in 1993 and 2003, respectively, and the Ph.D. degree in Computer Science and Engineering from Jawaharlal Nehru Technological University, Hyderabad, India, in 2013. She is currently a Professor in the Department of Computer Science and Engineering, PDA College of Engineering, Gulbarga. With over 22 years of teaching experience, she is guiding eight Ph.D. students and guided many undergraduate and post graduate projects. She has authored or co-authored over 100 articles in refereed international journals and conferences. Her current research interests include the image processing, video and sound processing, information retrieval, mobile applications and cloud computing. She has five Indian/International Patents grand and five published.

Heena Kouser Gogi received master's degree in Computer Science and Engineering from VTU, Belagavi, India. She is currently the Ph.D. scholar of the PDA College of Engineering, affiliated to VTU, Kalaburgi, India. Her research interests include image processing, cloud computing, video processing, and artificial intelligence. She has authored or co-authored over 10 articles in refereed international journals and conferences. She has two Indian/International Patents published.

Conversion of Natural Language to SQL Query

Chaitanya Nirfarake, Hrishikesh Vaze, Atharva Wagh, Zeeshan Mujawar, and Preeti Kale

Abstract This paper outlines a method for efficiently automating the conversion of Natural Language Query to Structured Query Language (SQL) queries. SQL is a standard tool for relational database management systems. The correct SQL query with proper keywords and syntax must be entered to obtain or manipulate data from such databases. Therefore, we put forward Natural Language Sequential Query Language "NLSQL", a system in natural language processing to get around this problem and convert natural language queries to SQL queries. To revamp prevailing advancements in this field, we propose three times state-of-the-art parsing and implementation for JOIN queries with our system's support for multiple tables in a rule-based system. Along with basic queries, complex commands can also be handled by this system. After experimenting and testing this work on a number of natural language queries we were able to achieve convincing results.

Keywords Automation · Natural language query · SQL · Query · Dynamic · Database

C. Nirfarake (✉) · H. Vaze · A. Wagh · Z. Mujawar · P. Kale
School of Computer Science Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India
e-mail: chaitunirfarake@gmail.com

H. Vaze
e-mail: hrishikesh.vaze01@gmail.com

A. Wagh
e-mail: a3wagh@gmail.com

Z. Mujawar
e-mail: zmujawar2@gmail.com

P. Kale
e-mail: preeti.kale@mitwpu.edu.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_33

1 Introduction

Natural language processing (NLP) is a field of computer science and artificial intelligence that deals with the interaction between computers and human language. NLP's major purpose is to enable computers to understand and interpret user queries in natural language. Natural language processing may be used to retrieve varied information from databases by asking inquiries in natural language and receiving the necessary results in the form of responses from the database. Natural language queries are a quick and easy way to access databases, especially for users who are inexperienced with complicated database query languages like SQL. NLP techniques are also used to translate natural language text from one language to another. This involves analyzing the grammatical structure of the source text, identifying the meaning of individual words and phrases, and generating a corresponding translation in the target language. NLP techniques can also be used to improve the accuracy of machine translation systems by identifying common errors and suggesting more appropriate translations.

This system is designed for anyone and everyone regardless of knowing SQL. The system proposes the architecture for processing the user input in terms of English language breaking it down to machine language and then procuring the requested results with the SQL query.

1.1 Motivation

Machine translation is now the most widely utilized NLP application for addressing linguistic barriers. As the amount of data available on the Internet expands, so does the need to access and comprehend it. Data may be transformed from one language to another via machine translation. NLP technologies let machines understand the meaning of phrases, boosting the efficacy of machine translation.

To determine or analyze the structure of a string of words presented to the computer, parsing is employed for this goal so that it may find its structure utilizing grammatical rules. When queries must be returned after translation, the kind of data is sure to be complicated and varied, and information from numerous tables may be required, hence JOINs in SQL are used to retrieve data from other tables based on logical relationships.

In this study, we propose the framework NLSQL in a rule-based system by combining the aforementioned elements.

1.2 Contribution

In order to aid machine translation, we propose a rule-based approach implemented in the system.

1. The system is dynamic and can work with any relational database. The adapter pattern is used to effectively generate the final SQL query (see Sect. 3). Although the system currently only has a MySQL driver, more drivers can be easily added in the future.
2. The system features parsing thrice with each one having a different function to maximize accuracy.
3. It also supports JOIN queries, which have been a limitation to several types of similar systems.
4. Dynamic table selection helps the system to select required tables only.

In this paper, we review the implementation of our model of NLSQL by analyzing previous related work and research in Sect. 2, discussing the implementation of our model in Sect. 3, demonstrating the functioning in Sect. 4, and concluding with results in Sect. 5.

2 Literature Review

The task of creating SQL representations from English queries has been researched in the NLP and DB fields since the 1970s [1]. Building a natural language interface for a database (NLIDB) has many critical challenges [2]. This happens in two ways: keyword mapping and joins inference [2]. The task of mapping tokens from a Natural Language Query (NLQ) to the database and its content is known as keyword mapping. Join inference is selecting the shortest way joining the required tables in the NLQ [2]. We can create NLIDB with two approaches, either a rule-based system or a model-based system. However, the commonly used system is model-based [1–4]. It uses long short-term memory (LSTM) recurrent neural network (RNN).

There has been extensive work in the NLP community on semantic parsing to logical representations that query a knowledge base, while work on mapping to SQL has recently increased [1]. To apply neural models to this task, we have to annotate large enough datasets of question-query pairs, which is a challenge. Recent research has generated large datasets automatically by using templates to form random queries and corresponding natural language-like questions, which humans then rephrased into English [5]. Another option is to use feedback-based learning, which involves the system alternately training and making predictions that the user rates as correct or incorrect [1].

Other semantic parsing tasks have been significantly improved by grammar-based decoding, but SQL and other general programming languages have complexities not found in logical formalisms, making writing hierarchical grammar difficult. For the NLIDB task, the paper [4] proposed a model that guides the decoding process using a dynamic, schema-dependent SQL grammar and a deterministic entity linking module [4].

The paper [6] contributes to proposing a system, NatSQL—a new intermediate representation, to convert natural queries into partially inferred SQL queries or intermediate results (IR) [6]. To test the NLP to SQL dataset which primarily aims to teach and test understanding of advanced SQL paper [7] suggests the use of spider dataset. Our work demonstrates how we can use a rule-based system to convert natural language queries into SQL queries.

Table 1 compares our system with formerly developed rule-based and model-based frameworks in terms of three major factors from Table 1 parsing implementation, selecting tables based on and joining queries that allow multiple table support. Also, the last tab "type of system" describes whether the paper talks about model-based system (M) or ruled-based system (R).

A rule-based system is effective at capturing a certain linguistic phenomenon: it will decode the linguistic links between words to interpret the sentence. As a result, it excels at sentence-level tasks like parsing and extraction. As a result, rule-based techniques are generally better suited to query analysis.

Table 1 Literature comparison for related work

Features			Paper No.	Type of system
Parsing	Dynamic table selection	Join inference		
Y	Y	Y	[1]	M
N	Y	N	[3]	M
N	Y	Y	[2]	M
Y	N	N	[8]	R
Y	Y	Y	[9]	M
N	Y	Y	[10]	M
N	N	N	[11]	M
Y	N	Y	[12]	M
N	N	N	[13]	M
Y	Y	Y	[4]	M
Y	Y	Y	[6]	M

3 Methodology

To convert natural text given by the user to a proper SQL query, steps used in our program can be classified into four types of steps:

A. Initialization

Here, system will do certain tasks for its initialization process like fetching required information related to the database. It will also initialize certain dictionaries which are required for further processing.

B. Lexical Analysis

In this step, system will get the user's input and will tokenize it for further processing.

C. Intermediate SQL Generation

This step includes parsing, reordering, and filtering of tokens, thus generating intermediate SQL query tokens.

D. Final SQL Generation

In this step, SQL tokens are parsed and joined generating the final executable SQL query.

3.1 Scope

For now our program only supports SELECT queries. So users can ask questions like "show …", "display …", "count …", or "which …", etc.

3.2 Setup

For our application, we use our property sale database, which includes the following tables:

I. Homes—This table has the following attributes: id (primary key), owner_id, address, and rooms column.
II. Owners—The owner's table has the following attributes: id (primary key), name, email, phone, and age.
III. Sales—The sales table is made of the following attributes: id (primary key), home_id, and price.

We test the system with following tools:

- Operating system: Windows 11
- Language: Python3.10
- Database: MySQL.

The MySQL system driver was also used (as discussed in Sect. 3.6.8). To test the accuracy of our system, we created 100 queries having four types of difficulty (see Sect. 5 for more details).

3.3 Inputs

User will enter credentials for his database and will start the program. Our program will only require the database credentials (along with its respective database driver, which is included in the program).

3.4 Data Structure

Keyword Dictionary

The system uses a Python dictionary that contains SQL keywords that can be matched with user input, Table 2 presents keywords from natural keywords (user input) mapped to tokens (smaller chunks of data) to replace them with machine-understandable values.

Dictionary mentioned in Table 2 cannot be used directly, so we convert it to this type of dictionary, wherein each word to replace has its value as tuple having first element as value to be replaced and second element as token type:

```
{'select': ('select', 'select'), 'fetch': ('select', 'select'),
'from': ('from', 'from'), . . . }
```

3.5 System Block and Flow Diagram

Figure 1 shows how the system interacts with user and databases. Figure 2 shows the flow of steps in the NLSQL algorithm starting from top to bottom. These steps are explained further in detail in Sect. 3.6.

Table 2 SQL keywords mapping

Expected natural query words	Replaced value	Token
SELECT, FETCH, GET, SHOW, LIST, DISPLAY, WHICH	SELECT	SELECT
FROM, IN, OF	FROM	FROM
ORDER BY, SORT BY, ORDER, SORT	ORDER BY	KEYWORD
NOT, NEITHER, NONE OF	NOT	
LIMIT, TOP, FIRST	LIMIT	
OFFSET, SKIP, AFTER	OFFSET	
AVG, AVERAGE, MEAN	AVG	FUNCTION
COUNT, TOTAL	COUNT	
AND, BOTH, TOGETHER	AND	CONJUCTION
OR, EITHER, ONE OF	OR	
EQUAL TO, IS	=	OPERATOR
>, GREATER	>	
<, LESS	<	
*, MULTIPLY	*	

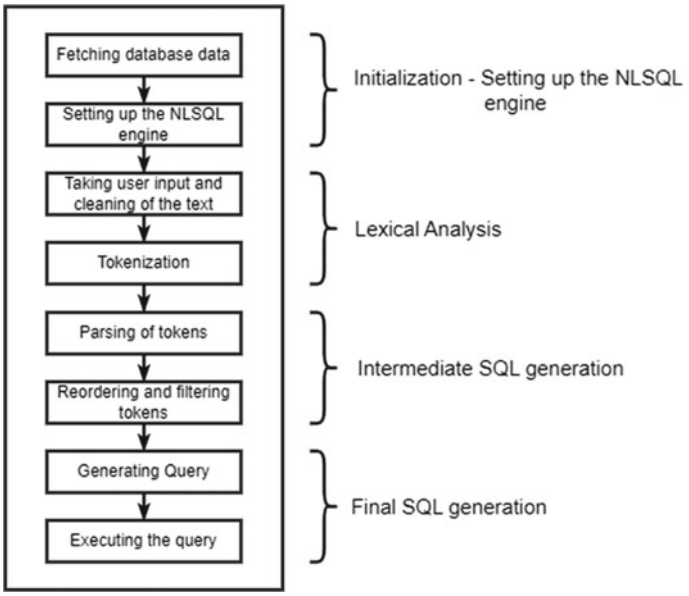


Fig. 1 System flow aggregation

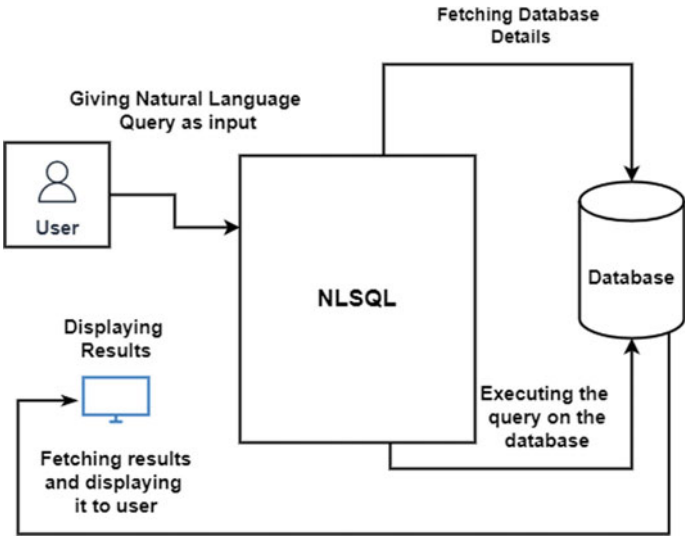


Fig. 2 System block diagram

3.6 Algorithm

Our algorithm works in eight major steps from fetching the natural query to executing the SQL query in the following way.

Step 1: Fetching the database data.
Step 2: Setting up the NLSQL engine.
Step 3: Taking user input and cleaning of the text.
Step 4: Tokenization.
Step 5: Parsing of tokens.
Step 6: Reordering and filtering tokens.
Step 7: Generating query.
Step 8: Executing the query.

Let's see what each step includes.

3.6.1 Fetching the Database Data

As soon as the program starts, the database driver will retrieve all of the information, including the list of tables, columns in each table along with its pseudo data types, foreign keys, and primary keys. This step will only be done once per system bootup.

Table 3 explains how the system maps database datatypes to token types.

Table 3 Data types and token types fetching

Original datatypes	Token types
Int, float, double, decimal	Number
Date, datetime, timestamp	Date
Rest all datatypes	String

3.6.2 Setting Up the NLSQL Engine

System will need to initialize a table selection dictionary. Selecting all tables from a database in every SQL query, even if they are not required, will result in a performance problem. We identify the tables required in our query and dynamically select them using the table selection dictionary. This selection will be carried out in step 5 (see Sect. 3.6.5).

3.6.3 Taking User Input and Cleaning of the Text

The program now accepts and processes user input. The first step is to clean up the text. Some words that are not useful in forming SQL will be ignored here. The following are some of the words used in ignore list in our program: the, a, are, is, to, in, then.

3.6.4 Tokenization

Following text cleaning, the next critical step is tokenization, in which the text is divided into a list of tokens for future processing.

3.6.5 Parsing of Tokens

Parsing includes the identification of each token into keywords, table names, and column names. It mainly consists of three steps:

1. Formatting: The system uses a regular expression to allow only digits, _, and characters in tokens. Next, we convert all characters to lowercase.

$$\text{Regular Expression used: } [\hat{}0-9a-zA-Z\_] \tag{1}$$

2. Lemmatization: It is an essential part of natural language processing. This process reduces the token into its root form ensuring the root form is valid according to the language.

3. Text Enrichment: Here, as the name suggests we will try to enrich the natural text. The program will try to find tokens that are almost similar to the keywords, table names, or column names and will convert them to the exact word.

The system will perform the parsing three times. Parsing will search the current tokens in the following order:

1. Keywords
2. Table name
3. Column name for tables which are selected.

Following cases can happen while text enrichment:

Case 1. The column name comes before the table name.

The parser cannot recognize this token as a column name, as there is yet no table selected. So it ignores the token and continues its parsing for the next tokens. As a result, when the next time the parser parses the unparsed tokens, it will be able to recognize the column as we have already done parsing of the other tokens, including the table names.

Case 2. Parser still does not recognize a token in the second attempt.

The parser considers it a value for the third time and finds the best-matched column for the token to be the value of.

3.6.6 Reordering and Filtering Tokens

a. Reordering

Most of the time natural query includes the word sequence "descending order", when converted to a stream of tokens, the token ORDER BY appears after the token DESC which is incorrect according to SQL syntax. Hence the system reorders these types of tokens.

b. Filtering

From SQL syntax, we know that a token of type SELECTOR should not appear after the token FROM. In the same way, tokens containing table names as the value should not appear after the token WHERE. This step filters out these types of unwanted tokens.

3.6.7 Generating Query

So the tokens have been properly parsed, filtered, and reordered based on the syntax of the selected database. The only process left is the creation of a SQL query. Because the output of this step will differ depending on the database, this step will be handled by a custom-made database driver (Fig. 3).

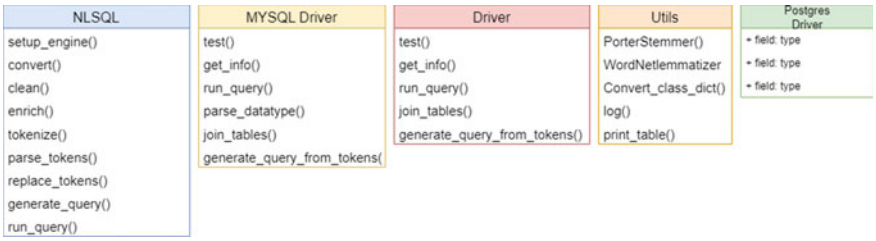


Fig. 3 Class diagram for driver

3.6.8 Executing the Query

In addition, after executing the generated query on the database, our program will also show the result. NLSQL Class:

This class will handle full NLSQL conversion. It will need one database driver.

$$\text{nlsql\_engine} = \text{NLSQL}(\text{db\_driver} = \text{db\_driver}) \tag{2}$$

3.7 Joining Tables

Consider the following input to the join algorithm. Let us see how the program converts them to join statement string. From the database information fetched before, we will use information about foreign keys and primary keys.

```
input = ['sales', 'homes']
```

Intermediate result:

```
joining_dictionary = { 'sales': [ ('homes', 'sales.home_id') ] }
```

Here we can see that we have sales joined with one table of homes with key sales.home_id as foreign key referencing the primary key "id" of the referenced table "homes".

Final join phrase:

```
" sales INNER JOIN homes ON sales.home_id = homes.id"
```

4 Results and Analysis

We ran our model on three benchmark databases. The model was tested on a set of several English questions whose corresponding SQL query was generated as well as the query returned the required data. In order to explain the accuracy of the model, we compared the queries returned by the model with the correct queries and have tabulated our comparison in Table 4.

Overall accuracy of the model was found to be 85.46%.

To test the accuracy of our model, we categorized the queries into four categories:

A. **Easy**

The SELECT and WHERE keywords are put to the test using a single table in the easy level.

For example, "show names of owners with age equal to 30"

SQL query output: `"SELECT owners.NAME FROM owners WHERE owners. Age = 30"`

B. **Medium**

The medium level evaluates numerous SQL operations on a single table, including AVG(), SUM(), and COUNT().

For example, "avg age of all owners"

SQL query output: `"SELECT AVG(owners.age) FROM owners"`

C. **Hard**

The hard level tests the JOIN functionality using two or more tables.

For example, "show email of owners with rooms more than 2"

SQL query output: `"SELECT owners.email FROM homes INNER JOIN owners ON homes.owner_id = owners.id WHERE homes.rooms > 2"`

D. **Extra hard**

The extra hard level evaluates the JOIN capability while applying SQL functions like AVG(), SUM(), and COUNT().

For example, "count price of rooms less than 20000"

SQL query output: `"SELECT COUNT(*) FROM sales INNER JOIN homes ON sales.home_id = homes.id WHERE homes.rooms < 20000"`

Table 4 Accuracy table

Category	Total count	Queries passed	Accuracy (%)
Easy	23	23	100
Medium	7	5	71.42
Hard	38	34	89.47
Extra hard	21	17	80.95

We discovered that the system joined tables unnecessarily in a few queries.

- For example, "show the homes for sale", there was no need to join tables homes and sales.
- In a few queries, the system either selected unnecessary columns or did not choose those columns that were not explicitly requested.

5 Conclusion and Future Work

We have proposed an innovative approach to translate natural language into SQL queries. Starting with the initial set of program tokens obtained by lexical parsing, our approach enters an iterative process of parsing the tokens, reorders, and filters the unwanted tokens. Our method also uses dynamic table selection, where the program will only select those tables which it finds appropriate to include in the query. The proposed methodology has been implemented as a tool called NLSQL, which will deliver a complete solution for producing SQL queries from natural language. Unlike other tools, this program will automatically retrieve the necessary data from the database. Users will be required to provide database credentials. The program will also run the query and display the results.

Limitations identified in our system were that tables which are directly joint can only be used for JOIN queries and we are working on expanding this work. Apart from this, the algorithm is not yet optimized for dates and therefore can be a prospective work hereafter. In the future, to effectively identify certain tokens, the system must also check for word combinations (which we refer to as the n-grams). The adapter pattern is used to effectively generate the final SQL query (see Sect. 3). The system currently only has a MySQL driver. Future work also includes creating more drivers for other relational databases.

References

1. Finegan-Dollak C et al (2018) Improving text-to-SQL evaluation methodology. arXiv preprint arXiv:1806.09029
2. Baik C, Jagadish HV, Li Y (2019) Bridging the semantic gap with SQL query logs in natural language interfaces to databases. In: 2019 IEEE 35th international conference on data engineering (ICDE). IEEE
3. Yu T et al (2018) TypeSQL: knowledge-based type-aware neural text-to-SQL generation. arXiv preprint arXiv:1804.09769
4. Lin K et al (2019) Grammar-based neural text-to-SQL generation. arXiv preprint arXiv:1905.13326
5. Kim H et al (2020) Natural language to SQL: where are we today? Proc VLDB Endow 13(10):1737–1750
6. Gan Y et al (2021) Natural SQL: making SQL easier to infer from natural language specifications. arXiv preprint arXiv:2109.05153
7. Eriksson S (2019) Comparing natural language processing to structured query language algorithms

8. Kate A et al (2018) Conversion of natural language query to SQL query. In: 2018 second international conference on electronics, communication and aerospace technology (ICECA). IEEE
9. Yaghmazadeh N et al (2017) SQLizer: query synthesis from natural language. Proc ACM Program Lang 1(OOPSLA):1–26
10. Popescu A-M, Etzioni O, Kautz H (2003) Towards a theory of natural language interfaces to databases. In: Proceedings of the 8th international conference on intelligent user interfaces
11. Kaur S, Bali RS (2012) SQL generation and execution from natural language processing. Int J Comput Bus Res 2229–6166
12. Singh G, Solanki A (2016) An algorithm to transform natural language into SQL queries for relational databases. Selforganizology 3(3):100–116
13. Banerjee T, Bhattacharjee U, Jansi KR (2021) Natural language querying and visualization system. In: Recent trends in intensive computing. IOS Press, pp 607–614

Design and Verification of Five-Stage Pipelined Architecture Using SV-Verification Methodology Module

Sidhant Priyadarshi, Saroja V. Siddamal, Apoorva Narode, Vaishnavi Chillal, and Vikash Sonar

Abstract RISC V architecture is finding its importance with semiconductor industry and academia. With the availability of open instruction, set design of the processor is possible. The RTL needs an extensive verification. Simulation-based methods are rampant, but exhaustive test generations are required. The papers reports design and System Verilog verification of the five-stage RISC V processor. Mentor Questa simulator is used to verify the design. The code coverage reported is 80%.

Keywords RISC V instruction set · Pipelined processor · Code coverage · Verification · System Verilog verification

1 Introduction

Over the past 10 years, people are using many electronic gadgets which perform the tasks efficiently; it's just because of processor present in the gadgets. Everyone wants their device to be smart and efficient enough with good performance. The designer can build a single-stage processor which can meet the needs of the user but is time-consuming process, the designer can upgrade the processor to five-stage pipelined architecture which works more efficient than single-stage pipeline in terms of clock rate, performance, etc. Pipelining typically reduces the processor's cycle time and increases the throughput of instructions. Multitasking is the best feature of contemporary computers. An unpipelined instruction cycle, also known as a fetch-execute cycle, is a slower method of processing instructions that increases in time. Unprecedented gains in size and performance are possible with pipelined computer

S. Priyadarshi · S. V. Siddamal (✉) · A. Narode · V. Chillal · V. Sonar
School of ECE, KLE Technological University, Hubballi, India
e-mail: sarojavs@kletech.ac.in

S. Priyadarshi
e-mail: sidhant1922@gmail.com

A. Narode
e-mail: anapoorva08@gmail.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_34

architecture. The architecture aims to understand the underlying workings of operating system instruction pipelining to improve computer architecture. To enhance the overall performance of CPU using five-stage pipelined architecture is designed and verified using verification methodology.

2 Related Work

To overcome the debugging challenge, the authors went through different architectures RISC and CISC. The author has concentrated on the RV32I basic integer instruction set in the paper [1]. RISC V is an open Standard Instruction Set architecture designed for academic and industrial applications. The five-stage pipelining soft processor is approximately 37% faster than the single-cycle design. The design can be improved by raising the maximum frequency of the suggested processor core which is given in paper [2].

The RISC V architecture and the instruction set to be designed are discussed in [3]. Pipelined processors are used to process multiple instructions in one clock cycle. This allows for the simultaneous execution of several instructions, which require a larger number of registers. All RISC instructions have the same length since instructions are fixed-size, and pipelining is simple as all instructions are executed in a constant amount of time. A one-click RISC processor is constructed using hardwired control unit [4]. Top manufacturers of microcontrollers and processors, such as TI, Freescale, Microchip, Renasis have a variety of ARM core-based controllers. Debugging is challenging due to the scheduling of instructions, which severely affects the processor's performance. Despite recent developments, the ARM core still rules the processor/controller industry as explained by authors in paper [5]. To overcome the debugging challenge the authors went through different architectures RISC and CISC. RISC is a series of straightforward processes and CISC is a small number of intricate instructions for a single activity. It is used for MIPS to improve the throughput of the entire workload by retrieving instructions from memory using a computer and enhancing PC. All instructions in a dynamically scheduled pipeline proceed through the issue stage in sequence, but instructions may be skipped in the second stage, which causes them to enter execution out of order. Dynamic scheduling is justified by rearranging the sequence in which instructions are executed in order to decrease delays, while preserving data flow [1]. Increasing the capability of verification tools at the same rate as designers are expanding the scope of the issue will be difficult, as it will lead to the bottleneck by how efficiently logical decision-making processes can be computed. Strategies may still be applicable for partial verification or in a hybrid simulation/verification tool if decision procedures cannot be expanded quickly enough [6]. The system-on-chip (SOC) design technique has been criticized for lack of a functional coverage metric. To address this, the framework for test creation and functional coverage estimates for pipelined processors is presented in the research paper [7] using Specman Elite. This methodology is implemented using a very-long instruction word (VLIW) Duplex architecture and the progress of the

verification process is tracked using a cover estimate. The metrics used to estimate coverage include register read/write, operation execution, and pipeline execution. In the research paper [8], the verification environment is improved, while the functionality of the design is verified using all possible situations. The following points will be taken into account based on the project's requirements.

- Reusability of code,
- Creation of stimuli, and
- Use the environment to verify the DUT.

The research paper [8] focuses on the verification environment and functionality of the design for the One Master One Slave configuration. The design is verified using System Verilog (SV) and constraint-based randomization techniques.

Code coverage is the number of code occurrences that occur during the verification process. It is used to identify erroneous routes and inactive code for a certain set of inputs. The verification tool QuestaSim computes code coverage and can be increased up to 100% by modifying the restrictions and test strategy [9]. The verification environment is not transferable to the verification of other designs [9].

The book [10] offers advice on verification methodology and reuse, and includes strategies for constraint-random stimulus generation, coverage-driven verification, assertion-based verification, formal analysis, and system-level verification. System Verilog also includes an industry standard for tying RTL design, test benches, assertions, and coverage together in a comprehensive and cohesive way.

The authors in this paper have designed a RV32I, five stages pipelined architecture and verified using SV. The rest of the paper is organized as follows. Section 3 discusses on the design of five-stage pipelined architecture. SV-verification methodology is discussed in Sect. 4. In Sect. 5 results are discussed and finally the conclusion.

3 Architecture for Five-Stage Pipelining of RV32I

To build a testing and verification environment, it's required to have a design under test (DUT). In this paper five-stage RV32I is used as DUT. The use case tested is a mod-6 up counter with BNE instruction, namely fetch, decode, execute, memory write, write back. In the fetch stage, the instructions are fetched from instruction memory that will be placed in Instruction Register (IR). In the decode stage, instruction will be decoded as of RV32I mnemonics, example 7'b0110011 for ADD operation, 7'b0110011 for MUL operation, and 7'b1100011 for Branch not equal operation. If instruction is other than 32 bit, the instructions are signed extended by zero padding. In execute stage, instructions will be executed to provide given operation for arithmetic logical unit [ALU]. If the instruction is branch, PC will directly go to fetch phase. If not it will go to next stage of pipelining. In memory stage, load/store instructions are performed. In the write back stage, result will be stored in the destination locations/Registers. Example: ADD X2, X1, X3: in fetch stage, above

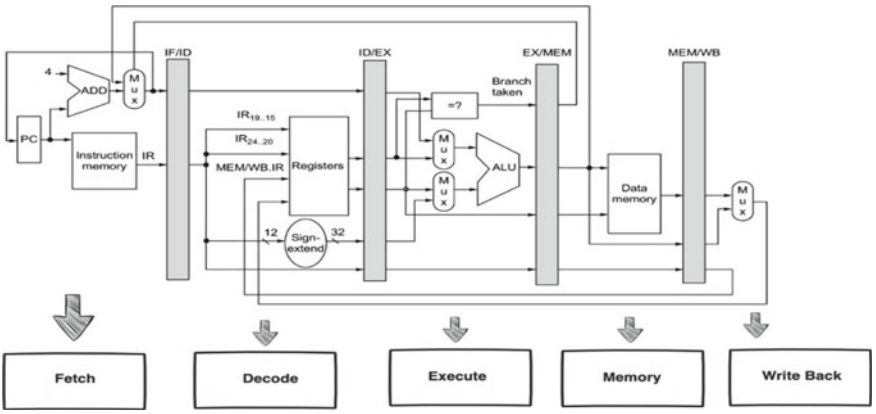


Fig. 1 Five-stage pipeline architecture

instruction is stored in form of binary code (0x00308133) in ROM is then fetched and stored in Instruction Register (IR) and Program Counter (PC) is incremented to next location. In decode stage, data of x2 is 2, x1 is 1 and x3 is 3 and opcode will be decoded. In execute stage, based on opcode, funct3 and funct7 instruction is executed in ALU and gives output as 4. In memory stage, output from ALU will be passed to next stage as add instruction does not have any memory operations. In write back stage, output from ALU is stored to destination register x2 (Fig. 1).

4 System Verilog and Universal Verification Methodology

Design verification is the most important aspect of the product development process, consuming as much as 80% of the total product development time. The intent is to verify that the design meets the system requirements and specifications. Here, design is verified with two programs having some sets of instructions loaded performing hex to decimal and mod-6 up counter program. The figure shows the SV-VMM architecture, which is a framework for creating scalable and modular verification environments for complex designs.

It is composed of four main layers: the testbench layer, verification components layer, verification infrastructure layer, and DUT layer. The testbench layer is responsible for driving the design under test and verifying its outputs. The verification infrastructure layer manages the testbench and verification components. Finally, the DUT layer contains the design being verified. The SV-VM architecture promotes modularity and easy reuse of verification components across projects. Layered testbench architecture as shown in Fig. 2 consists of transaction class, driver class, monitor class, scoreboard class, environment class, interface, etc. Transaction class will transact instructions to the DUT, it also contains memory units for transaction

for future process. Sequencer class will sequence the inputs for next iterations and puts data in mailbox to send data to driver. Driver class will drive the inputs from transaction to interface for transferring to DUT. Interface class is going to create interface between DUT and driver, as to communicate between blocks and also allowing smooth refinement from abstract system level through successive steps down to lower RTL. Monitor class samples the interface signals and changes the activity at the signal level to the transaction level. Sends the sampled transaction through mailbox to scoreboard. Scoreboard class compares the read data with the data in local memory if the transaction type is "read". If the transaction type is "write" the wdata will be written to local memory. Environment class encapsulates all components and objects of testbench and helps in starting communication between DUT and transactor. SV provides a framework for coverage-driven verification (CVD). CVD combines automatic test generation, self-checking test benches, and coverage metrics to significantly reduce the time spent verifying a design. Firstly a smart test bench is created that generates legal stimuli and sends it to the DUT. Coverage monitors are added to the environment to measure progress and identify non-exercised functionality. Checkers are added to identify undesired DUT behavior. Simulations are launched after both the coverage model and test bench have been implemented. Verification then can be achieved.

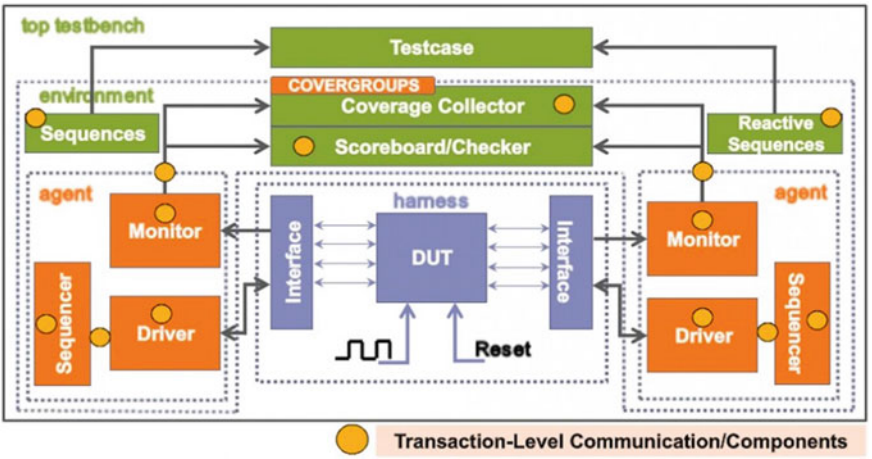


Fig. 2 Testbench architecture for DUT

Table 1 Hardware utilization when implemented in Spartan 6

Parameter	Value
Delay (ns)	2.822
No. of FF's	722
No. of LUT's	1102
Clock frequencies (MHz)	340

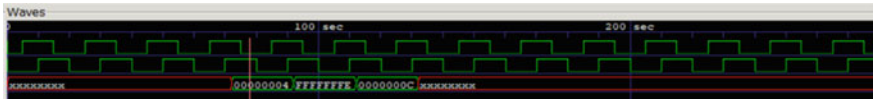


Fig. 3 Functional simulation of five-stage pipelined architecture

5 Results and Discussion

To verify the working of architecture, two kinds of verification are performed. First, functional simulation is done to check RTL working in Xilinx ISE. Functional verification is done next to verify corners of architecture with some sets of instructions in using Mentor Questa simulator.

5.1 Five-Stage Pipeline Processor Synthesized Results

The architecture is designed and synthesized using Xilinx 14.2 ISE. Table 1 gives the hardware requirements when implemented in Spartan 6 FPGA. Figure 3 represents the simulation output of five-stage pipelined RV32I ISA. There are two clocks used with a phase difference of 180° named as clk1 and clk2, respectively. Instruction to add 1 and 3 (add x2, x1, x3) is fetched and executed at fourth cycle of clk1 which is next cycle to write back stage results in 4. Second instruction of subtracting 1 and 3 (sub x2, x1, x3) is fetched and results in -2 at fifth cycle of clk1. Third instruction of multiplying 2 and 3 (mul x1, x2, x3) is fetched and results in 12 where x2 is updated as 4 from first instruction as shown in Fig. 3.

5.2 SV-VMM Verification Results

In this section, the authors have explained the verification results for five-stage pipeline architecture, which includes instruction monitoring, output matching results, and code coverage report. For twenty-four instructions, testbench gives coverage report of 80%. Table 2 gives the set of instruction tested on the DUT. Two use cases

Table 2 Set of instruction tested on the DUT

Instructions	Machine language
ADDI x1, 010 [x0]	00001111111100000000000010010011
ADD x2, x16, x0	00000001000000000000000100110011
DIV x2, x1, x2	00000010001000001100000100110011
ADD x3, x2, x0	00000000001000000000000110110011
ADD x2, x16, x0	00000001000000000000000100110011
DIV x2, x1, x2	00000010001000001100000100110011
ADD x1, x5, x0	00000000010100000000000010110011
ADD x3, x0, x0	00000000000000000000000110110011
ADD x2, x1, x0	00000000000100000000000100110011
Count: ADD x3, x1, x0	00000000000100000000000110110011
SUB x1, x1, x2	01000000001000001000000010110011
BNE x1, x0, count	00000100000000001001010001100011

are tested, first is to convert hexadecimal into decimal and second is to perform mod-6 up counter.

Figure 4 shows the various stages of the instruction passing through the layered testbench from driver to scoreboard. If there is a mismatch in instruction at driver and scoreboard, it's verified as mismatch as shown by black arrow.

5.3 Coverage Report

Figure 5 shows a coverage report for twelve instructions used for implementing two programs which are hex-to-decimal converter and mod-6 up counter. Twelve instructions used for implementing two programs are hex to decimal converter and mod-6 up counter. Out of twelve, three instructions are branched five times; eventually twenty-four instructions are transmitted and matched with DUT's instructions. For twenty-four instructions, testbench gives coverage report of 80%.

```
 5000 [DRIVER] : IR = 00001111111100000000000010010011
----------------------------
----------------------------
 5000 [Output Monitor] : IR = 00001111111100000000000010010011
----------------------------
----------------------------
15000 [DRIVER] : IR = 00000001000000000000000100110011
----------------------------
----------------------------
15000 [Input Monitor] : IR = 00001111111100000000000010010011
----------------------------
15000 [SCOREBOARD] : IR = 00001111111100000000000010010011 Matched!
15000 [SCOREBOARD] : MissMatched! should be 16 or 5 but it is 00000000000000000000000000000000
----------------------------
20000 [Output Monitor] : AluOut = 00000000000000000000000000000000
----------------------------
20000 [SCOREBOARD] : IR = 00001111111100000000000010010011 Matched!
20000 [SCOREBOARD] : MissMatched! should be 16 or 5 but it is 00000000000000000000000000000000
----------------------------
25000 [DRIVER] : IR = 00000010001000001100000100110011
----------------------------
----------------------------
25000 [Output Monitor] : IR = 00000010001000001100000100110011
----------------------------
----------------------------
35000 [DRIVER] : IR = 00000000001000000000000110110011
----------------------------
----------------------------
35000 [Input Monitor] : IR = 00000010001000001100000100110011
----------------------------
35000 [SCOREBOARD] : IR = 00000010001000001100000100110011 Matched!
35000 [SCOREBOARD] : MissMatched! should be 16 or 5 but it is 00000000000000000000000000000000
```

Fig. 4 Layered testbench verification of instruction

Fig. 5 Code coverage report

```
# ------------------------------------
#                    0 [GENERATOR] : IR
# ------------------------------------
# ------------------------------------
# Coverage =          80 %
# ------------------------------------
# ------------------------------------
#                 5000 [DRIVER] : IR =
# ------------------------------------
```

6 Conclusion

The paper has presented the design and verification of five-stage pipelined architecture using RV32I ISA. RTL is designed using Xilinx ISE. Functional simulation is performed on Icarus-Verilog and the design is verified using SV-VMM (layered testbench architecture) on QuestaSim verification with coverage of 80%.

References

1. Olanrewaju RF, Fajingbesi FE, Junaid SB, Alahudin R, Anwar F, Pampori BR (2017) Design and implementation of a five stage pipelining architecture simulator for RiSC-16 instruction set. Indian J Sci Technol 10(3):1–9
2. Phangestu AE, Mujiono IT, Kom MI et al (2022) Five-stage pipelined 32-bit RISC V base integer instruction set architecture soft microprocessor core in VHDL. In: 2022 international seminar on intelligent technology and its applications (ISITIA). IEEE, pp 304–309
3. Asanovic K, Waterman A (2019) The RISC V instruction set manual volume I: unprivileged ISA document version 20191213
4. Java at point. The-classic-five-stage-pipeline-for-RISC-processor-with-diagram
5. Ghosh A, Kumari S, Pan N, Kumari S, Kumari A, Chowdhury TR (2019) Analysis of 5 stage pipelined operations of arm 32/64 bit. Int J Adv Res Ideas Innov Technol 5:432–435
6. Burch JR, Dill DL (1994) Automatic verification of pipelined microprocessor control. In: International conference on computer aided verification. Springer, pp 68–80
7. Mishra P, Dutt N, Kashai Y (2004) Functional verification of pipelined processors: a case study. In: Fifth international workshop on microprocessor test and verification (MTV'04). IEEE, pp 79–84
8. Polsani P, Priyanka V, Sai P et al (2020) Design verification of serial peripheral interface (SPI) protocol. Int J Recent Technol Eng (IJRTE) 8(6). ISSN: 2277-3878
9. Giridhar P, Choudhury P (2019) Design and verification of AMBA AHB. In: 2019 1st international conference on advanced technologies in intelligent control, environment, computing communication engineering (ICATIECE). IEEE, pp 310–315
10. Bergeron J, Cerny E, Hunter A, Nightingale A (2006) Verification methodology manual for System Verilog, vol 1. Springer

Prediction of Heart Disease with Machine Learning Methods: Comparison and Analysis

Bharat Singh, Savita Nandan, Sahil Kumar, and Nidhi Kushwaha

Abstract Heart stroke is a serious medical condition that can have devastating consequences if not detected early. Lately, scientists have explored the chance of utilizing AI calculations to foresee a singular's gamble of having a coronary episode. This has led to the development of various predictive models that use a range of risk factors, such as age, blood pressure, and cholesterol levels, to make accurate predictions. In this paper, three models have been applied for early detection. By using these models, healthcare providers can identify individuals who are at high risk of having a heart stroke and provide them with appropriate interventions to prevent or mitigate its effects. The machine learning model's application in heart stroke prediction has the capability to ameliorate patient condition and reduce healthcare costs.

Keywords Machine learning · Heart disease · Prediction · Precision · Recall · Confusion matrix

1 Introduction

Around the world, cardiovascular diseases (CVDs) represent 17.9 million passings yearly and are the main source of death. Cardiovascular diseases (CVDs) include rheumatic heart disease, coronary heart disease, cerebrovascular disease, and others. Respiratory failures and strokes represent in excess of four out of each and every

B. Singh (✉) · S. Nandan · S. Kumar · N. Kushwaha
Indian Institute of Information Technology, Ranchi, Ranchi, India
e-mail: bsingh@iiitranchi.ac.in
URL: https://www.iiitranchi.ac.in

S. Nandan
e-mail: savita.btech.cs19@iiitranchi.ac.in

S. Kumar
e-mail: sahil.btech.cs19@iiitranchi.ac.in

N. Kushwaha
e-mail: nidhi@iiitranchi.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_35

five passings from cardiovascular illness, and 33 of these passings happen rashly in individuals younger than 70 [1]. Heart stroke, also known as a stroke, is a medical condition that occurs due to the disruption of blood flow to the brain. This disruption can be caused by a blockage or rupture of a blood vessel, and can result in brain damage or even death [2, 3]. In fact, heart stroke ranks as the second most common cause of death globally and is also a significant contributor to long-term disability [4, 5]. Lately, scientists have explored the chance of utilizing AI calculations to foresee a singular's gamble of having a coronary episode. These algorithms can analyze a range of risk factors, such as blood pressure, age, smoking history, and cholesterol levels to make accurate predictions about an individual's likelihood of having a heart stroke [6–9].

By identifying individuals who are at high risk of having a heart stroke, healthcare providers can provide appropriate interventions to prevent or mitigate its effects. The use of machine learning algorithms in heart stroke prediction has the potential to significantly improve patient outcomes and reduce healthcare costs. By identifying individuals who are at high risk of having a heart stroke, healthcare providers can intervene early to prevent the onset of the condition or minimize its effects [6, 10]. Furthermore, machine learning algorithms can help healthcare providers to better allocate resources by identifying high-risk patients who require more intensive monitoring and care. Overall, the use of machine learning algorithms in heart stroke prediction is an exciting and promising area of research that has the potential to make a significant impact on public health.

2 Machine Learning Algorithms Used

2.1 Logistic Regression (LR)

Logistic regression is a factual strategy used to display the likelihood of a paired result, such as the presence or absence of a certain condition or the occurrence of a particular event [11]. It works by estimating the probability of the outcome based on one or more independent variables, such as age, gender, and medical history. The model then calculates the odds of the outcome occurring, and uses this information to predict the likelihood of the outcome in new cases [2, 9]. It is a popular method because it is relatively simple to implement and interpret, and can be used with both continuous and categorical independent variables. However, it is important to ensure that the assumptions of logistic regression are met, such as linearity, independence, and absence of multicollinearity. Overall, logistic regression is a powerful statistical tool that can help researchers and healthcare providers to make predictions about the likelihood of binary outcomes, and inform decisions about diagnosis, treatment, and prevention.

2.2 Support Vector Machine (SVM)

SVM tracks down the most ideal decision limit that expands the margin between two classes by transforming the input data into high-dimensional space. It aims to locate the hyperplane with the greatest margin of separation between the two classes, thereby improving classification accuracy. It is especially helpful in dealing with high-dimensional datasets and can handle both linearly and non-linearly separable data by using different kernels [12]. It is also less prone to overfitting than other machine learning algorithms. SVM is also computationally efficient and scales well with large datasets. However, SVM can be sensitive to the choice of kernel function and requires careful selection of hyperparameters. SVM can also be sensitive to imbalanced datasets, where the number of examples in each class is not equal, which can lead to biased classification [11].

2.3 Random Forest (RF) Classifier

It is a strategy for outfit learning in which different choice trees are built and their results are consolidated to make expectations. Every decision tree in the Random Forest is based on an irregular subset of the preparation information and an irregular subset of the highlights, which lessens the gamble of overfitting and works on the speculation of the model [2, 13]. The Random Forest algorithm then aggregates the output of each decision tree to make a final prediction. One of the main advantages of the Random Forest Classifier is its ability to handle complex datasets with high-dimensional features. Random Forest Classifier is also relatively easy to implement and interpret, making it a popular choice for machine learning practitioners. Inclusively, RF Classifier is a prevailing and versatile ML algorithm that can handle complex datasets, reduce the risk of overfitting, and achieve high accuracy. Its versatility and ease of use have contributed to its widespread adoption in various industries [14, 15].

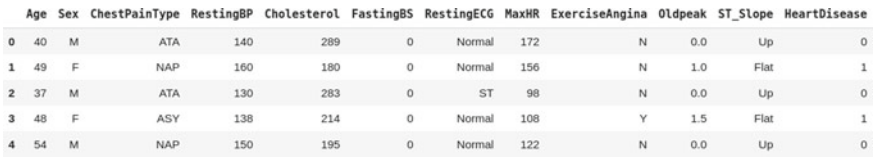
3 Dataset Description

The dataset that we have used in our experiment was obtained from the Kaggle website [16] which including cholesterol level, maximum HR achieved, age, sex, chest pain type, resting ECG results, resting BP, exercise-induced angina, fasting blood sugar, ST slope, and Old Peak. The target label for the model is stroke, with the ID attribute excluded from the analysis. A summarized description of heart stroke data has been given in Table 1.

Table 1 summarize the data analyzed in work. The features were presented separately because the data set contains both numerical and categorical variables.

Table 1 Summary and distributions of heart stroke data

Feature name	Min	Max	Std. dev.	Median	Mean
Age	28	77	9.50	54	52.86
Oldpeak	0	6.2	1.07	0.50	0.90
Maximum heart rate	69	202	24.54	140	140
Cholesterol	85	603	59	237	244
Resting blood pressure	92	200	17.28	130	133



	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
0	40	M	ATA	140	289	0	Normal	172	N	0.0	Up	0
1	49	F	NAP	160	180	0	Normal	156	N	1.0	Flat	1
2	37	M	ATA	130	283	0	ST	98	N	0.0	Up	0
3	48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat	1
4	54	M	NAP	150	195	0	Normal	122	N	0.0	Up	0

Fig. 1 The heart stroke dataset screenshot

In contrast, the numerical features were depicted using minimum, maximum, median, mean, and standard deviation. Also, in Fig. 1 an instance of dataset has been shown.

4 Proposed Framework Description

This paper proposes a model to predict the likelihood of an individual experiencing a heart stroke depend on different input attributes such as age, gender, smoking status, work type. Flowchart of the proposed model has given as Fig. 2. This heart data is trained on three machine learning algorithms namely logistic regression, SVM, and Random Forest, and their execution are evaluated to determine the best method for accurate prediction. The comparative analysis of each algorithm is presented through a plotted graph of the obtained accuracy, as depicted in Fig. 5. The dataset undergoes pre-processing to eliminate null and duplicate values. The dataset is partitioned into separate training and testing parts, which are then used as inputs to several classification models in order to obtain predictions. The accuracy of the models is evaluated using the confusion matrix. Additionally, the model is tested with a custom input value to determine its accuracy.

4.1 Data Pre-processing and Encoding

To enhance the accuracy of the model, it is crucial to eliminate any null values in the attribute, which could impede the model's performance. Therefore, we have acquired a pre-processed dataset that is free of null values, and we have converted

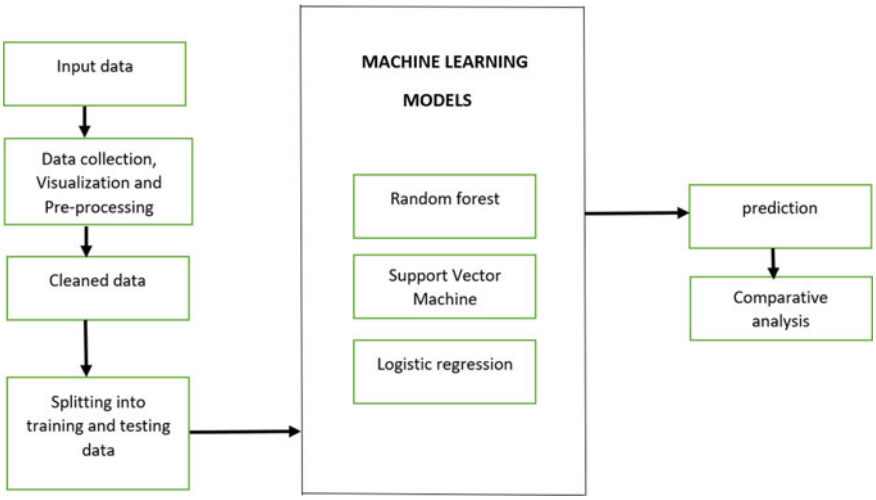


Fig. 2 Flowchart of the proposed model

	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
0	40	1	1	140	289	0	1	172	0	0.0	2	0
1	49	0	2	160	180	0	1	156	0	1.0	1	1
2	37	1	1	130	283	0	2	98	0	0.0	2	0
3	48	0	0	138	214	0	1	108	1	1.5	1	1
4	54	1	2	150	195	0	1	122	0	0.0	2	0

Fig. 3 Table shows the labeled encoded data

the categorical inputs into numerical inputs by applying the 'LabelEncoder' technique. Since the methods that we are applying for training can only be performed on numerical inputs that undergo attribute standardization, text data (such as gender: male or female) have been label encoded into numeric values (0, 1) to establish correlations between different attributes. An instance of dataset depicted in Fig. 3 after the encoding and pre-processing.

4.2 Data Correlation

Correlation Matrix: It is a tabular representation of correlation coefficients between different variables. **Correlation Coefficient**: It is a mathematical metric used to determine the degree of linear association between two variables (if two variables are highly co-related then the impact of both on the target variable will be the same). A correlation among all the features that was present in the data has been studied and shown in Fig. 4.

	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
Age	1.000000	0.055750	-0.077150	0.254399	-0.095282	0.198039	-0.007484	-0.382045	0.215793	0.258612	-0.268264	0.282039
Sex	0.055750	1.000000	-0.126559	0.005133	-0.200092	0.120076	0.071552	-0.189186	0.190664	0.105734	-0.150693	0.305445
estPainType	-0.077150	-0.126559	1.000000	-0.020647	0.067880	-0.073151	-0.072537	0.289123	-0.354727	-0.177377	0.213521	-0.386828
RestingBP	0.254399	0.005133	-0.020647	1.000000	0.100893	0.070193	0.022656	-0.112135	0.155101	0.164803	-0.075162	0.107589
Cholesterol	-0.095282	-0.200092	0.067880	0.100893	1.000000	-0.260974	-0.196544	0.235792	-0.034166	0.050148	0.111471	-0.232741
FastingBS	0.198039	0.120076	-0.073151	0.070193	-0.260974	1.000000	0.087050	-0.131438	0.060451	0.052698	-0.175774	0.267291
RestingECG	-0.007484	0.071552	-0.072537	0.022656	-0.196544	0.087050	1.000000	-0.179276	0.077500	-0.020438	-0.006778	0.057384
MaxHR	-0.382045	-0.189186	0.289123	-0.112135	0.235792	-0.131438	-0.179276	1.000000	-0.370425	-0.160691	0.343419	-0.400421
erciseAngina	0.215793	0.190664	-0.354727	0.155101	-0.034166	0.060451	0.077500	-0.370425	1.000000	0.408752	-0.428706	0.494282
Oldpeak	0.258612	0.105734	-0.177377	0.164803	0.050148	0.052698	-0.020438	-0.160691	0.408752	1.000000	-0.501921	0.403951
ST_Slope	-0.268264	-0.150693	0.213521	-0.075162	0.111471	-0.175774	-0.006778	0.343419	-0.428706	-0.501921	1.000000	-0.558771
HeartDisease	0.282039	0.305445	-0.386828	0.107589	-0.232741	0.267291	0.057384	-0.400421	0.494282	0.403951	-0.558771	1.000000

Fig. 4 Co-relation matrix of various attributes on heart stroke dataset

We can reduce the attributes by removing one of the two highly correlated variables as they both will have nearly same impact on the target variable by using the concept of dimensionality reduction.

5 Results and Analysis

In this section, we present the outcomes achieved from the implementation of Random Forest, Logistic Regression, and SVM. The evaluation of algorithmic performance was based on several metrics, including Accuracy score, Precision (P), Recall (R), and $F1$-measure. Specifically, the Precision metric quantifies the accuracy of positive predictions, while Recall assesses the number of actual positives that are correctly predicted. $F1$-measure test's accuracy (Fig. 5).

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP})$$

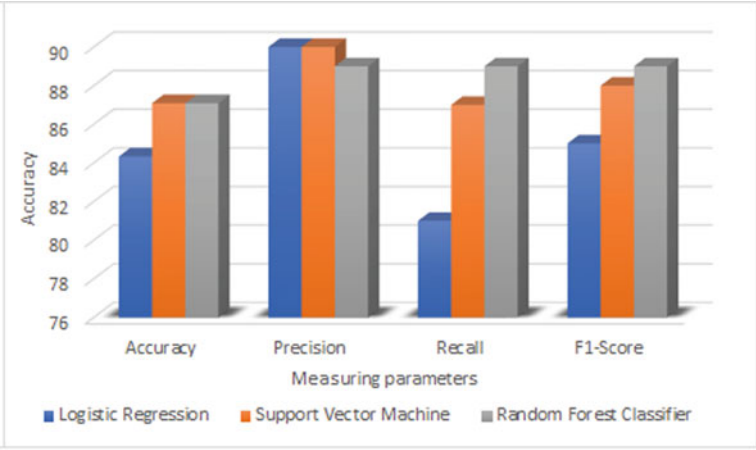
$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN})$$

$$F1\text{-Measure} = 2 * \text{Precision} * \text{Recall}/(\text{Precision} + \text{Recall})$$

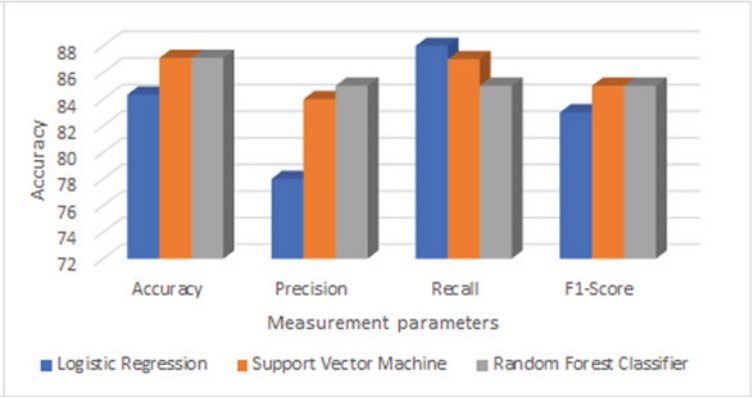
$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{Total number of samples})$$

$$\text{Support of class} = \text{Count of actual values of class}$$

- TP (True positive): It depicts the scenarios in which the model correctly predicted the positive class.
- FP (False positive): It depicts the scenarios in which the model predicted a positive class, but the actual class was negative.



(a) Results for the target class Yes

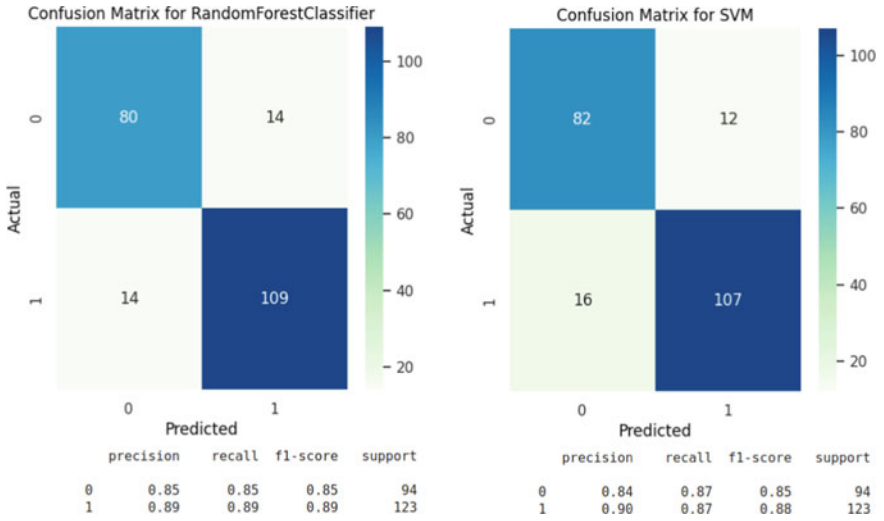


(b) Results for the target class No

Fig. 5 Performance of models for accuracy, recall, precision, and $F1$-measure of three methods

– TN (True negative): It depicts the scenarios in which the model correctly predicted the negative class.
– FN (False negative): It depicts the scenarios in which the model predicted a negative class, but the actual class was positive.

The performance metrics mentioned above are derived from the confusion matrix, which yields an assessment of the model's performance as a whole. A confusion matrix is a performance evaluation tool commonly used in machine learning to assess the accuracy of a classification model. It is a square matrix that summarizes the number of true positives, true negatives, false positives, and false negatives produced by a binary classification algorithm. From Fig. 6 we find that the SVM model is better than the other two with respect to type II error. Furthermore, the results that we achieved in our study has depicted in Table 2.



(a) Random Forest (b) Support Vector Machine

Fig. 6 Confusion matrix of SVM and Random Forest for heart stroke data

Table 2 Accuracies and others measurements achieved using LR, SVM, and RF classifier algorithms

Algorithm	Precision	Recall	$F1$-score	Accuracy	Sensitivity
Logistic regression	0.86	0.86	0.86	84.21	0.88
Support vector machine	0.84	0.89	0.86	84.56	0.87
Random Forest classifier	0.86	0.88	0.87	85.36	0.85

6 Conclusion

Heart stroke prediction is a crucial task that can help to prevent and manage cardiovascular diseases, which are among the main sources of death around the world. Machine learning algorithms such as LR, SVM, and RF Classifier have shown promising results in predicting heart strokes and identifying high-risk individuals. The most effective algorithm was RF classifier with an accuracy of 85.26% among the above three models. By leveraging the power of machine learning, healthcare professionals can improve the accuracy and efficiency of heart stroke prediction, ultimately leading to better patient outcomes and a healthier population.

References

1. World Health Organization (2023) https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1
2. Ozcan M, Peker S (2023) A classification and regression tree algorithm for heart disease modeling and prediction. Healthc Anal 3:100130. ISSN: 2772-4425. https://doi.org/10.1016/j.health.2022.100130
3. Ahsan MM, Siddique Z (2022) Machine learning-based heart disease diagnosis: a systematic literature review. Artif Intell Med 102289
4. Mittal RA (2017) Increasing heart attacks in young Indians, 2017. https://timesofindia.indiatimes.com/life-style/health-fitness/every-heart-counts/increasing-heart-attacks-inyoung-indians/articleshow/56295257
5. Anbuselvan P (2020) Heart disease prediction using machine learning techniques. Int J Eng Res Technol (IJERT) 9(11). ISSN: 2278-0181. http://www.ijert.org
6. Xavier A (2020) Heart disease prediction using machine learning and data mining technique. Int J Eng Res Technol (IJERT). ISSN: 2278-0181. NTASU-2020 conference proceedings. http://www.ijert.org
7. Kasabe R (2020) Heart disease prediction using machine learning. Int J Eng Res Technol (IJERT) 9(08). ISSN: 2278-0181. http://www.ijert.org
8. Limbitote M (2020) A survey on prediction techniques of heart disease using machine learning. Int J Eng Res Technol (IJERT) 9(06). ISSN: 2278-0181. http://www.ijert.org
9. Shah D, Patel S, Bharti SK (2020) Heart disease prediction using machine learning techniques. SN Comput Sci 1:345
10. Dwivedi AK (2018) Performance evaluation of different machine learning techniques for prediction of heart disease. Neural Comput Appl 29:685–693
11. Kumar NK (2020) Analysis and prediction of cardio vascular disease using machine learning classifiers. In: 6th international conference on advanced computing and communication systems (ICACCS)
12. Govindarajan P, Soundarapandian RK, Gandomi AH, Patan R, Jayaraman P, Manikandan R (2020) Classification of heart stroke disease using machine learning algorithms. Neural Comput Appl 32(3):817–828
13. Gavhane A (2018) Prediction of heart disease using machine learning. In: 2nd international conference on electronics, communication and aerospace technology (ICECA)
14. Amini L, Azarpazhouh R, Farzadfar MT, Mousavi SA, Jazaieri F, Khorvash F, Norouzi R, Toghianfar N (2013) Prediction and control of heart stroke by data mining. Int J Prev Med 4(Suppl 2):S245–S249
15. Latha CBC, Jeeva SC (2019) Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. Inform Med Unlocked 16:100203. ISSN: 2352-9148. https://doi.org/10.1016/j.imu.2019.100203
16. https://www.kaggle.com/datasets/bunyaminpolat/heart-stroke-prediction. Accessed 04 Jan 2023

A Threat and Risk Analysis of Internet Communication Technology Use

Kanika Panwar, Vikas Sihag, Gaurav Choudhary, and Nicola Dragoni

Abstract The Internet is abuzz with opportunities and challenges for youth wherein challenges assume the form of threats and risks at times. Online grooming is one such risk faced by youth that can be detrimental to their physical, mental, and sexual well-being. Using suitable concepts from risk studies, information security and cyber sociology, this study identifies the threats and vulnerabilities faced by the youth. Using exploratory research methods we analyze threats such as exposure to inappropriate content, cybercrimes, the presence of sexual predators, cyber bullying, and cyber stalking. A detailed taxonomy of vulnerabilities identified in the literature is also evolved and interpreted. Furthermore, we perform risk analysis and design a risk assessment matrix to evaluate threats on the basis of their impact and likelihood.

Keywords Security · Internet · Vulnerabilities · Online Threats · Risk Analysis

1 Introduction

The Internet offers youth new opportunities for self-discovery, accessing health information, and building relationships. However, it also poses risks, including online grooming. Research shows increased levels of online risks faced by young people, challenging previous views that portrayed the Internet as both empowering and potentially harmful to their well-being. Concerns surrounding online social media and networking among youth are growing due to its unmediated nature [10, 16]. Studies reveal that young people are exposed to unsolicited sexual content, along with various negative experiences such as unwanted contact, social exclusion, threats, rep-

K. Panwar · V. Sihag (✉)
Sardar Patel University of Police, Security and Criminal Justice, Jodhpur, India
e-mail: vikas.sihag@policeuniversity.ac.in

K. Panwar
e-mail: kanika@policeuniversity.ac.in

G. Choudhary · N. Dragoni
DTU Compute, Technical University of Denmark, Copenhagen, Denmark
e-mail: ndra@dtu.dk

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_36

Table 1 Comparative review of literature of online threats faced by youth

Works	Inappropriate content	Pornographic content	Online strangers	OSS	Cyber bullying	Privacy and security threats	Online addiction	Parent-child relation
[9]			✓	✓		✓	✓	✓
[18]	✓			✓	✓			
[13]	✓		✓	✓	✓			
[23]		✓		✓	✓	✓	✓	
[7]		✓	✓		✓			
[4]		✓		✓				
[20]	✓	✓	✓	✓			✓	
[2]		✓	✓	✓			✓	
[19]	✓	✓	✓		✓	✓	✓	

utation damage, fraud, viruses, and lack of permission [1]. The prevalence of sexting and Online Sexual Solicitation has been extensively studied. The increasing interest in online risks against youth highlights the need for understanding the potential consequences. Table 1 provides a list of major research on online threats faced by youth.

Various studies have highlighted the online threats and risks faced by youth, but there is a lack of clarity regarding their severity and vulnerabilities [2]. One prominent risk mentioned is online sexual solicitation (OSS), which is associated with a high level of severity. This study aims to explore the vulnerabilities of youth in relation to OSS by classifying and identifying threats, vulnerabilities, and risks based on information security concepts [14]. Using an exploratory approach, relevant articles were extracted from platforms like Google Scholar and Science Direct. The review focused on literature published in the past two decades, specifically examining the intersection of the internet, youth (aged 10–18), and the risks they face. Searches were conducted using keywords and synonyms related to youth, Internet, online, risk, threat, and sexual solicitation.

The objective of the paper is to explore the category of youth vulnerable to OSS. It seeks to analyze the risks associated with OSS and evaluate the severity of these risks for youth using a risk assessment matrix.

2 Conceptualization

For the purpose of the present study, the concepts of vulnerability, risk, and threat were adapted from the literature in information security studies. The variables used are conceptualized below:

Vulnerability The state of being exposed to the possibility of physical or emotional harm is referred to as vulnerability. Vulnerabilities can be exploited as weaknesses in a person or a system. They can be considered as a web threat and a cause for concern, as they can enable other threats. Vulnerabilities may arise from any kind

of weakness in human behavior, technology systems, or other resources that may lead to a harmful exploit or incident. Additionally, vulnerability can be described as an interactive process between a young person's social context and underlying factors that place them at risk of negative outcomes. Disadvantaged environments such as substance-abusing families, abusive/violent households, or families with mental illness, as well as individual characteristics like aggressive temperament or risk-taking tendencies, can result in vulnerabilities.

Threat It can be defined as something that can intentionally or unintentionally exploit a vulnerability. Online threats refer to the dangers that people and computer systems face in the digital world. This category encompasses a wide range of risks, including commonly known threats such as phishing and computer viruses. For the purposes of this study, threats are defined as anything that can potentially lead to Online Sexual Solicitation (OSS) by exploiting the vulnerability of young people.

Risk It refers to the possibility of loss, damage, or destruction that may occur when a vulnerability is exploited by a threat. It is defined as the likelihood that a threat agent will take advantage of a system weakness and cause harm to the system. In this particular study, Online Sexual Solicitation (OSS) is regarded as the risk that arises when young people's vulnerabilities are exploited by various threats.

3 Analysis

To simplify understanding, the risk assessment does not involve assigning numerical values to human factors in this context. As previously mentioned, youth are exposed to various online threats that can become risks when vulnerabilities are exploited. To assess the risks, it is important to identify the threats and the agents behind them, as well as the vulnerabilities that can lead to risk. Risk is the combination of the probability of an event and its consequence.

Figure 1 illustrates the threat model and depicts threats presented to youth. A threat can be a potential cause of harm Whereas, the term *threat agent* is used to denote an individual or group that can manifest a threat.

T_1:*Age-inappropriate content* The Internet offers a wide range of content that may not be suitable for all age groups, posing risks of accidental or intentional exposure. Youth, equipped with Internet-enabled mobile devices, face the danger of accessing age-inappropriate material like pornography and adult content also raise concerns related to coercion, consent, and illegal activities. Research highlights the prevalence of young people, particularly boys aged 14–16, being exposed to online pornography, with a significant percentage regularly watching it [8].

T_2:*Self Harm* The act of purposefully hurting oneself as an emotional coping mechanism that may or may not lead to suicide is self-harm. Youths may start these practices with self-harm, and repetition may become a factor leading to attempting suicide. Websites have been reported to invite users to self-harm by promoting suicide, anorexia, and hurting themselves. The Internet has been identified as a cat-

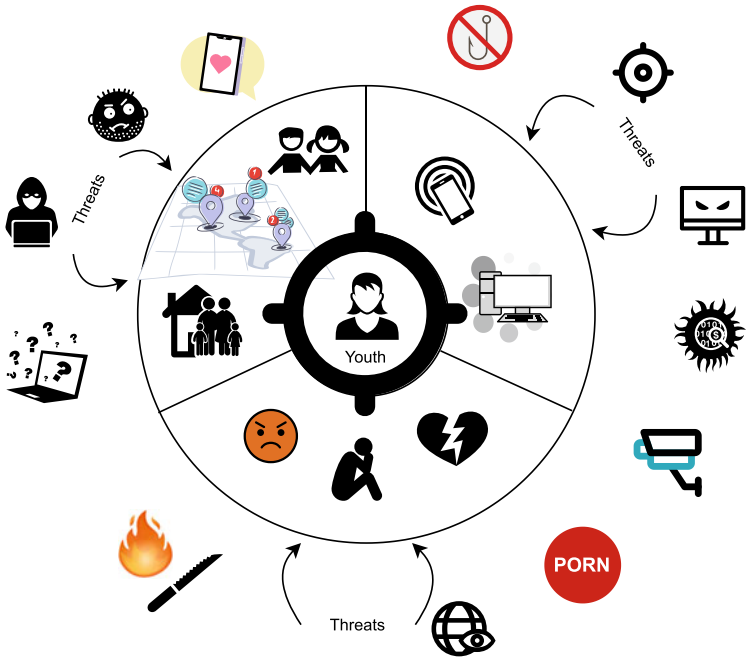


Fig. 1 Threat model of youth with reference to OSS

alyst for self-harm practices, influencing over 10% of individuals through self-harm content on social networking sites [17].

T_3:*Internet Addiction* Internet addiction, characterized by compulsive online behavior that disrupts daily life and strains relationships, poses a significant threat to today's youth. The COVID-19 pandemic further intensified reliance on screens among youth, as academic and social activities shifted to virtual platforms, leading to increased screen time and potential addiction risks.

T_4:*Cyber Crimes and frauds* Cybercrime and frauds targeting youth have become prevalent in today's digital age. Young individuals are vulnerable to various forms of online criminal activities. Identity theft is a significant concern, as attackers steal personal information to commit fraud and cause harm. Phishing scams trick unsuspecting youth into revealing sensitive data, such as financial information. Moreover, young people are exposed to fraudulent websites, counterfeit products, and malware disguised as enticing offers. Clickbait tactics lure them into clicking misleading links.

T_5:*Presence of Sexual predators/pedophiles* The internet can be a dangerous place for children due to the presence of sexual predators. These individuals hide their adult identity and use strategies to befriend young people in order to exploit them sexually. The use of private devices with Internet access and the presence of youth on social media platforms increase their vulnerability to this threat. Websites, video gaming

platforms, instant messaging, and social media are potential platforms where sexual predators may prey on young individuals.

T_6:*Online Bullying* Cyber-bullying refers to the use of electronic means to bully a person, usually through intimidating and threatening messages, sharing of embarrassing photos and videos, or creating fake profiles on social media. It can also involve spreading false rumors on online forums or engaging in abusive behavior during online gaming, including using offensive language and targeting other players' avatars with violent actions.

T_7:*Cyber Stalking* It involves using internet technology to invade someone's privacy, monitor their activities, and make online threats. With the Internet constantly evolving, cyberstalkers have become more sophisticated, using various tools to increase the risk to their victims. The consequences of cyberstalking can be severe, causing fear, anxiety, depression, sleep problems, and social isolation. It can even lead to more serious conditions such as post-traumatic stress disorder, suicidal thoughts, and fear for one's safety.

Vulnerabilities Vulnerability refers to a state of being exposed and poorly protected, particularly for young people who are more susceptible to dangerous situations and exploitation. These vulnerabilities can stem from demographic, social, economic, psychological, and health factors. The youth's behavior in the online world is influenced by their offline life circumstances, combined with the threats and risks they encounter online [3]. Table 2 presents various indicators of vulnerabilities as reported in various relevant literature reviewed. Demographic vulnerabilities include mid-adolescence, gender, sexual orientation, and single-parent families. High-risk internet use factors include excessive online activities, lack of supervision, communication with strangers, and risks of malware. Psycho-social vulnerabilities encompass low parental supervision, cyberbullying, social isolation, and offline victimization. Family-related vulnerabilities involve troubled homes, youth in care, and caregiving responsibilities. Lastly, behavioral vulnerabilities include aggressive behavior, meeting online acquaintances offline, engaging in sexting or discussing sexual activities, visiting explicit websites, and experiencing depression or low self-esteem. These vulnerabilities, when combined with certain demographic and high-risk internet behaviors, significantly increase the risk of online sexual solicitation.

Risk Analysis A risk analysis has been conducted to assess the impact and likelihood of various threats, indicating which threats are more likely to result in the risk of Online Sexual Solicitation (OSS). The analysis considers high-impact threats, which can cause significant harm to the physical, social, and mental well-being of young individuals as shown in Table 3. Moderate-impact threats may lead to moderate pathologies, while low-impact threats have minimal impact. The likelihood of each threat is categorized as low, medium, or high based on its probability of occurrence.

For instance, Threat T_1, which involves exposure to age-inappropriate content, was found to present a moderate risk of OSS. The impact of this threat varied among youth, particularly considering demographic vulnerabilities such as age. Mid-age adolescents were found to be more vulnerable to OSS. Exposure to age-inappropriate content, including violence, pornography, and other explicit material, can signifi-

Table 2 Taxonomy of vulnerabilities of youth related to online sexual solicitation

Indicators	Factors	Related works
Demographic	Age gender sexual orientation family composition living standard	[2, 6, 11, 12, 22]
High-risk internet use	Time spent online gaming private internet device access online contact with strangers low online security hacking private data exposure	[2, 3, 5, 7, 21]
Psycho-social	Parental supervision lack of knowledge cyber bullied minors social isolation offline victimization	[2, 10, 12, 21]
Family	Troubled home living in care caregiver	[3, 22]
Behavioral	Aggression meeting online people in real talking about sex with strangers visiting pornographic websites receiving inappropriate pictures sexting depression low self-esteem	[10, 21, 22]

Table 3 Risk analysis of threats

[Low: ○, Moderate: ◓, High: ●]

Threats	Impact	Likelihood	Risk of OSS
T_1 Age-inappropriate content	◓	●	◓
T_2 Tendency of Self Harm	●	○	◓
T_3 Internet Addiction	○	◓	○
T_4 Cyber Crimes and frauds	◓	◓	◓
T_5 Presence of Sexual predators or pedophiles	●	○	◓
T_6 Online Bullying	●	◓	●
T_7 Cyber Stalking	●	◓	●

cantly impact the mental well-being of young people. While this threat does not directly result in OSS, it can indirectly contribute to the risk by exposing individuals to inappropriate sites.

The likelihood or probability of risk refers to the chances of it happening and is classified into probable, possible, and improbable classes. The severity of the impact is categorized as Acceptable (little or no effect on the health and well-being of youth);

Table 4 Risk probability matrix

Risk probability		Severity of impact		
		Acceptable	Tolerable	Unacceptable
Likelihood	*R*3: Improbable (Unlikely)			*T*5
	*R*2: Possible (Likely to occur)	*T*3	*T*4	*T*6, *T*7
	*R*1: Probable (High likely)		T1	

Tolerable (causes some effects on the health and well-being of youth but does not result in significant harm); and Unacceptable (causes major disruption to the life of youth or may lead to serious injury or death).

These categories were derived from Table 3, which presents a risk analysis of the threat. Using the likelihood and impact of each risk, they were placed into a risk assessment matrix and assigned a level of severity: Low, Moderate, or High. The criteria for identifying the severity of the risk are based on the risk assessment matrix from Table 4. For example, a risk with an acceptable impact and low probability would have a low severity level.

The analysis reveals that threats from sexual predators, online bullying, and cyberstalking can lead to the risk of Online Sexual Solicitation (OSS) [15]. The study explores vulnerabilities of youth and categorizes them accordingly. Demographic vulnerabilities, high-risk Internet use, and certain behaviors are key factors to consider. Girls and queer youth face a higher risk, as well as those with risk-taking tendencies and conflict with social norms. Sexting, responsiveness to sexual communication, and experiences of online bullying and cyberstalking increase vulnerability to OSS. Addressing this issue requires family and school involvement, with offline and online safety programs, caregiver support, and constructive parenting approaches to reduce the risk of OSS.

4 Conclusion

This paper explores the risks and vulnerabilities faced by young people regarding Online Sexual Solicitation (OSS) in the digital society. It presents a detailed analysis of threats, vulnerabilities, and their impact, using an information security-based risk analysis. The research reveals that youth face threats such as exposure to inappropriate content, self-harm, predator danger, online bullying, and stalking, which, combined with vulnerabilities, increase the risk of OSS. Factors like sexting, sexual interactions, low parental supervision, and meeting strangers contribute to vulnerability. To mitigate these risks, regular digital literacy workshops involving students, parents, and peer counselors are recommended, along with strategies to empower youth and facilitate permanent content deletion. Collaboration between parents, the community, and society is vital in minimizing the severity of OSS.

References

1. eSafety Commissioner: State of play-youth, kids, and digital dangers (2018)
2. De Santisteban P, Gamez-Guadix M (2018) Prevalence and risk factors among minors for online sexual solicitations and interactions with adults. J Sex Res 55(7):939–950
3. El Asam A, Katz A (2018) Vulnerable young people and their experience of online risks. Hum-Comput Interact 33(4):281–304
4. Jones LM, Mitchell KJ, Finkelhor D (2013) Online harassment in context: trends from three youth internet safety surveys (2000, 2005, 2010). Psychol Violence 3(1):53
5. Kostić J, Panić T, Pedović I (2016) Problematic online behaviors and experiences of adolescents: correlations with parental mediation, information disclosure and personality characteristics. Br J Educ, Soc Behav Sci 18(4):1–14
6. Kowalski RM, Giumetti GW, Schroeder AN, Lattanner MR (2014) Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. Psychol Bull 140(4):1073
7. Livingstone S, Helsper EJ (2013) Children, internet and risk in comparative perspective. J Child Media 7(1):1–8
8. Mascheroni G, Cuman AD (2014) Net children go mobile: final report. deliverables d6. 4 and d5. 2
9. Mitchell KJ, Finkelhor D, Wolak J (2001) Risk factors for and impact of online sexual solicitation of youth. Jama 285(23):3011–3014
10. Mitchell KJ, Finkelhor D, Wolak J (2007) Online requests for sexual pictures from youth: Risk factors and incident characteristics. J Adolesc Health 41(2):196–203
11. Mitchell KJ, Jones LM, Finkelhor D, Wolak J (2014) Trends in unwanted online experiences and sexting
12. Montiel I, Carbonell E, Pereda N (2016) Multiple online victimization of spanish adolescents: Results from a community sample. Child Abuse Negl 52:123–134
13. Nigam H, Collier A (2010) Youth safety on a living internet: report of the online safety and technology working group
14. Panwar K (2020) Internet use among tweens and teens: Threats, risks and concerns. Explor, ISS e-J 4(1):92–118
15. Panwar K, Sihag VK (2020) Changing forms of cyber violence against women and girls. Indian Police J 111
16. Panwar K, Sihag VK (2023) Protest movements, social media, and the role of law enforcement. Int J Inf Syst Soc Change (IJISSC) 14(1):1–13
17. Rasmussen S, Hawton K et al (2014) Adolescent self-harm: a school-based study in northern ireland. J Affect Disord 159:46–52
18. SKFORCE I.A.T.: enhancing child safety & online technologies (2008)
19. Smahel D, Machackova H, Mascheroni G, Dedkova L, Staksrud E, Ólafsson K, Livingstone S, Hasebrink U (2020) Eu kids online 2020: survey results from 19 countries
20. Teimouri M, Benrazavi SR, Griffiths MD, Hassan MS (2018) A model of online protection to reduce children's online risk exposure: empirical evidence from Asia. Sex Cult 22(4):1205–1229
21. Wachs S, Wolf KD, Pan CC (2012) Cybergrooming: risk factors, coping strategies and associations with cyberbullying. Psicothema 628–633
22. Wolak J, Finkelhor D, Mitchell KJ, Ybarra ML (2010) Online "predators" and their victims: myths, realities, and implications for prevention and treatment
23. Wurtele SK, Berkower F (2010) Off limits: a parent's guide to keeping kids safe from sexual abuse. Safer Soc Press

Implementation of Recommendation System's Service Model Using Amazon E-commerce Dataset

Riddhi Dhage, Suyash Nehete, Sarvesh Hon, Tanuja Patankar, and Laxmi Kale

Abstract Amazon, Flipkart, Spotify, Netflix are renowned names in the world of online businesses. What makes them so successful is their ability to provide users with the right content they are looking for without complex search. These platforms embed powerful recommendation systems to recommend products and contents to users which is of their interest. Recommendation engines make use of user's previous purchase history, likes and dislikes, reviews, ratings, and the frequently visited items to recommend such personalized contents. Thus, better recommendation system helps businesses to improve their sales, customer engagement, and ultimately the profit earned. But for small-scale businesses and startup, implementing recommendation systems from scratch is quite difficult task and is often ignored by them. Solution to this problem would be implementing recommendation system as a cloud-based service that can be easily called and used by the businesses. This paper describes the proposed recommendation system which will be deployed as a service and could be used by any online business firm to generate recommendations for their customers by making simple API calls. The paper explains the algorithms that are used to implement recommendation system, also, describes the dataset generated by scraping the Amazon e-commerce site and the results obtained by training the recommendation model on that dataset. Evaluation metrics used are accuracy and hit rate. Results are determined for every cycle after tuning the batch size for 20 epochs each. With

R. Dhage (✉) · S. Nehete · S. Hon · T. Patankar · L. Kale
Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune 411044, India
e-mail: riddhi.dhage20@pccoepune.org

S. Nehete
e-mail: suyash.nehete20@pccoepune.org

S. Hon
e-mail: sarvesh.hon20@pccoepune.org

T. Patankar
e-mail: tanuja.sali@pccoepune.org

L. Kale
e-mail: laxmi.adhav@pccoepune.org

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_37

batch size 512, at fifth epoch, model is trained well for the given dataset and gives maximum training and validation accuracy with hit rate increased to 64%.

Keywords Recommendation systems · Services · Collaborative filtering · Content-based filtering · Hybrid recommendation system

1 Introduction

Recommendation systems—sometimes also called as recommendation engines are the part of information filtering systems that suggest the products or content having the highest probability to be brought or visited by the user. These systems play an important role in today's online businesses including content providers and e-commerce by engaging maximum number of users in the business. Recommendation systems use following two main approaches to recommend contents.

1.1 Collaborative Filtering

This system uses collaborative filtering method to generate recommendations which is a popular technique used to predict the interests of a user based on the past behaviors and preferences of other users who are similar to them. The idea behind collaborative filtering is that users who have similar preferences in the past are likely to have similar preferences in the future, e.g., say User A has watched and rated "A Space Odyssey" and "Seven Samurai" highly. The recommendation system finds that User B has also watched and rated these movies highly, as well as "Interstellar". The system then recommends "Interstellar" to User A, as he has not watched it yet but has similar preferences to User B.

1.2 Content-Based Filtering

Content-based filtering is a technique that makes recommendations based on the similarity between the features or attributes of items. This approach recommends items to a user that are similar in content to items that the user has previously liked or interacted with, e.g., say User A has previously read and liked the books such as "The Jungle Book", "The White Man's Burden", and "Recessional". There is high probability for User A to be recommended with "Lispeth", "The White Seal" or "Mowgli's Brothers", a few creations by Rudyard Kipling.

1.3 Hybrid Recommendation System

A hybrid recommendation system is a type of recommendation system that utilizes multiple recommendation techniques to generate more accurate and diverse recommendations. By combining different methods, hybrid systems can overcome the weaknesses of one method by leveraging the strengths of another.

One common type of hybrid system is a content-based and collaborative filtering hybrid. Content-based filtering recommends items based on their attributes, while collaborative filtering recommends items based on the preferences of similar users. By combining these two techniques, the hybrid system can provide more personalized and accurate recommendations to users.

The proposed system uses hybrid model by sequentially employing content-based and collaborating filtering approach.

2 Literature Survey

We studied various papers for implementing recommendation model. Key findings and conclusions by different authors are described below.

This paper talks about the user-based collaborative filtering suggestion algorithm based on the Hadoop-Mahout machine learning framework and the Naive Bayes classifier to create a user model and make a personalized recommendation [1]. Hadoop platform is a cloud computing tool that is free to use. It uses distributed-file systems to make parallel computing work better. This method can greatly increase the number of transactions, so e-commerce websites use it a lot. Also, the recommendation system built on the user model and the Mahout framework can give web users good suggestions, which makes it much easier for them to find information.

In this work, different filtering methods, such as BayesNet, decision table, logistic, k-NN, JRip, LibSVM, randomized filtered classifier, random forest, and random tree, are compared with each other [2]. The Kappa statistic, mean absolute error (MAE), root mean squared error (RMSE), F-measure, ROC area, and accuracy are all used to measure how well something works. Using the WEKA tool, it compares the algorithms on the MovieLens dataset to find the best one. With an accuracy of 99.39%, the random forest method seems to give the best results.

This study suggests a deep neural network architecture that uses not only ID embeddings but also extra information like job postings and candidates in the case of a job recommendation system [3]. It uses the interaction between the user and the item for CF and extra knowledge about both the user and the item for content-based filtering. DeepHybrid is a way to predict the top-n list of things. It is made up of matrix factorization and deep neural network. The paper shows the results of experiments done on a real dataset from a job website. The results are better than with the deep learning model that only uses ID embeddings.

This paper confers the collaborative filtering approach which is customized and personalized based on both vital factors like users and items [4]. The experimentation was carried out on the public dataset MovieLens-100K and visualization and analysis was performed to enhance the effectiveness of the model. The experiments of the research paper were carried out using MATLAB platform. The changes in training and testing splits were made to obtain optimal results. The root mean square error (RMSE) parameter was utilized to calculate the accuracy. It summarizes that hybrid recommendation system is an efficient technique to resolve the defects in working of system.

This paper introduces to a novel approach which is deep learning-based recommendation system [5]. The orthodox methodologies like hybrid of content-based system and collaborative filtering approach have certain limitations, these are surmounted by deep learning algorithms. The hybrid model proposed has two distinct cases incomplete cold start case and complete cold start case, this consolidated model is compared with hybrid feature section algorithms and the statistical analysis is performed. It is concluded that the mean absolute error measure is preferable over the hybrid feature selection measure. Thus the graphical results depict the similar observations.

3 Proposed Work

3.1 Datasets

Model is trained on two datasets—user and product dataset. These datasets are generated by scraping Amazon's e-commerce website in order to gather real-time user ratings and product features.

1. **Rating.csv for collaborative filtering**

This dataset has 731,786 records and is used for collaborative-based recommendation systems. It includes columns like "user_id", "product_id", "rating", and "timestamp". The "user_id" represents the unique identifier for users, "product_id" represents the unique identifier for products, rating represents user's ratings on a scale of 1–5, and timestamp represents the time when the rating was given. This dataset is used to train and evaluate recommendation algorithms that provide personalized product recommendations based on user's past ratings and preferences (Fig. 1).

2. **Product.csv for content-based filtering**

This dataset contains 55,094 records used for content-based recommendation. It includes columns like "product_id", "title", and "description", providing information about products. The dataset can be used to build recommendation algorithms based on product titles and descriptions. Proper text preprocessing techniques, such as tokenization and feature extraction is applied to prepare the data for machine learning algorithm. It is important to ensure data accuracy and reliability (Fig. 2).

Fig. 1 Rating.csv

	user_id	product_id	rating	timestamp
0	A3J3BRHTDRFJ2G	B00005TQ09	2.0	1029283200
1	AXNQ3UXQ3P5HE	B00005TQ09	4.0	1052006400
2	AUNJSK273Z367	B00005TQ09	5.0	1028678400
3	A2RATH5QG78TYE	B00005TQ09	4.0	1074902400
4	A2YHVQ83C7VLGV	B00005TQ09	3.0	1080604800

	product_id	title	price	description
0	B00005TQ09	Comfort Perfume Deluxe After Wash Fabric Condi...	₹238.00	\n About this item Add comfort fabric cond...
1	B0001Y7UAI	VillageTiger Washable, Reusable 5 Layer Respir...	₹249.00	\n About this item Care Instructions: Hand...
2	B00020BJA8	FIREWELD Welding Safety Face Shield/Head-Mount...	₹279.00	Super lightweight design,plasma cutter face...
3	B000BQ7GW8	3M 1100 Ear Plug Corded, Extra Soft, Reusable ...	₹332.00	\n About this item Smooth, dirt resistant ...
4	B000EPR7XO	Cadbury Bournvita Chocolate Health Drink, 2kg ...	₹670.00	This is a Vegetarian product. \n Ab...

Fig. 2 Products.csv

3.2 System Architecture

The system is composed of three main components: client frontend, client backend, and recommendation service. Client frontend interacts with client backend to perform various operations such as login, registration, and recommendation requests. Client backend interacts with the recommendation service to enable users to receive recommendations based on their inputs. The client can upload two CSV files, products.csv and ratings.csv, at the beginning of the process. The data from these files is preprocessed, and product and user IDs are converted to a numerical format and stored in a mapping file. The recommendation service includes both content-based filtering and collaborative filtering models. The content-based filtering model uses TF-IDF vectorizer and cosine similarity to make recommendations based on the content of the products. The collaborative filtering model uses a PyTorch Lightning neural network to make recommendations based on the user's previous ratings. Both models store their mappings in a pickle file for future use. When a user requests a content-based recommendation, they provide a product ID, which is processed by the content-based filtering model. The model returns a list of recommended product IDs. When a user requests a collaborative filtering recommendation, they provide a user ID, which is processed by the collaborative filtering model. The model returns a list of recommended product IDs based on the user's previous ratings. For hybrid recommendations, the system first generates a content-based recommendation list based on the provided product ID. Then, the user ID and the content-based recommendation list are passed to the collaborative filtering model, which generates a final list of recommended products. Here we are using sequential approach for generating hybrid recommendations (Fig. 3).

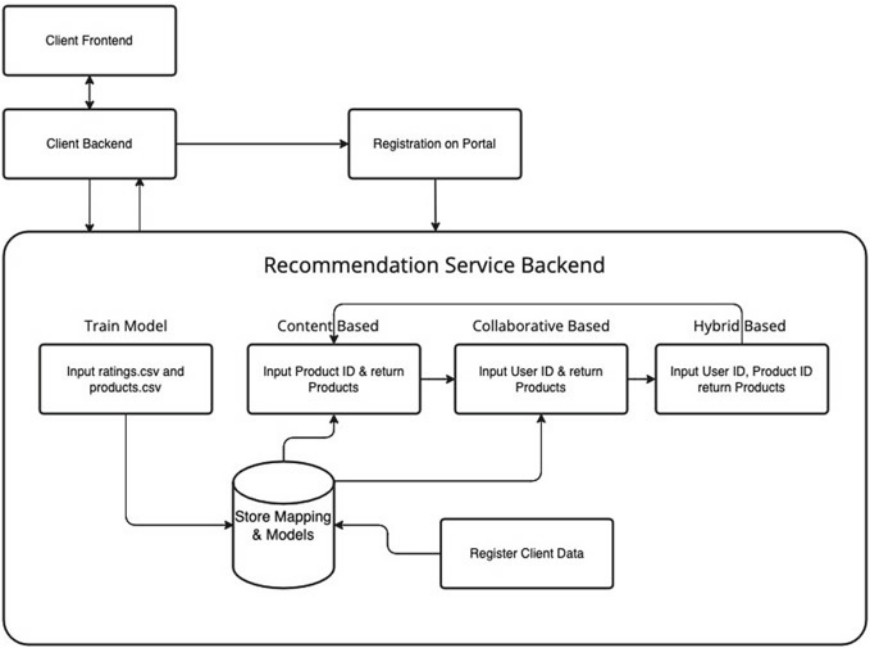


Fig. 3 System architecture diagram

3.3 Algorithms Used

Proposed system is a hybrid recommendation engine which filters and recommends data using both, collaborative and content-based approach. It sequentially filters data collaboratively (i.e., taking all its past rating and reviews for the products in consideration and finding the similar kind of users) and based on contents (i.e., considering attributes or the characteristics of the items), respectively. Thus, system follows sequential approach. Proposed system uses neural collaborative filtering algorithm for collaborative filtering and TF-IDF technique for content-based filtering.

Neural Collaborative Filtering: Neural collaborative filtering (NCF) is a deep learning-based approach which extends collaborative filtering by incorporating neural networks, which can learn more complex patterns in the user-item interactions than traditional collaborative filtering techniques. NCF typically uses a multilayer perceptron (MLP) neural network to learn the nonlinear mapping between the user and item features and the corresponding ratings. The neural network is trained to predict the rating of a user for an item based on the user and item embeddings learned during training. The user and item embeddings are learned using a matrix factorization technique—singular value decomposition (SVD). The embedding vectors are initialized randomly and updated during training using backpropagation and stochastic gradient descent (SGD) to minimize the loss function. The loss function

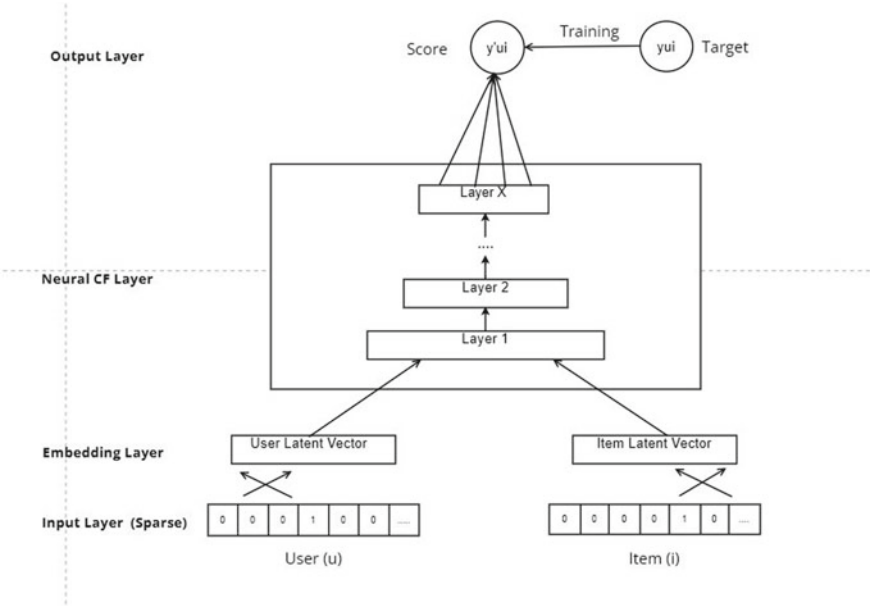


Fig. 4 Neural collaborative filtering architecture

typically measures the difference between the predicted ratings and the actual ratings (Fig. 4).

Term Frequency—Inverse Document Frequency (TF-IDF): It is a statistical measure that reflects how important a word is to a document in a collection of documents.

The term frequency (TF) part of TF-IDF measures the frequency of a word in a specific document. It is calculated as the ratio of the number of times a word appears in a document to the total number of words in that document. Higher TF value indicates that a word appears more frequently in a document.

The inverse document frequency (IDF) part of TF-IDF measures the relevance of a word in the entire collection of documents. It is calculated as the logarithm of the ratio of the total number of documents in the collection to the number of documents that contain the word. Higher IDF value indicates that a word is less common and more important in the overall collection of documents.

The TF-IDF vectorizer multiplies the TF and IDF values for each word to obtain a combined score, which represents the importance of each word in a document relative to the entire collection of documents. The resulting TF-IDF scores are then used as features to represent the textual content of documents in a numerical form and provide recommendations accordingly.

4 Experimental Results

Our recommendation model is trained with the help of binary cross-entropy loss (BCELoss) and stochastic gradient descent (SGD). We trained the model with different batch sizes of 32, 128, and 512 to see how they affected its performance. We are also measuring the hit rate, which is the average number of times the model recommends a correct product to a user. We consider it is a "hit". We use these metrics to look at how well the model works. Following are the results obtained from these experiments.

4.1 Results for Collaborative Filtering

The model was trained with a batch size of 32 and 20 epochs using the BCELoss and SGD optimization algorithms. The training accuracy started at 80.99% and gradually went up to 86.17%, while the validation accuracy started at 97.16% and peaked at 95.34% after the sixth epoch before going down to 91.55% on the last epoch (Fig. 5).

The hit rate was also used to plot the graphs along with training and validation loss and accuracy. The hit rate, which is the average number of times a product is recommended, started at 0.5794 and went up to 0.6067 after the last epoch. Overall, the model did well in terms of accuracy and hit rate, but the validation accuracy changed a little bit after the sixth epoch (Fig. 6).

The model was then trained for 20 epochs with 128 batch size. For each epoch, the training and validation accuracy and loss are shown above.

Overall, both the training set and the validation set show that the model is very accurate, with the highest validation accuracy of 0.8968 being reached in the 15th epoch. In later epochs, however, it looks like the model may have started to overfit to the training data, as the validation accuracy starts to go down, while the training accuracy keeps going up.

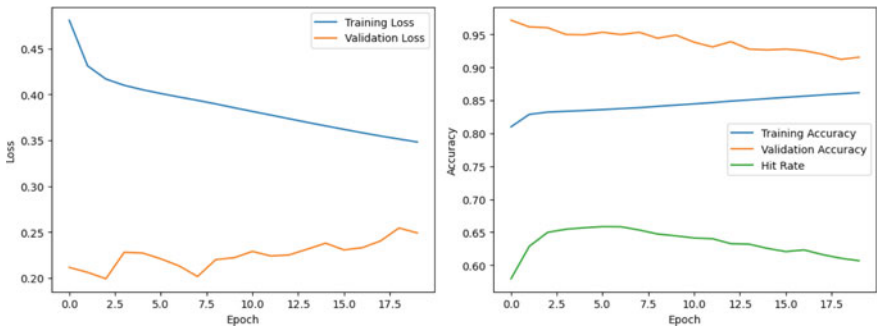


Fig. 5 20 epochs and 32 batch size

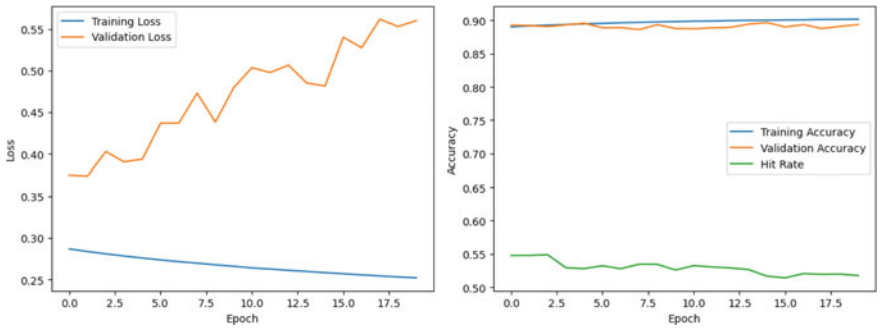


Fig. 6 20 epochs and 128 batch size

The hits metric also follows a similar pattern. In the second epoch, the hits score was 0.5478, which was the highest it ever got.

Overall, the model is very accurate based on the data, but there may be some overfitting in later epochs, which could affect how well it works with new data (Fig. 7).

The model was trained for five epochs, and each batch had 512 data points. Over iterations, the training accuracy starts at 80% and slowly goes up to 86.4%, while the validation accuracy starts at 99% and goes down to 89.4%. Both the training loss and the validation loss go down as time goes on. The hits metric, on the other hand, which measures how well the protection system works, starts at 0.38 and goes up to 0.64. Thus, we get best hit rate at fifth epoch with batch size 512.

Overall, the model seems to do well on the training data, with a high level of accuracy and a lower loss. But the accuracy of validation drops a lot over time, which could be a sign of overfitting. The rise in the number of hits is a good sign that the recommendation system is getting better.

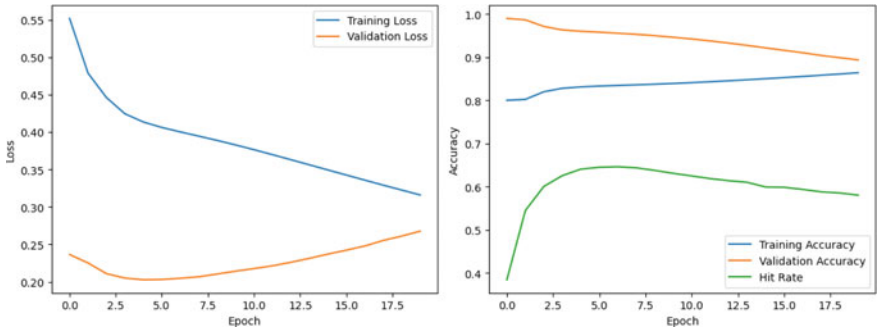


Fig. 7 20 epochs and 512 batch size

4.2 Results for Content-Based Filtering

The graph shows that as the number of features goes up, the average difference between the original and modified kernels goes down. This means that the linear kernel is less sensitive to changes in the number of features. But the rate at which the norm of the difference goes down slows down as the number of features goes up. This suggests that adding more features past a certain point doesn't make a big difference in how well the linear kernel works (Figs. 8 and 9).

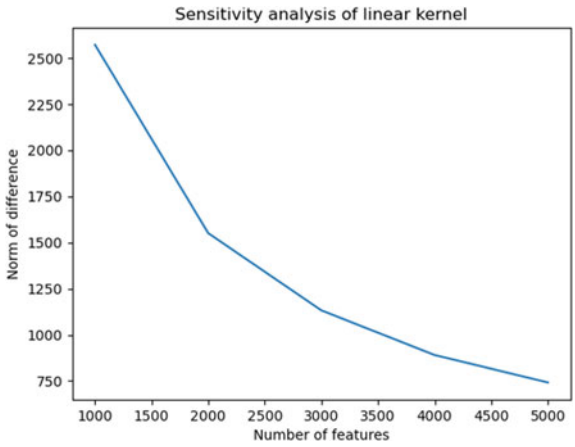


Fig. 8 Sensitivity analysis of linear kernel

Product Id	B00005TQ09	
Product Name	Comfort Perfume Deluxe After Wash Fabric Conditioner Desire 850 ml, Liquid Fabric Softener with Fine French Fragrance for Freshness, Softness & Shine	
User Id	A3J3BRHTDRFJ2G	
Collaborative Filtering	Content Based Filtering	Hybrid Filtering
Bildos 3 Ply Disposable Surgical Face Mask with Nose Pin and Ultrasonic Welded Ear Loop (100)	Comfort Fabric Conditioner Natural - Imported, 2L	VENUS V-Shwas Regular FFP2 Respirator, ISI Certified, Unisex White Anti pollution Face Mask with Ear loop fitting
SURGICOMFORT Non Woven Disposable Surgical Head Cap (100 pcs, Blue Color) for Unisex	Puer Fabric Conditioner | Floral Enigma | 21 Days Fragrance & 30% Extra Softness | Ecocert & COSMOS Actives | Eco-Friendly & Plant-Based Fabric Softner(1 Litre)	QUARANT Melt Blown - SMMS 3 Ply Disposable Face Mask with Nose Pin and Reusable Travel Pouch (Blue, Pack of 100) for Unisex
Production and Characterization of Activated Carbon from Raffia Palm Seeds and Shells	VIWA Drax Perfume for Men 40ML	Scalemox Liquid Detergent 5 Liter All in One Washing Machine No Harmful Chemicals and Acid with Phosphate Free and Eco-Friendly Liquid Detergent pack of 1
100 Disposable Cap Stretchable Blue Bouffant Caps/Surgical Caps/Cooking Caps (100 Pieces)	Comfort After Wash Fabric Conditioner refill pouch, super saver pack Lily fresh variant for all day freshness and lasting fragrance. 2 L	Kidbee Baby Solid Super Soft Cotton Hosiery Bland Color Babies Boys & Girls Regular Shorts (Multicolor (Set of 5), 12-18 Months)

Fig. 9 Comparison of resultant product list

5 Conclusion

Based on our experiments and evaluation, we conclude that the proposed recommendation system as a service (RSaaS) is an effective approach for providing personalized recommendations to users. We utilized a hybrid recommendation approach combining collaborative and content-based filtering techniques to provide accurate and diverse recommendations. We also utilized a sequential approach to combine the results of the content-based and collaborative models. Our experimental results showed that the highest hit rate of 64% was achieved with a batch size of 512 and 5 epochs.

We also evaluated the performance of our model with multiple batch sizes and found that the hit rate improved as the batch size increased. This indicates that our model can scale well with larger datasets and can handle more concurrent user requests. Overall, our RSaaS approach can be a valuable tool for businesses and online platforms looking to provide personalized recommendations to their users.

6 Future Scope

In future work, we plan to incorporate user feedback to refine the recommendations. We can also investigate the use of deep learning models such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs) to improve the accuracy of the recommendations. Furthermore, we plan to extend our recommendation system to handle more complex use cases such as recommending items based on multiple input criteria. Also, we aim to evaluate the effectiveness of our recommendation system on larger datasets and in different domains to further validate its performance.

Lastly, we plan to explore the integration of our recommendation system as a service with other applications and platforms to provide a seamless user experience. We will also investigate the deployment of the recommendation system on cloud platforms to improve its scalability and accessibility.

Acknowledgements Inspiration and motivation always play a key role in the success of any venture. We would like to express our gratitude and appreciation to all those who have been its true source. We sincerely thank Dr. Govind Kulkarni, Director, Pimpri Chinchwad College of Engineering Pune. We pay our deep sense of gratitude to Dr. Sonali Patil, HOD-IT Department to encourage us to the highest peak and to provide us with the most valuable guidance throughout the research work.

References

1. Song B, Gao Y, Li X-M (2020) Research on collaborative filtering recommendation algorithm based on mahout and user model. J Phys Conf Ser 1437:012095. https://doi.org/10.1088/1742-6596/1437/1/012095

2. Alam MT, Ubaid S, Shakil, Sohail SS, Nadeem M, Hussain S, Siddiqui J (2021) Comparative analysis of machine learning based filtering techniques using MovieLens dataset. Procedia Comput Sci 194:210–217. ISSN 1877-0509. https://doi.org/10.1016/j.procs.2021.10.075
3. Çakır M, Oguducu SG, Tugay R (2019) A deep hybrid model for recommendation systems. ISBN 978-3-030-35165-6. https://doi.org/10.1007/978-3-030-35166-3_23
4. Song Y, Liu S, Ji W (2017) Research on personalized hybrid recommendation system. In: 2017 international conference on computer, information and telecommunication systems (CITS), Dalian, China, pp 133–137. https://doi.org/10.1109/CITS.2017.8035321
5. Ouhbi B, Frikh B, Zemmouri E, Abbad A (2018) Deep learning based recommender systems. In: 2018 IEEE 5th international congress on information science and technology (CiSt). https://doi.org/10.1109/cist.2018.8596492

Quantifying Vertical Farming Potential Using Digital Twins

Anushk Naval, Vaibhav Kumar, and Kumar Gaurav

Abstract This study evaluates the potential of vertical farming as a sustainable method for urban agriculture, specifically for building facades. To achieve this, we utilized digital twins of the study area, IISER Bhopal, to simulate irradiance, temperature, and humidity for various 3D buildings for a period of two months. We then validated the simulation outcomes using an in-house IoT device and found that the simulation outcomes were accurate and followed a similar trend to the IoT measurements from various vertical sites. Using the optimal values for irradiance, humidity, and temperature, we generated vertical growth potential maps for various vegetable species. Additionally, we conducted a detailed analysis of the impact of neighborhood built-up design on the potential for vertical farming on building facades. The results showed that not only rooftops but also facades of buildings have great potential for growing various vegetables vertically. However, the effect of the neighborhood was significant, leading to a significant reduction in the potential. As a decision-making tool, this study has the potential to enhance the accuracy and effectiveness of quantifying vertical farming areas. This can be a game-changer in developing sustainable cities and societies.

Keywords 3DGIS · IoT · Digital twin · Sustainable cities · Vertical farming

A. Naval (✉) · K. Gaurav
Department of Earth and Environmental Science, Indian Institute of Science Education and Research Bhopal, Bhopal 462066, India
e-mail: yash.anushk567@gmail.com

K. Gaurav
e-mail: kgaurav@iiserb.ac.in

V. Kumar
Department of Data Science and Engineering, Indian Institute of Science Education and Research Bhopal, Bhopal 462066, India
e-mail: vaibhav@iiserb.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_38

1 Introduction

Urban Agriculture (UA) includes utilizing privately and publicly owned land and water bodies throughout intra-urban and peri-urban areas to produce, process, and market food, fuel, and other outputs within or near cities [1]. Vertical Farming (VF) a form of UA involves growing crops in layers stacked vertically under artificial illumination in a controlled environment. It is gaining huge popularity, primarily because it can produce crops using precise fertilizer management and requires less space and water, VF promises excellent agricultural yields, year-round production, and less environmental impact [2]. Moreover, it reduces the burden of traditional agricultural techniques leading to more opportunities for sustainable development and food security in India [3].

However, the existing methods to access the potential of VF in buildings are done through manual inspection of lighting and climate parameters which are the key drivers [4]. These inspections are highly resource intensive. Further, the lack of dynamic considerations such as seasons, weather, and neighborhood details has been a bottleneck in the urban farming potential analysis.

In this paper, the use of digital twins is proposed as a potent solution for simulating and analyzing various scenarios in vertical farming. Digital twins can create a virtual replica of the physical system and simulate the light sources and meteorological parameters, such as geographic location, time of year, and weather patterns [5]. By simulating the angle and intensity of the sun's rays at different times of day and year, the amount of incident solar energy can be determined, which when correlated with suitable conditions for vegetable species, can generate potential growth maps for the building. This study has validated the irradiance information generated using simulations through the use of a developed IoT device. Additionally, the paper demonstrates the potential of high-fidelity twins in analyzing the impact of the neighborhood on the growth potential of vegetable species, which is a void in the research.

Contributions

The objective of this study is to develop a decision-making tool that offers guidelines for constructing sustainable urban spaces with vertical farming-driven architecture. The main contributions of this research are as follows:

- Simulation of solar irradiance at 1 m vertical and horizontal resolution for the study area using developed Level of Detail-2 (LOD2) 3D building models.
- Validation of simulated irradiance at various vertical levels using a developed IoT device.
- Utilizing the correlation between solar radiation, temperature, and humidity with different vegetable species and identifying their optimal growing conditions for the production of vertical vegetable growth potential maps.
- Analysis of the impact of neighborhoods on growth potential.

This study has the potential to optimize the design process of urban built-up spaces by quantifying the feasibility of vegetable farming, ultimately leading to the paradigm of vertical farming-driven built-up spaces.

2 Methodology

The methodology has been implemented in various stages. In the data collection step, we gathered the meteorological data and thereafter created a 3D model of the study area. The solar irradiance simulations were performed on the 3D model, and the data was then validated using the real data gathered by the IoT sensors. The digital twin was created by combining the data with the 3D model. The solar radiation analysis, which quantifies the areas suitable for vertical farming, was then conducted using the digital twin. The neighborhood impact on the model was then obtained by combining the digital twin with the architectural alterations.

2.1 Development of 3D Model

We developed the LoD2 3D model of the study area as the virtual replica of the real environment [6–8]. Firstly, the initial spatial information of sites' were acquired by GPS. CADmapper, an open-source software, was used to create the CAD file for the area with static information–topography, building heights, road widths, and shapes. The height of the building is then cross-verified using "Mobile GPS". The curves are then extruded to their respective heights to create the LOD 2 model for the location site. The 3D model was divided into 13872 grids, each with a dimension of 1 m $\times$ 1 m.

2.2 Irradiance Simulation Using Sun-Angle Geometry

EnergyPlus Weather (EPW) files contain meteorological data, such as temperature, humidity, and incident radiation. The data has the sun angles recorded at the nearest weather stations which are then extrapolated to the site locations as sky matrix [9]. The sky matrix for the period March and April 2024 was integrated with the 3D model to produce the irradiance values at 1m vertical and horizontal resolution.

2.3 Irradiance Validation Using IoT

Development of IoT unit for meteorological measurements
The Internet of Things (IoT) is used for meteorological monitoring. IoT can provide real-time data gathering, processing, and transfer for remote data collection [10]. The article focuses on IoT unit that collects three meteorological parameters–temperature, humidity, and solar irradiance. The unit comprises of Arduino Uno R3 microcontroller, DHT-22 temperature, and humidity sensor, BH-1750 digital light sensor module, and ESP-8266 Wi-Fi module for wireless communication and IoT. The data is collected remotely from the IoT device to "ThinkSpeak" servers. The microcontroller was programmed to gather information from sensors every two minutes. The information was then transformed into a single-bus format and transmission to the "ThinkSpeak" servers via the IoT device's channel API. The acquired data was then analyzed using statistical analysis, data visualization, and trend analysis. The light intensity sensor measures the values in "lux" which is converted to W/m^2 by multiplying the values with a factor of 0.008 [11].

Sensor deployment
The sensors were deployed in various hostel buildings of the Indian Institute of Science Education and Research, Bhopal. The buildings consist of 5 floors with each floor having the height of 3 m. Each room has a facade window measuring 1.5 m × 2 m. These facades were used to do the solar analysis for the potential of microsites for vertical farming. The sensors were placed parallel to the facade on the locations for a seven-day data collection period. The sensors were placed at different heights. Microsite-1 was on floor 4, Microsite-2 and Microsite- 4 was on the ground floor, and Microsite-3 was on the 3rd floor.

Simulation data validation using IoT
Accurate data is critical in developing high-fidelity digital twins. The degree to which the virtual representation. We compared the simulated irradiance and metrological data with real data using IoT outcomes. We observed the trend of the data to validate the simulated outputs.

2.4 Quantification of Vertical Farming Area

To find the suitable sites for specific vegetable growth we took the simulated irradiance output which then uses the optimal ranges of the photosynthetic photon flux density (PPFD) [12] values converted to kWh/m^2 for different species shown in tables below [13], for potential farming area quantification. Then using Ladybug we took the building of interest as the subject geometry and the neighboring building as content. The content's shadow impact was analyzed to give the affected solar potential analysis. The output, in turn, was used to compute the impacted potential area for vertical farming.

3 Results

3.1 Irradiance Simulation Using Sun-Angle Geometry

The EPW file was used to compute the sun sky matrix using the "SkyMatrix" function. The analysis period for computation was taken from 1st March to 30th April, starting from morning 7 AM to 5 PM in the evening. The generated sky matrix with the 3D model was passed into the "IncidentRadiation" function. The function was executed on gridded 3D model with and without the architectural content to get the results, Figs.1 and 2. The results show that for this period, the vertical profile receives a cumulative radiation of 80–220 kWh/m^2. Figure 1 shows a constant incident radiation received, without surrounding affecting the incident sunlight. Figure 2 vertical profile shows the variations due to the effect of neighborhood buildings.

3.2 Validating Irradiance Using IoT Measurements

Trend analysis for microsites

The study was performed for four microsites. The M-1 facade faces 330° northwest, M-2, and M-4, 155° southeast, and the M-3 facade faces 60° northeast. For M-1, it was expected that the facade would not receive direct sunlight from the sun during the initial hours of the day. This is because the facade faces opposite to the direction of the rising sun in the morning. As time progresses, in the afternoon, it is expected

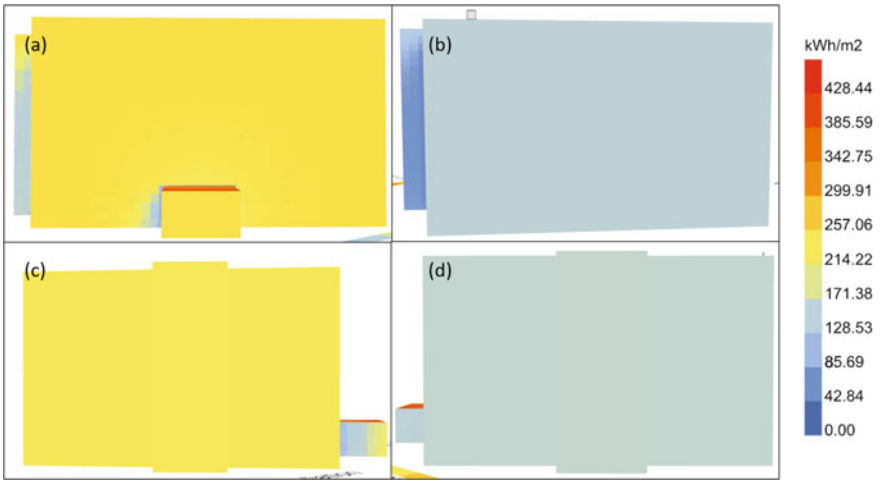


Fig. 1 Results for solar analysis for building faces **a** south face; **b** north face; **c** east face; **d** west face without the neighborhood effects

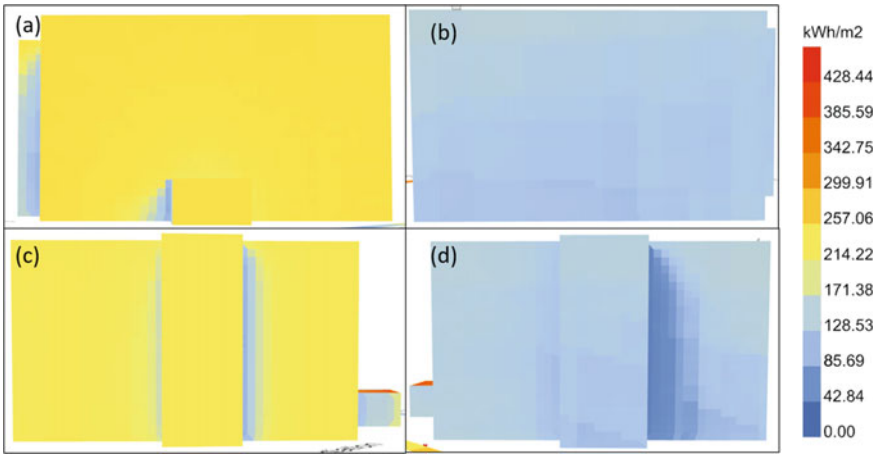


Fig. 2 Results for solar analysis for building faces **a** south face; **b** north face; **c** east face; **d** west face considering the neighborhood effects

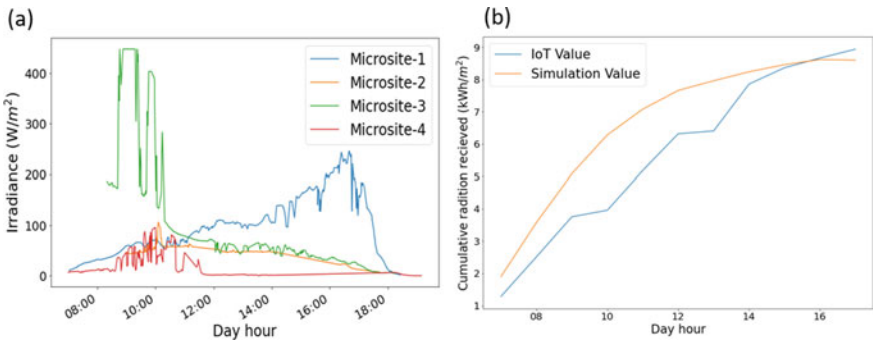


Fig. 3 **a** The time series plot for the 4 microsites, plotting iiradiance received with time for 7th April 2023. **b** Plot showing the cumulative radiation received for Microsite-1 averaged for 7 days for both IoT and simulated data

that the facade will start receiving direct sunlight. Similarly, for M-2 and M-4, it was expected that the facade would receive direct sunlight from the sun during the initial hours of the day. Furthermore, for M-3, it was expected that the facade would receive direct sunlight from the sun during the initial hours of the day. Fig. 3a shows the plots for the radiance received for the 4 microsites.

When the IoT sensor data was plotted with the simulation data, it was observed that the data was similar. Figure 3b shows the plot of the cumulative radiance from morning 7 AM to evening 5 PM averaged for seven days. Both the data and simulation show a positive correlation of 0.97. Both the data have a positive covariance of 5 and a standard deviation of 2.4, showing that the data follow the same and similar trend. Thus, the data can be trusted and used for simulating solar analysis.

	Site-1	Site-2	Site-3	Site-4
Day-1	-0.92	0.00	-0.81	-0.94
Day-2	-0.93	-0.44	-0.94	-0.96
Day-3	-0.94	-0.28	-0.67	-0.84
Day-4	-0.97	0.14	-0.81	-0.92

Fig. 4 Correlation heat-map of temperature and humidity for different sites. The deeper the color represents the more negative the value

Correlation between temperature and humidity

A correlation matrix between temperature and humidity was generated for the four sites. Based on the data analyzed in Excel, a general trend of decreasing humidity with increasing temperature had been observed across all sites. However, Microsite-2 stands out as an exception, showing a different pattern. Instead of a negative correlation between temperature and humidity, Microsite-2 exhibits either a weaker negative dependence or a positive correlation between temperature and humidity.

Fig. 4 shows the correlation matrix heat map of temperature and humidity for four sensors on four randomly selected days. This deviation in the correlation between temperature and humidity at Microsite-2 compared to the other sites was found out to be because of a tree in front of the site. This tree was a barrier to our sensor, making the observational site cooler by casting shadows over the site.

3.3 Quantifying Vegetable Yield Potential

Analysis of vegetable growth on building

The results from the irradiance analysis and Table 1 were used to identify the potential locations and total area for growing vegetable species. Figure 5 shows the potential areas on the vertical profiles of the building

The simulation provides that the optimal area for tomato production is 26%, i.e., 3574 m^2. Other plant species, like cannabis flowering, require higher radiations for optimal growth, providing a different statistic that only 12% of the area is optimal for farming. Table 2 shows the quantification results for different plant species.

Table 1 Light levels (converted PPFD) for different species of corps

Plant	Min (kWh/m^2)	Max (kWh/m^2)	$Typical$ (kWh/m^2)
Tomato	73	149	115
Pepper	51	128	98
Cucumber	51	149	98
$Cannabis vegetative growth$	119	235	149
Cannabis flowering	277	640	427

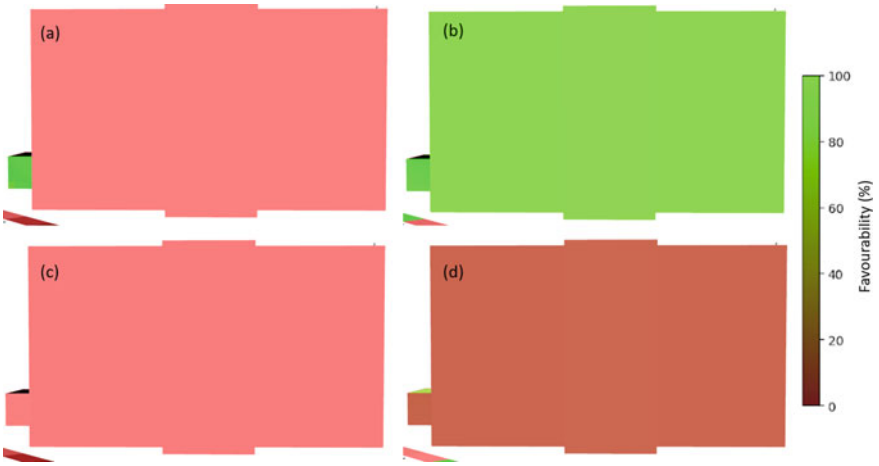


Fig. 5 Figure showing the potential sites (without considering neighboring content) for the growing of **a** tomato, **b** cannabis, **c** pepper, **d** cannabis flower

Table 2 Quantified farming area based on different species with and without considering the content effects

	Without content		With content		
Plant	Potential grids	Percentage area	Potential grids	Percentage area	Percentage change
Tomato	3560	25.6	3574	25.8	0.2
Pepper	5139	37.1	7093	51.2	14.1
Cannabis vegetative	4238	30.5	2948	21.3	-9.2
Cannabis flowering	1751	12.6	1709	12.3	-0.3

Analyzing neighborhood impact on growth potential

The results from the solar analysis with the neighborhood content and Table 1 were used to quantify the favourability of building for crop species growth. We simulated the digital twin for the study from 1 March to 30 April. As a result, Fig. 6 was obtained. It was observed that the shadow from the content buildings caused some grids, which were not optimal before, to now preferred grids for farming. Low light-requiring crops like pepper can be grown in more quantity than before, but there is a decrease in the number of grids for high light levels requiring plants. Table 2 shows the quantification results for different plant species.

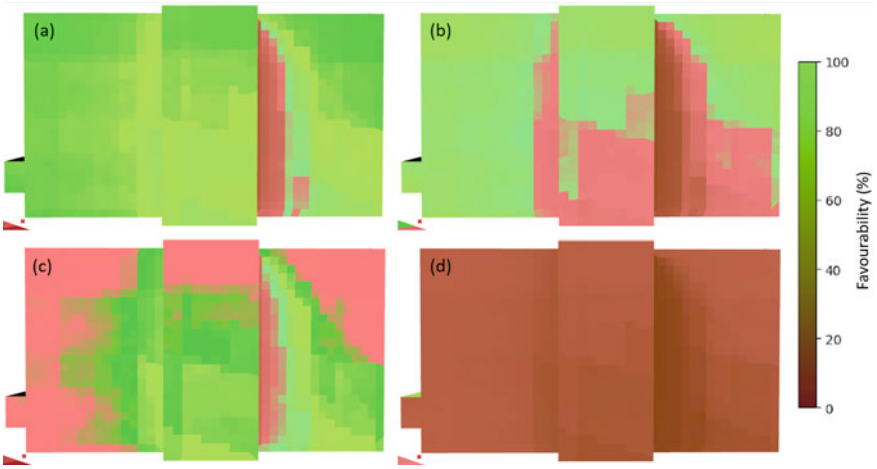


Fig. 6 Figure showing the potential sites (by considering neighborhood information) for the growing of **a** tomato, **b** cannabis, **c** pepper, and **d** cannabis flower

4 Conclusion

This study explored the potential of vertical farming as a sustainable method for urban agriculture on building facades. The use of digital twins of the study area allowed for accurate simulations of irradiance, temperature, and humidity for various 3D buildings. Vertical growth potential maps were generated for various vegetable species using optimal values for irradiance, humidity, and temperature. The validation through IoT proves the potential of digital twins combined with sun-angle geometry for accurate irradiance simulation. Further, IoT was also used to validate the meteorological data. A key finding remains that a LOD2 model is enough to increase the scale of the outcomes to derive city-scale farming potential.

The outcomes also suggest that neighborhood built-up design can significantly influence the growth potential. Hence, it should be a part of policy instruments before designing urban spaces. The findings of this study have important implications for sustainable urban development and the implementation of smart agricultural techniques, which may lessen the carbon impact of conventional agriculture. By enhancing the accuracy and effectiveness of quantifying vertical farming areas, this study can serve as a valuable decision-making tool for urban planners and policymakers, leading to the development of sustainable cities and societies.

References

1. Viljoen A, Bohn K (2014) Second nature urban agriculture: designing productive cities. Routledge. Available online:https://books.google.co.in/books?hl=en&lr=&id=jxwWBAAAQBAJ&oi=fnd&pg=PP1&dq=The+significance+of+vertical+farming+as+a+crucial+element&ots=_KAP6kNFU9&sig=gnvTLnVRz1TTbvymkaTA12XMSFM&redir_esc=y#v=onepage&q&f=false
2. Despommier D (2013) Farming up the city: the rise of urban vertical farms. Trends Biotechnol 31(7):388–389. https://doi.org/10.1016/j.tibtech.2013.03.008
3. Orsini F, Kahane R, Nono-Womdim R, Gianquinto G (2013) Urban agriculture in the developing world: a review. Agron Sustain Develop 33:695–720. https://doi.org/10.1007/s13593-013-0143-z
4. Sahasranaman M (2016) Future of urban agriculture in India. Instit Res Anal Policy 2(10):1-24. Available online:http://irapindia.org/images/irap-Occasional-Paper/IRAP-Occasionalpaper-10.pdf
5. Monteiro J, Barata JT, Veloso M, Veloso L, Nunes JOR (2018) Towards sustainable digital twins for vertical farming. IEEE 234–239. https://doi.org/10.1109/icdim.2018.8847169
6. Azhar S, Khalfan M, Maqsood T (2012) Building information modeling (BIM): now and beyond. Australasian J Constr Econ Build 12(4):15–28. https://doi.org/10.3316/informit.013120167780649
7. Fai S, Rafeiro J (2014) Establishing an appropriate level of detail (LoD) for a building information model (BIM)-West Block, Parliament Hill, Ottawa, Canada. ISPRS Ann Photogram Remote Sens Spat Inf Sci 2(5):123. https://doi.org/10.5194/isprsannals-II-5-123-2014
8. Biljecki F, Ledoux H, Stoter J (2016) An improved LOD specification for 3D building models. Comput Environ Urban Syst 59:25–37. https://doi.org/10.1016/j.compenvurbsys.2016.04.005
9. Crawley DB, Lawrie LK, Pedersen CO, Winkelmann FC (2000) Energy plus: energy simulation program. ASHRAE J 42(4):49–56. Available online: https://www.researchgate.net/profile/Drury-Crawley/publication/230606369_EnergyPlus_Energy_Simulation_Program/links/546a40d80cf2f5eb18077919/EnergyPlus-Energy-Simulation-Program
10. Gaikwad SV, Vibhute AD, Kale KV, Mehrotra SC (2021) An innovative IoT based system for precision farming. Comput Electron Agricult 187:106291. https://doi.org/10.1016/j.compag.2021.106291
11. Michael PR, Johnston DE, Moreno W (2020) A conversion guide: solar irradiance and lux illuminance. J Measur Eng 8(4):153–166. https://doi.org/10.21595/jme.2020.21667
12. Songa S, Cheonga JC, Leea JSH, Tana JKN, Chiama Z, Aroraa S, Pnga KJQ, Seowa JWC, Leonga FWS, Palliwalb A et al (2022) Home gardening in Singapore: a feasibility study of vertical gardening for self-sufficiency in high-rise public housing apartment buildings. https://doi.org/10.1016/j.ufug.2022.127755
13. MechaTronix: Typical PPFD and DLI values per crop. Available online: https://www.hortigrowlight.com/en-gb/typical-ppfd-dli-values-per-crop

Security Analysis of Multi-factor Authentication with Legitimate Key Exchange for Vehicle to Vehicle Communication

Stuti Bhojak, Vipul Chudasama, Daiwat Vyas, and Devendra Vashi

Abstract Internet of Vehicles (IOV) in which the vehicles share information like traffic, road safety, location sharing, toll payment, road accident, etc. with each other wirelessly. Vehicle Ad-hoc Network (VANET) includes reliable data transmission on a network, frequently changing topology, mobility of vehicles, and security of each component where vehicles can communicate securely and effectively. However, the existing framework has some security issues for which the security properties like nonce and multi-factor authentication needs to consider. At last, we suggest that our framework is more secure and efficient for V2V communication.

Keywords IoV · VANET · V2V communication · Security · Performance · Authentication

1 Introduction

In IoV, the vehicles communicate with vehicles and also with some system objects like roadside infrastructure, pedestrians, mobile devices, and public networks with a number of benefits including comfort, speed, and traffic safety [1]. VANET has different types communication: V2X called Vehicle-to-Everything, V2V called Vehicle-to-Vehicle, and V2I is called Vehicle-to-Infrastructure. It uses the designated short-range contact medium to exchange information about moving cars [1]. Which employs the IEEE 802.11p standard that allows for fast implementation, efficient network

S. Bhojak (✉) · V. Chudasama · D. Vyas · D. Vashi
Nirma University, Sarkhej–Gandhinagar Highway, Ahmedabad, India
e-mail: 21MCEI10@nirmauni.ac.in

V. Chudasama
e-mail: vipul.chudasama@nirmauni.ac.in

D. Vyas
e-mail: daiwat.vyas@nirmauni.ac.in

D. Vashi
e-mail: devendra.vashi@nirmauni.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_39

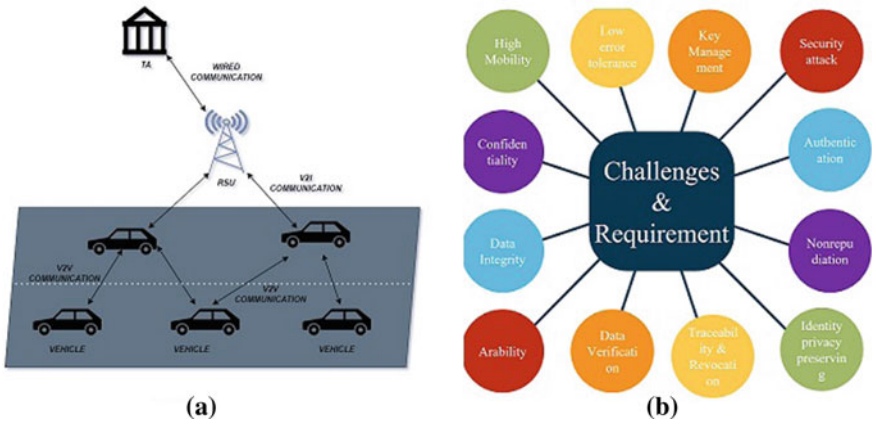


Fig. 1 **a** VANET architecture, **b** challenges and requirement

access, and dynamic spectrum access [2]. VANET has unique characteristics like high mobility, unbounded network, frequent exchange of information, time effective, and cost-effectiveness which makes a network more effective and easier to build and understand [3]. All cars are connected and authenticated by the Road side unit and each RSU is connected to the trusted entity (TA) and must be authenticated by it. Each car has a temper-proof device (TPU) which stores all the information about the vehicle. Vehicles transfer the data with each other for communication so we need to provide security like that so data can't be harmed and protected from malicious users [4]. VANET architecture has three components. First, the mobile unit contains all the portable devices for V2V and V2I communication [5]. Second, the infrastructure domain contains all RSUs and transceivers for V2I communication [5]. Third, the management domain comprises of an administration system-perhaps a server-and a surveillance program [5]. Figure 1 tells how the different types of components connect with each other in VANET. To apply this security we faced some challenges and issues like key management, low error tolerance, high mobility, confidentiality, data integrity, etc., it shows in Fig. 1

1.1 Research Contribution

- We present a framework that optimizes the security and performance of V2V communication using cryptography.
- We perform the single vehicle registration to the TA, for achieved the Ephemeral key leakage resistance.
- If the private key is revealed, then your previous session key should be secure like this, we can achieve forward security.

Table 1 Litarature table

Study	Approach	Proposed technique	Persevered Attack	Type of network	Limitation/future scope
[6]	Secure message transmission	Mutual Authn. using RA and TA	MITM attacks, anonymity attacks, replay attacks	V2V	Long Authn. process
[7]	Secure message authentication	Symatric cryptosystem for security	MITM attacks attacks, replay attacks	V2I	Authn authentication only done by RSU
[8]	RSU controls whole system	Group signature scheme	Tracking attacks eavesdropping	V2V	Reduce performance efficiency and not time effective
[9]	Authentication in VANET	Apply multi-factor authentication	Impersonation, message replay, identity disclosure	V2V	Whole security depens on CA's master key
[10]	Intelligent transportation	Blockchain apply for secure system	Black hole, DOS, Jamming	V2X	Blockchain query overhead
[11]	Data dissemination method for privacy	lightweight encryption fundamentals	MITM, side channel, identity, replay	V2V, V2R	Increase security and privacy, formal security assessments
[12]	Intelligent resource allocation	ML algo. used (RNN, ARIMA)	DOS attack	V2X	ARIMA has a less accuracy
[13]	Secure the data sharing	Proof-of-stake consensus method to improve suitability	Message spoofing attacks, bad mounting attacks	IOV	More time consuming
[14]	Apply security and privacy on IOV	Use blockchain secuirty	DOS, MIMT, blackhole, flooding	IOV	High energy consumption, block transaction limit, more overhead
[15]	Secure from black hole attack	Upgraded Ad-Hoc DVR SIN, and IDAD	Black hole attack	V2I	Not secured from all the attack

- For more security we apply the noce and multi-factor authentication.
- In last, we present the efficiency of the proposed framework which contains the cost of communication, and cost of storage.

1.2 Organization of the Paper

The remaining part of the paper is arranged in this way, the work done in this domain is presented in Sect. 2, Literature Review. The proposed framework for secure communication. presented in Sect. 3. Finally, the security and cost analysis are presented in session 4.

2 Literature Review

Various techniques illustrated below which are used for security in VANET: Table 1 shows the review of some security schemes.

3 Proposed Approach

3.1 Traffic Simulation

Using the SUMO simulator we can create the traffic simulation Fig. 2. To create the simulation, first using the osmWebWizard we can open the real-world map using the Open street map and then import that real-world map to the sumo simulator with mobility. tcl file which is used for network simulation [16].

3.2 Proposed Enhanced AKE Framework

This part introduces the suggested enhanced Authenticated Key Exchange (AKE) framework for V2V communication. Communicating parties share the same secret key securely [17]. The Public Key Generator (PKG) is the third organization that may be regarded as a reliable source for creating secret keys for vehicles. The five stages of the suggested AKE framework are setup, private key extraction, generating nonce, key exchange, and multi-factor authentication. Some of the symbols used in this work are enumerated in Table 2 for ease of presentation before they are addressed.

3.3 Setup of PKG

At this point, the PKG initializes a few significant numbers and makes them accessible publicly [18]:

1. PKG chooses a q prime value which is in a bit, where its private code is q. The PKG generates P work like a generator for a GG (cyclic additive group) of the k-order on the area $\mathbb{EE}/\mathbb{F}_p$.

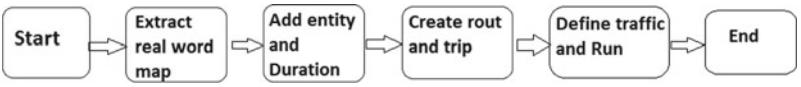


Fig. 2 Workflow for sumo and NS

Table 2 Symboles table

Symbols	Explanation	Symbols	Explanation
s_{pkg}	Secrete key of PKG	$sk_i = (x_i, p_i)$	i's secrete key
PUB_{pkg}	PKG's public key	n_i	i's Ephemeral private key
HS_1, HS_2	Hash functions	N_i	i's Ephemeral public key
tm_i	Valid duration of ω_i	θ_i	i's Public information
$pk_i = (X_i, R_i)$	i's public key	TT_i	i's transaction information
K_{ij}, K_{ji}	i's and j's shared secret key	SK_{ij}, SK_{ji}	i's and j's session key

2. PKG uses the master secret key, which is the master private key, $s_{\text{pkg}} \in \mathbb{Z}_p^*$ to create the public key $PUB_{\text{pkg}} = s_{\text{pkg}} P$.
3. The hash functions used in the framework are defined by the PKG as follows: $HS_1 : \{0, 1\}^* \| GG \| GG \| \rightarrow \mathbb{Z}_p^*$ and $HS_2 : \{0, 1\}^* \| \{0, 1\}^* \| \{0, 1\}^* \| GG \| GG \rightarrow \{0, 1\}^q$, here q denotes the how the large is the session key.
4. The factors $q, k, EE/\mathbb{F}p, P, GG, PUB_{\text{pkg}}, HS_1$ and HS_2 are published by PKG.

3.4 Private Key Extraction

After this, Half public and secret keys created by PKG for all nodes [18]:

1. Each vehicle selects a random partial private key x_i from the set of non-zero integers. Using this partial private key, the vehicle computes a partial public key X_i, which is equal to the multiply of x_i with a set point P on an elliptic curve. The vehicle also constructs a message $\theta_i = (\text{ID}_i || X_i || tm_i)$ that includes its identification, the computed partial public key X_i, and the vehicle's legitimate time t_i. The message is then sent to the PKG.
2. Once the PKG receive θ_i , it validate the sustainability of the timestamp tm_i. If it still holds true, $r_i \in \mathbb{Z}_p^*$ is randomly selected, and $R_i = r_i P$, $h_i = HS_1(\text{ID}_i || R_i || X_i)$, and the partial privet key as $p_i = r_i + h_i s_{\text{pkg}}$ (modular of k) are computed. The partial public key (R_i), and partial private key(p_i) are sent by PKG over a safe channel to the i-vehicle.
3. Now vehicle has its own partial public key (X_i), PKG through generated partial public key (R_i) and has its own partial secrete key (x_i), PKG through generated partial secrete key (p_i)

3.5 Generate Nonce

In light of the fact that vehicle A, using the identity ID_A, is a beginner while vehicle B, using the identity ID_B, is a responder.

1. Vehicle i dispatches to PKG $\theta_i = (\mathrm{ID}_i||\mathrm{ID}_j||tm_i)$, where ID_i known as the vehicle i's identity and tm_i is the θ_i's valid time.
2. After PKG taking θ_i, it validate the sustainability of the timestamp tm_i if it looks valid then generates random the $n_1 \in \mathbb{Z}_p^*$. Then, the PKG sends this nonce to both vehicles.

3.6 Key Exchange

In this phase, vehicles send the secret key to each other securely [18].

1. By calculating $N_A = n_A P$, which is Vehicle A's ephemeral public key. As its ephemeral private key, Vehicle A chooses a random element $n_A \in Z_p^*$. Then, tm'_A, which denotes the valid period of θ'_A, is calculated as $TT_A = TT_{A1}||TT_{A2}$ and $\theta'_A = (\mathrm{ID}_A||X_A||tm'_A||h(n_1))$, are then computed as $TT_{A1} = p_A N_A$ and $TT_{A2} = x_A N_A$,where tm'_A is the sustainable time of θ'_A. Then, vehicle A forward vehicle B on its way along with θ'_A, R_A, and TT_A.
2. Now vehicle B selects the private key $n_B \in Z_p^*$, Vehicle B calculates $N_B = n_B P$to make a transient public key. Following this, Vehicle B determines $TT_{B1} = p_B N_B$ and $TT_{B2} = x_B N_B$, sets $TTB = TT_{B1}||TT_{B2}$ and $\theta'_B = (\mathrm{ID}_B||XB||tm'_B||h(n_1))$. Vehicle B transmits. θ'_B, R_B, and TT_B to vehicle A.
3. Vehicle A first verifies the accuracy of tm'_B and the $h(n_1)$ is the same as the PKG sent and whether ID_B is the intended identity after getting the message from vehicle B. In the case that all of them are accurate, car A determines $K_{AB1} = n_A p_A(R_B + HS_1(ID_B||X_B||R_B)PUB_{\mathrm{pkg}}) + p_A TT_{B1}$, $K_{AB2} = x_A n_A TT_{B2}$, and session key
$SK_{AB} = HS_2(ID_A||ID_B||\theta_A||\theta_B||TT_A||TT_B||K_{AB1}||K_{AB2})$.
4. Vehicle B first verifies the accuracy of tm'_A and the $h(n_1)$ is the same as the PKG sent, and whether ID_A is the intended identity after getting the message from vehicle A. Vehicle B determines the session secret if all of them are accurate:
$SK_{BA} = HS_2(ID_A||ID_B||\theta_A||\theta_B||TT_A||TT_B||K_{AB1}||K_{AB2})$,
$K_{BA1} = n_B p_B(R_A + HS_1(ID_A||R_A||X_A)PUB_{\mathrm{pkg}}) + p_B TT_{A1}$,
$KAB2 = x_B n_B TT_{A2}$

Node A and B have the similar session key.

3.7 Multi-factor Authentication

Now both vehicles have the same secret key for communication. To be more secure about the secret key the multi-factor authentication can be added to confirm that both vehicles have the same secret key communication.

1. Vehicle A and Vehicle B each possess a physical token that stores a secret key.
2. When Vehicle A initiates the key exchange process, it sends a message to Vehicle B that includes a challenge based on a random number generated by Vehicle A: $C_A = (g^r)_{X_i}$.
3. Vehicle B receives the challenge and generates a response by apply a hash function to the challenge and a secret key stored in its physical token: $s_B = H(C_A, K_{BA1})$. where K_{BA1} is the secret key stored in Vehicle B's physical token, and cryptography hash function is known as a H.
4. Vehicle A verifies the response by using the same hash function to the challenge and the secret key stored in its physical token: $s_A = H(C_A, K_{AB1})$. where K_{AB1} is the secret key stored in Vehicle A's physical token. Vehicle A checks if $s_A = s_B$. If the computed response matches the received response, Vehicle A considers Vehicle B to be authenticated.

After the whole key distribution process, to transfer the message between nodes "RC4" algorithm is used which is a stream cipher method with variable duration [19].

4 Security Analysis and Correctness of Proposed Framework

4.1 Security Analysis

Forward Security is achieved using ephemeral keys. Impersonation Resilience is achieved by considering the nonce and time. Man-in-the-middle attacks prevented using the nonce and multi-factor authentication. Even if your ephemeral key is disclosed, the session key is not affected by that. PKG has the details of which nodes are communicating with each other using this you can prevent the impersonation [13]. We have provided the correct evidence in this part to demonstrate the consistency of the session keys for the two entities. Due to K_{AB1}, K_{BA1}, K_{AB2}, and K_{BA2} being the only differences between the two groups determine the session secret. We only need to demonstrate that $K_{AB1} = K_{BA1}$ and $K_{AB2} = K_{BA2}$, and the following formulas show that they are correct:

Table 3 comparative study of security features

Frameworks	AB1	AB2	AB3	AB4	AB5	AB6	AB7	AB8
Li et al. [20]	P	P	P	P	P	P	NP	NP
Lei et al. [18]	P	P	P	P	P	P	P	NP
Gupta et al. [21]	P	P	P	P	NP	P	NP	NP
Proposed framework	P	P	P	P	P	P	P	P

**Note* Frequently changes key(AB1); Used strong temporary keys(AB2); PKG changes keys(AB3); Key-negotiation spoof resilience(AB4); Fading confidential reveal protection(AB5);Key control(AB6); PKG can't access full key of node(AB7);Multi-factor authentication(AB7); Preserve(P); Not preserved(NP) ;

$$
\begin{aligned}
K_{AB1} &= n_A p_A (R_B + HS_1(ID_B||X_B||P_B)PUB_{\text{pkg}}) + p_A TT_{B1} \\
&= n_A p_A (R_B + h_B(PUB_{\text{pkg}})) + p_A p_B n_B P \\
&= n_A p_A (r_B + h_B s_{\text{pkg}})P + p_A p_B n_B P \\
&= n_A p_A p_B P + p_A p_B n_B P \\
&= n_B p_B p_A P + p_B p_A n_A P \\
&= n_B p_B (h_A s_{\text{pkg}} + r_A)P + TT_{A1} p_B \\
&= p_B n_B (H_1(ID_A||X_A||R_A)PUB_{\text{pkg}} + R_A) + TT_{A1} p_B \\
&= K_{BA1}
\end{aligned}
$$

4.2 Comparative Analysis of Proposed Framework with Other Frameworks with Respect to Security, Computing and Communication Cost

The comparative outcome in terms of security features is displayed in Table 3. Clearly, the suggested AKE framework complies with all of the security requirements listed in the chart including multi-factor authentication others can't have that. On the basis of Table 4, key extraction and exchange times for each enhanced AKE framework can be approximated. It shows the time taken by each phase and storage cost with the message size used by the framework.

5 Conclusion and Future Scope

We proposed a framework that resists security attacks like disclosure of ephemeral keys and absence of complete key escrow and provides multi-factor authentication which gives it higher security in V2V communication. We highlighted the challenges and issues that are at present in IoV. With security, we presented a rough estimation of computation cost for each phase computation time, message size and storage cost

Table 4 Comparison of computation cost

Computation cost				
Framework	Key extraction (ms)	Key exchange and multi-factor authentication (ms)	Storage cost (bytes)	Message size (bytes)
Li et al. [20]	6MP + 4HF ≈ 2.00	12MP + 6HF ≈ 3.90	200	200
Gupta et al. [21]	$2PM_{G1}$ + 2HF ≈ 2.18	4MP + $8PM_{G1}$ + 3HF ≈ 11.27	296	360
Lei et al. [18]	4MP + 2HF ≈ 1.30	14MP + 4HF ≈ 4.40	240	360
Proposed framework	4MP + 2HF ≈ 1.30	15MP + 8HF ≈ 9.5	240	360

**Note MP* point multiplication using elliptic curve cryptography, *HF* hash function, PMG1 group multiplication operation in G1 that utilizes pairings

which suggests good efficiency. As an outcome, we can say that our framework optimizes security for effective V2V communication. In the future, the implementation of proposed work in different scenarios help to optimize it further.

References

1. Alharthi A, Ni Q, Jiang R (2021) A privacy-preservation framework based on biometrics blockchain (BBC) to prevent attacks in VANET. IEEE Access 9:87299–87309
2. Zhou H, Xu W, Chen J, Wang W (2020) Evolutionary v2x technologies toward the internet of vehicles: challenges and opportunities. Proc IEEE 108(2):308–323
3. Wang B, Wang Y, Chen R (2019) A practical authentication framework for VANETs. Secur Commun Netw
4. Lu Z, Liu W, Wang Q, Qu G, Liu Z (2018) A privacy-preserving trust model based on blockchain for VANETs. IEEE Access 6:45655–45664
5. Ayyappan B, Kumar PM (2016) Vehicular ad hoc networks (VANET): Architectures, methodologies and design issues. In: 2016 second international conference on science technology engineering and management (ICONSTEM), pp 177–180
6. Khan MA (2021) Secure message transmission for v2v based on mutual authentication for VANET. Wirel Commun Mob Comput
7. Zarbi S, Mortazavi SA, Salehpour P (2020) Security analysis of an efficient authentication scheme for vehicular ad hoc networks. In: 2020 17th international ISCISC. IEEE, pp 44–47
8. Hasrouny H, Bassil C, Samhat AE, Laouiti A (2015) Group-based authentication in v2v communications. In: 2015 fifth international conference on DICTAP. IEEE, pp 173–177
9. Alfadhli SA, Lu S, Chen K, Sebai M (2020) Mfspv: A multi-factor secured and lightweight privacy-preserving authentication scheme for VANETs. IEEE Access 8:142858–142874
10. Song Y, Fu Y, Yu FR, Zhou L (2020) Blockchain-enabled internet of vehicles with cooperative positioning: a deep neural network approach. IEEE IoT J 7(4):3485–3498
11. Bodkhe U, Tanwar S (2021) P2iov:privacy preserving lightweight secure data dissemination scheme for internet of vehicles. In: 2021 IEEE Globecom workshops (GC Wkshps), pp 1–6
12. Muhammad A, Khan TA, Abbass K, Song WC (2020) An end-to-end intelligent network resource allocation in iov: a machine learning approach. In: 2020 IEEE 92nd vehicular technology conference (VTC2020-Fall), pp 1–5

13. Cui J, Ouyang F, Ying Z, Wei L, Zhong H (2021) Secure and efficient data sharing among vehicles based on consortium blockchain. IEEE Trans Intell Transp Syst
14. Labrador M, Hou W (2019) Implementing blockchain technology in the internet of vehicle (IoV). In: 2019 international conference on intelligent computing and its emerging applications (ICEA), pp 5–10
15. Stępień K, Poniszewska-Marańda A (2020) Security methods against black hole attacks in vehicular ad-hoc network. In: 2020 IEEE 19th international symposium on network computing and applications (NCA). IEEE, pp 1–4
16. Amina B, Mohamed E (2018) Performance evaluation of VANETs routing protocols using sumo and NS3
17. Harn L, Hsin WJ, Mehta M (2005) Authenticated Diffie–Hellman key agreement protocol using a single cryptographic assumption. IEE Proc Commun 152(4):404–410
18. Meng L, Xu H, Xiong H, Zhang X, Zhou X, Han Z (2021) An efficient certificateless authenticated key exchange protocol resistant to ephemeral key leakage attack for v2v communication in IoV. IEEE Trans Veh Technol 70(11):11736–11747
19. Xu H, Zeng M, Hu W, Wang J (2019) Authentication-based vehicle-to-vehicle secure communication for VANETs. Mob Inf Syst
20. Li Q, Hsu CF, Raymond Choo KK, He D (2019) A provably secure and lightweight identity-based two-party authenticated key agreement protocol for vehicular ad hoc networks. Secur Commun Netw 2019:1–13
21. Gupta DS, Islam SH, Obaidat MS, Vijayakumar P, Kumar N, Park Y (2020) A provably secure and lightweight identity-based two-party authenticated key agreement protocol for IIoT Environ IEEE Syst J 1–10

ChatGPT: An Advanceds Natural Language Processing System for Conversational AI Applications—A Comprehensive Review and Comparative Analysis with Other Chatbots and NLP Models

Kumar Keshamoni

Abstract ChatGPT is an OpenAI-developed big-language model architecture based on the GPT-3.5 standard. It is aimed at engaging users in natural language talks on a variety of themes. With its comprehensive knowledge of English and the world, input from users can result in replies that resemble those of humans. It can understand and react to a variety of linguistic patterns and idioms because it was trained on a vast corpus of Internet content. Numerous programmes, such as customer support systems, virtual assistants, and chatbots have already implemented ChatGPT. Potentially useful for communication, education, and entertainment.

Keywords ChatGPT · Google · User · Data · Responses · Model · API · Tasks · Processing

1 Introduction

Elon Musk, Sam Altman, Greg Brockman, Ilya Sutskever, John Schulman, and Wojciech Zaremba co-founded OpenAI. Elon Musk, Sam Altman, Greg Brockman, Ilya Sutskever, John Schulman, and Wojciech Zaremba also developed a language model. The goal of OpenAI is to create AI that is both beneficial and safe for everyone. One of the company's most important products in reaching this goal is ChatGPT. The outcome of years of deep learning and natural language processing research by OpenAI and its partners is ChatGPT. A language model called Chat Generative Pre-trained Transformer for Conversational AI (ChatGPT) was created by OpenAI, a research group committed to the safe and useful development of artificial intelligence. Elon Musk, Sam Altman, Greg Brockman, Ilya Sutskever, John Schulman,

K. Keshamoni (✉)
Assistant Professor, Department of ECE, Vaagdevi Engineering College, Warangal, India
e-mail: kumar.keshamoni@gmail.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_40

and Wojciech Zaremba were among the co-founders of OpenAI [1]. Alec Radford, Jeffrey Wu, Rewon Child, David Luan, and Dario Amodei led a team of OpenAI researchers that created ChatGPT. OpenAI made the first iteration of GPT, known as GPT-1, available in June 2018, and GPT-3 will follow in June 2020 [2].

A language model called ChatGPT was created by the OpenAI artificial intelligence research organization. It is a subset of the Generative Pre-trained Transformer (GPT) family of models, which are built on the transformer architecture and pre-trained using unsupervised learning methods on substantial amounts of text data. ChatGPT aims to produce human-like answers to inputs in natural language. It is intended to carry out a variety of tasks, such as text summarization, question response, and language translation. ChatGPT is perfectly suited for tasks requiring natural language semantic and contextual interpretation. One distinguishing feature of ChatGPT is its capacity to produce text that is coherent, grammatically accurate, and written in a style like that of human writing. The model can recognize long-range relationships in the input sequence and provide output based on the context of the full input, thanks to the use of a self-attention mechanism.

ChatGPT comes in different versions, like the GPT-2 and GPT-3 models, which are bigger and more powerful than the ones before them. GPT-3 has gotten a lot of attention because it does a great job at a wide range of natural language processing tasks, such as translating languages, answering questions, and filling in blanks in text [3]. Even though ChatGPT and other language models have made a lot of progress in natural language processing, there are worries about how they could be used wrongly or lead to unintended results. This includes the production of false news, hate speech, and other damaging content. As a result, researchers continue to investigate means of mitigating these risks and ensuring the responsible development and application of language models.

OpenAI's Chat Generative Pre-trained Transformer (ChatGPT) is a pre-trained language model. It is a state-of-the-art language model that generates natural language text in response to a given prompt using deep neural networks. ChatGPT is trained on a large corpus of diverse text data and has a parameter capacity of up to 13.5 billion, making it one of the largest and most potent language models currently available.

One thing that makes ChatGPT stand out is that it can make high-quality text that makes sense and fits the situation. It can do many things with natural language, like translate, answer questions, finish sentences, and make up dialogue. ChatGPT can also be fine-tuned for specific tasks using additional training data, allowing it to generate responses that are even more accurate and contextually relevant. ChatGPT builds on the success of earlier language models such as BERT and GPT-2, but incorporates architectural and methodological enhancements to achieve even higher performance. It employs a transformer-based architecture, which enables it to recognize long-term dependencies in text and generates high-quality responses consistent with the given context. ChatGPT has practical applications in several industries, including customer service, chatbots, and content generation, in addition to its use in research and development. It is important to note, however, that ChatGPT is still

a machine learning model and may have limitations and biases that must be taken into account when implementing it in real-world applications.

2 Literature Review

Chat Generative Pre-trained Transformer (ChatGPT) is a large language model that has received considerable attention in the field of natural language processing (NLP) since its release in 2020 by OpenAI. The model is pre-trained on a massive amount of text data and is based on the GPT-3 architecture to generate human-like responses to a given prompt. Several evaluations of ChatGPT's performance and capabilities have been conducted. Below is a literature review of several of these studies.

In a study by Adiwardana et al. [4], the authors evaluated ChatGPT's performance on a variety of language tasks, including translation, question-answering, and text summarization. ChatGPT outperformed several state-of-the-art models on these tasks, according to the research. Liu et al. [5] evaluated the capacity of ChatGPT to produce human-like responses in a conversational context. The authors conducted a human evaluation of ChatGPT's responses and found that the model could generate coherent and contextually relevant responses that were difficult to distinguish from those generated by a human. Zhang et al. [6] investigated the use of ChatGPT in a customer service environment. The authors trained the model using a dataset of customer service conversations and assessed its performance using a test set of customer queries. The study found that ChatGPT was able to generate contextually relevant and informative responses. In a study conducted by Huang et al. [7], the authors evaluated ChatGPT's performance on a variety of medical-related language tasks. ChatGPT was able to generate responses that were highly pertinent to medical queries and exhibited a solid grasp of medical terminology, according to the study.

In general, the literature indicates that ChatGPT is a highly capable language model that can generate human-like responses to a wide variety of language tasks. The model has demonstrated particular promise in conversational settings and has applications in customer service, medical information, and other fields. However, additional research is required to fully comprehend the model's capabilities and limitations.

3 Integration and Implementation

3.1 Implementation

ChatGPT can be implemented using OpenAI's API or by training your own model using the available codebase. Here's a general overview of the steps involved in implementing ChatGPT:

(i) Data Collection: Collect relevant data to train the ChatGPT model. This can include chat logs, customer service transcripts, and other text-based data.
(ii) Pre-processing: The data needs to be pre-processed to ensure that it is in a format suitable for training the ChatGPT model. This can include tasks such as tokenization, data cleaning, and data augmentation.
(iii) Training: Train the ChatGPT model using the pre-processed data. This involves using deep learning techniques such as backpropagation to adjust the model's weights and biases.
(iv) Validation: Validate the model to ensure that it is generating high-quality responses. This can be done by using a separate validation set of data to test the model's performance.
(v) Deployment: Deploy the ChatGPT model in a suitable environment, such as a web server or chatbot platform, where it can generate responses to text prompts in real-time.
(vi) Maintenance: Continuously monitor the ChatGPT model's performance and update it as necessary. This may involve retraining the model with new data or tweaking the model's hyperparameters to improve its performance.

Overall, implementing ChatGPT requires a deep understanding of machine learning, natural language processing, and software engineering principles.

3.2 Chronological Development

ChatGPT is part of a family of language models developed by OpenAI, each of which has its own history, versions, and revisions. Here is a brief overview of the history of ChatGPT:

- **GPT-1**: OpenAI introduced the first iteration of Generative Pre-trained Transformer (GPT) in June 2018. It was pre-trained on a vast corpus of text data and had 117 million parameters.
- **GPT-2**: OpenAI published GPT-2 in February 2019, a more advanced language model with 1.5 billion parameters. GPT-2 generated controversy due to its capacity to generate highly plausible false news articles and other deceptive content.
- **GPT-3**: With 175 billion parameters, GPT-3 was introduced in June 2020 and is presently the most advanced version of the language model. It can generate highly coherent and contextually appropriate responses to a broad variety of text prompts, making it suitable for a wide range of applications, including chatbots, content creation, and language translation.
- **GPT-3.5**: In November 2021, OpenAI announced an update to GPT-3, referred to as GPT-3.5. This update added new features such as better handling of code and programming languages, as well as improvements to the model's memory and efficiency.

Each version of ChatGPT has undergone multiple revisions and improvements, as OpenAI continues to refine and optimize the language model. These revisions have included changes to the model architecture, the training data used, and the fine-tuning process.

3.3 Integration

Integrating ChatGPT into your website or application can be done using an API provided by OpenAI. Here are the basic steps to integrate ChatGPT:

(a) Sign up for OpenAI API access: Go to OpenAI's website and sign up for access to the GPT API.
(b) Choose a plan: Choose a plan that suits your needs based on your expected usage and budget.
(c) Get your API key: Once you have signed up and selected a plan, you will receive an API key that you will use to authenticate your requests.
(d) Integrate the API into your website or application: Use the API documentation provided by OpenAI to integrate the ChatGPT API into your website or application. This may involve using a software development kit (SDK) provided by OpenAI or writing custom code to interact with the API.
(e) Test and refine: Once the integration is complete, test it thoroughly to ensure it is working as expected. Refine the integration as needed based on user feedback and usage patterns.

Keep in mind that integrating ChatGPT into your website or application is just the first step. You will need to continually monitor and improve the integration to ensure it is delivering the best user experience and meeting your business goals.

3.4 Security

ChatGPT is a language model developed by OpenAI, and the security of the platform is a top priority. OpenAI takes several measures to ensure the security and privacy of the platform, including:

(a) Data Privacy: OpenAI uses strict data privacy policies to ensure that user data is kept confidential and secure.
(b) Encryption: All communication with the ChatGPT platform is encrypted using secure protocols to prevent unauthorized access.
(c) Access Control: OpenAI restricts access to the platform to authorized personnel only, and all activity on the platform is monitored and audited.
(d) Regular Updates: OpenAI regularly updates the platform's security protocols and software to protect against new threats and vulnerabilities.

However, as with any online platform, there is always a risk of security breaches, and users should take necessary precautions, such as not sharing sensitive information or personal data, when using the ChatGPT platform. To ensure the privacy and security of users, it is important to follow best practices for data protection, including encryption of data in transit and at rest, secure storage and handling of user data, and regular security audits and testing.

4 Alternatives and Competitor's

Depending on the specific use case and requirements, several other language models and AI systems could be considered counterparts to Chat Generative Pre-trained Transformer for Conversational AI (ChatGPT). Here are some of the most notable rivals:

- **BERT (Bidirectional Encoder Representations from Transformers)**: Google's BERT language model is extensively used in natural language processing (NLP) applications such as question-answering, sentiment analysis, and language translation.
- **Transformer-XL**: Transformer-XL is a language model developed by Carnegie Mellon University that handles lengthy text sequences more efficiently than traditional transformer models, making it suitable for language modelling and text generation.
- **CTRL (Conditional Transformer Language Model)**: Salesforce's CTRL language paradigm is designed to generate text based on specific control codes, enabling more precise and targeted text generation.
- **GShard**: Google's GShard is a framework for distributed machine learning designed to efficiently train large-scale language models using parallel processing and distributed computation.
- **T5 (Text-to-Text Transfer Transformer)**: Google's T5 language model is designed to manage a wide variety of NLP tasks using a single, unified architecture, making it a versatile and flexible language processing solution.
- **Microsoft DialoGPT**: Microsoft's DialoGPT is a language paradigm designed specifically for the generation of human-like dialogues. It is pre-trained on a large corpus of dialogues and can be customized for particular duties, such as customer service or chatbots.
- **Amazon Lex**: Amazon Lex is a conversational artificial intelligence service that can be used to create chatbots and virtual assistants. It employs machine learning algorithms to comprehend natural language and generate responses that mimic human behaviour.

Chat Generative Pre-trained Transformer for Conversational AI (ChatGPT) is a state-of-the-art language model that has set a high standard for natural language processing. However, there are several other language models and conversational AI

systems that could be considered competitors or alternatives to ChatGPT. Here are a few examples.

5 Chatbots Versus ChatGPT

There are many chatbots available in the market, and each has its own unique features and capabilities. Here are some popular chatbots and how they compare to ChatGPT:

- **Siri**: Siri is a voice assistant developed by Apple that can understand natural language commands and answer questions. However, Siri's capabilities are limited to predefined commands and cannot generate human-like responses like ChatGPT.
- **Alexa**: Alexa is a voice assistant developed by Amazon that can perform a variety of tasks, including playing music, setting reminders, and answering questions. However, like Siri, Alexa's responses are limited to predefined commands and cannot generate human-like responses.
- **Google Assistant**: Google Assistant is a voice assistant developed by Google that can perform a wide range of tasks, including setting reminders, playing music, and answering questions. However, like Siri and Alexa, Google Assistant's responses are limited to predefined commands and cannot generate human-like responses.
- **Mitsuku**: Mitsuku is a chatbot that has won several awards for its conversational abilities. However, its responses are still limited to predefined scripts and cannot generate human-like responses like ChatGPT.
- **Xiaoice**: Xiaoice is a Chinese chatbot developed by Microsoft that has been praised for its conversational abilities. However, like Mitsuku, its responses are still limited to predefined scripts and cannot generate human-like responses like ChatGPT.
- **ELIZA**: ELIZA is one of the earliest chatbots, created in the 1960s. It used simple pattern matching and substitution to generate responses to user inputs. ELIZA was designed to simulate a Rogerian psychotherapist, and it could ask users open-ended questions and provide reflective responses. Compared with ChatGPT, ELIZA is much simpler in terms of its capabilities and technology.
- **Replika**: Replika is a chatbot that is designed to be a personal AI companion. It uses machine learning to understand and respond to user inputs, and it is designed to learn about the user over time and personalize its responses. Replika is also capable of generating its own questions and conversation topics. Compared with ChatGPT, Replika has a more narrow focus on providing emotional support and companionship to users.
- **ChatGPT**: ChatGPT is a language model developed by OpenAI and based on the GPT architecture. It uses deep learning and natural language processing to generate responses to user inputs. ChatGPT is not designed to have a personality or personalization capabilities, but it is capable of generating highly sophisticated

and coherent responses on a wide range of topics. Compared with the other chatbots mentioned here, ChatGPT is more advanced in terms of its capabilities and technology.

Overall, each of these chatbots has its own advantages and disadvantages, and the best one for a given situation will hinge on the application's particular requirements and objectives. ChatGPT, on the other hand, is a large language model that has been trained on a massive text dataset and can generate human-like responses to a broad variety of questions and topics. In addition to learning and adapting to new information, ChatGPT is more versatile and adaptable than its predecessors.

6 Applications

ChatGPT has a wide range of applications in various fields, including:

- Chatbots: ChatGPT can be used to create chatbots capable of conversing with consumers and providing them with customized recommendations and solutions.
- Virtual assistants: ChatGPT can be used to create virtual assistants capable of performing a variety of duties, including scheduling appointments, setting reminders, and answering questions.
- ChatGPT can be used to improve customer service by providing prompt and accurate responses to customer inquiries.
- ChatGPT can be used to develop language translation systems capable of translating text from one language to another. ChatGPT can generate content such as articles, product descriptions, and social media postings.
- Education: ChatGPT can be used to create educational tools, such as chatbots that aid students in their studies.
- Healthcare: ChatGPT can be used to create virtual assistants capable of assisting patients with medical questions and providing personalized recommendations.

Overall, ChatGPT has vast potential to revolutionize various industries by providing intelligent and personalized conversational experiences to users.

7 Conclusions

Here are some concluding remarks for ChatGPT based on its characteristics, implementation, applications, multi-language support, security, chatbots, and Google search engine integration:

Characteristics: ChatGPT is a large language model based on the GPT-3.5 architecture, designed to provide information, answer questions, and assist in various tasks. It has a vast knowledge base and can understand and generate human-like language.

Implementation: ChatGPT is implemented using deep learning techniques and is trained on a massive dataset of text data. It is hosted on servers and can be accessed through various interfaces, such as chatbots, web applications, or APIs.

Applications: ChatGPT has many applications, including customer support, virtual assistants, education, healthcare, and more. It can also be used for content generation, language translation, and even creative writing.

Multi-language support: ChatGPT supports multiple languages, making it useful for a global audience. It can understand and generate text in various languages, including English, Spanish, French, and more.

Security: ChatGPT's security is paramount, and measures are taken to ensure user data is safe and secure. Confidentiality and integrity of user data are a top priority.

Chatbots: ChatGPT can be integrated with chatbots to provide a conversational interface for users to interact with. This makes it easier for users to get the information they need quickly and efficiently.

Google search engine integration: ChatGPT can also be integrated with Google search engine to provide relevant information to users based on their queries.

Overall, ChatGPT is a powerful language model with numerous applications, multi-language support, and robust security measures. Its integration with chatbots and Google search engine makes it even more versatile and useful.

References

1. Dwivedi YK, Kshetri N, Hughes L, Slade EL et al (2023) "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int J Inf Manag
2. Qureshi R, Irfan M, Ali H, Khan A et al (2023) Artificial intelligence and biosensors in healthcare and its clinical relevance: a review. Institute of Electrical and Electronics Engineers (IEEE)
3. Chu Z (2020) On learning and applying text representations. The University of Chicago
4. Adiwardana D, Luong MT, Kudlur M, So DR, Fergus R (2020) Towards a human-like open-domain chatbot. Advances in neural information processing systems, 33
5. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V(2020) Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach. arXiv preprint arXiv:1911.05287.
6. Zhang Y, Sun S, Galley M, Chen YC, Brockett C, Gao X, Gao J, Liu J, Dolan B (2021) ReCoSa: detecting the relevant contexts with self-attention for multi-turn dialogue generation. arXiv preprint arXiv:2109.00240.
7. Huang KY, Lin JY, Chen CH, Lee CH (2021) RefGPT: a dual-attention model for retrieval-augmented coherent text generation. arXiv preprint arXiv:2108.06722

Two-Step Method for Extracting Emotion Cause Pairs from Texts

Rohini Mutalikdesai, K. G. Karibasappa, and K. Karibasappa

Abstract The main purpose of emotion cause extraction (ECE) is to find all the causes of certain emotions. This task has become an important topic for research in recent years since it may be used to examine the causes of distinct emotions in the text. However, annotated datasets are rare in research, particularly in languages other than Chinese. The main drawback of the ECE task is that emotions should be labeled before the real cause can be identified, which is practically challenging. To address this issue, emotional cause pair extraction (ECPE), a new work, has been suggested. It attempts to find potential emotion pairings and the causes of those pairings in a given document. Given a document, the first stage identifies all emotion and cause sentences, and the second stage identifies the correct emotion-cause pairs. This work demonstrates the task's performance in extracting emotion-cause pairs and identifying the causes of emotions in texts.

Keywords Pair extraction · Emotion cause extraction · Corpus · Semantic rules · Correlation

1 Introduction

Text analysis is essential for understanding the moods and emotions behind text-mediated communication. Recent research has looked at the fundamental elements of emotions in addition to categorizing them. The emotion cause extraction (ECE) task requires a deeper comprehension of the text's emotional content and the recognition of possible causes for a specific emotion. Xia and Ding [1] and Gui et al. [2] provided

R. Mutalikdesai · K. G. Karibasappa
School of Computer Science and Engineering, KLE Technological University, Hubballi, Karnataka 580031, India
e-mail: karibasappa_kg@kletech.ac.in

K. Karibasappa (✉)
Department of Computer Science and Engineering, Graphic Era (Deemed to be) University, Dehradun 248002, India
e-mail: kwadikiKaribasappa.cse@geu.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_41

datasets to deal with word-level cause extraction, respectively. The ECE requires the annotation of emotions prior to the extraction of causes, which takes time.

An example of this corpus is illustrated in Fig. 1a. There are five sentences in Fig. 1a. The word "happy" appears in the fourth sentence. This is known as an emotional sentence. The emotional sentence "happy" has two interconnected reasons: "a mother returned home from work with a surprise gift" and "told her daughter that the gift was for her", in the second and third sentences, respectively. These are known as causal clauses. This is the example clause that contains the cause. However, the existing ECE task requires emotions to be labeled before determining causes, which limits its utility. Emotion cause pair extraction (ECPE) is a technique used to identify all correct emotional pairs and underlying causes from a file.

The distinction between traditional ECE and new ECPE tasks is represented in Fig. 1. Traditional ECE seeks to extract causal clauses from documents with annotated sentiments, while new ECPE tasks yield emotion-cause pairs without needing annotations. If the sentiment annotation in Fig. 1a is "happy", the aim of the ECE task is to extract the associated causal clauses: "a mother returned home from work with a surprise gift" and "told her daughter that the gift was for her". While the objective of the ECPE job is to instantly extract all potential emotion-cause pairs, including ("The daughter was very happy", "a mother returned home from work with a surprise gift") and ("The daughter was very happy", "told her daughter that the gift was for her"), without providing the emotion annotation "happy". The objective of emotion cause pair extraction (ECPE) is to investigate all possible emotion pairings and the causes that correspond with them in a file. In contrast to the thoroughly studied emotion cause extraction (ECE), ECPE does not call for sentiment phrases to be annotated.

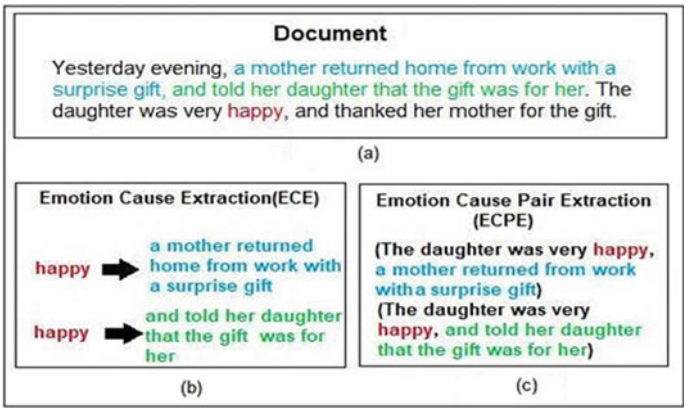


Fig. 1 Example of the corpus (**a**), distinction between the ECE (**b**) and ECPE (**c**) tasks

2 Related Work

We have divided all the literature survey papers into three categories based on: (1) Co-reference, (2) Semantic features, and (3) Both Co-reference and Semantic features.

2.1 Co-reference

Xia and Ding [1] proposed a two-stage approach for the emotion-cause pair extraction task, which showed a significant increase in accuracy scores. Yu et al. [3] also proposed a distillation method to mitigate the effects of error propagation. A basic pipelined strategy based on two separate encoders was presented [4], which improved the absolute $F1$ by 5.20% and 5.85%, respectively. Li et al. [5] proposed a common learning framework for emotion root cause span extraction and sentiment classification using BERT and a standard dataset. The ECSP task is presented in [6], which involves extracting and classifying emotions and causes from documents. The suggested strategy successfully achieved an outstanding $F1$ score of 88.71%. A multi-task learning strategy is proposed [7] to extract emotions, causes, and emotion-cause pairs. The model learns to link cause clauses to emotion clauses and considers pair extraction as a link prediction issue. Two machine learning-based root cause detection techniques have been created [8], with the rule-based method having an accuracy rate of 68.30% and the conditional random field model-based technique having 77.57%. UTOS is a sequence model [9] that uses end-to-end unified sequencing to detect causal pairs and extract emotion-cause pairs in a single pass of sequential labeling.

2.2 Semantic Features

Gui et al. [2] proposed an emotion-cause extraction method from news text using a kernel technique to identify the emotion's underlying cause. This work must be translated into English, and a bilingual solution must be looked into. A special computational multi-granularity bidirectional attention stream network is built [10] using a machine-understanding frame and can increase the $F1$ by 2.6%. Finding the right emotion in a phrase can be difficult due to adversative words. The proposed social emotion cause extraction system [11] uses a three-component lexicon-enhanced network to learn emotions through dynamic training with high precision, recall, and $F1$ score. A dual MRC framework is introduced [12], which extracts emotion-cause pairings in both directions, in the E-C and C-E directions. The C-E strategy performs 5.61 and 3.78% better in $F1$ than the top baseline method, ISML. A proposed neural network [13] uses span association to extract emotion and cause pairs from sentences. The model surpassed its $F1$ score by 1.86, 3.10, and 5.66% on three subtasks. Multi-Granularity Semantic-Aware Graph Model [14] that combines both precise- and

large-grained semantic characteristics has the highest $F1$ score for the ECPE task. DQAN was developed [15] to address the error propagation issue and outperforms the best ECPE-based method by 7.3% and CANN-E by 84% on the Chinese dataset. Cao et al. proposed [16] the Causal Narrative Comprehension Model (CNCM) for emotion cause extraction, which uses a narrative-aware causal association (NCA) unit and a result-aware emotion. This model outperforms the advanced RHNN model by 6.6% in terms of $F1$ score but performs significantly worse overall. The hierarchical inter-clause interaction network (HICIN) framework is used [17] to model inter-clause interaction. On Chinese and English datasets, the suggested model outperforms the RTHN model by 0.81% and 0.65%, respectively. However, the system's usefulness is limited, as the emotion must first be explicitly marked before the cause can be derived.

2.3 Both Co-reference and Semantic Features

Wang et al. [18] published Multimodal Emotion Cause Pair Extraction in Conversation, which uses the graph attention network and the pre-training model BERT to gather clause semantic data. Bi-LSTM-based emotion and cause extractors are used [19] to extract useful properties from semantic data, and the model performs 3.4% better in the $F1$ score than other models. Future sentiment category predictions based on MC-ECPE would likely be more difficult due to the task's poor performance. The DCN-ECPE model is [20] an optimum full model for emotion-cause pair extraction, which builds dual-channel representations of emotion-cause pairings to take into account contextualized interactions of the sentence pairs and probable causal elements. Xu et al. [21] proposed an event extraction framework with an event representation strategy based on a tree structure for detecting the emotional causes of occurrences. Evaluations show that the suggested strategy outperforms the other approaches by 7.04%. SINA City News is the most popular corpus in the reviewed research [22] for emotion cause extraction. Labeled datasets are significantly deficient in the study, particularly for languages apart from Chinese. A novel method based on learning to rank has been developed [23] to identify emotion causes from the perspective of information retrieval. The suggested strategy performs significantly better in terms of precision, recall, and $F1$ measure than the baseline methods. The most highly rated clause for each emotion is then utilized as the explanation, with the other emotions' various causes being disregarded. A Bi-LSTMCRF network based on attention was used [24] to record context information and semantic links between candidate sentence clauses and emotion expression. It was successful in the trial, with an $F1$ score of 88.40% emotion expression.

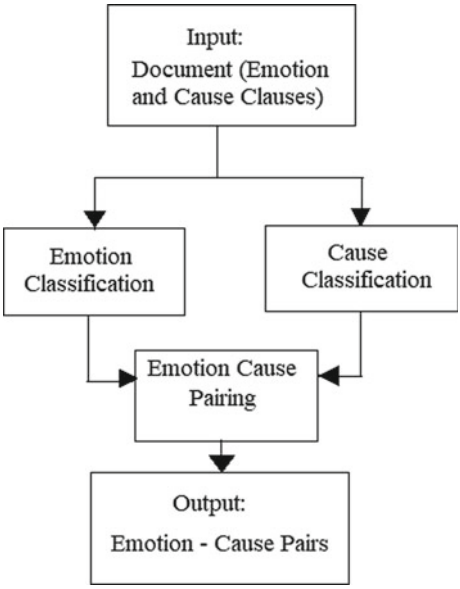


Fig. 2 Block diagram of two-step method

3 System Design

We have created a two-step strategy to implement this ECPE task.

Step 1: Emotion and Cause Classification

We have divided the ECPE task into two subtasks: emotion extraction and cause extraction. This process is divided into a collection of emotional clauses and a collection of causal clauses for each document.

Step 2: Pairing of Emotion Cause

In the second step, Emotion set E and Cause set C are then paired using a non-EUI join method. This provides us with a list of all possible emotion cause pairs and filters the pairs based on specified conditions (Fig. 2).

4 Methodology

For a given document D with n clauses, this task is defined as follows: if the document $D = (C_1, C_2, \ldots, C_n)$ then a clause is a group of words, where $Ci = (Wi_1, Wi_2, \ldots, Wi_n)$, then collection of emotion and cause pairs is $G = (\ldots, (Ce, Cc), \ldots)$ where Ce is emotion clause and Cc is cause clause. Emotion cause pair extraction task is addressed with a two-stage approach as described below:

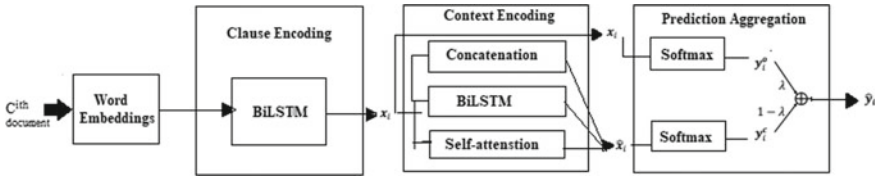


Fig. 3 Bi-LSTM model architecture

Step 1: Emotion and Cause Classification

Using an independent multi-task learning classifier, each clause in document D is categorized as either an emotion clause or a cause clause. The lowest layer captures the word-clause document structure using a Bi-LSTM network with two layers. A bidirectional LSTM is used to find the hidden state of the jth word in the ith clause $h_{i,j}$. Next, the attention method is used to obtain S^i, a clause representation $S^i = [S_1, S_2, \ldots, S_n]$. We shall observe that the upper layer has two components: cause extraction and emotion extraction. The independent clause representations are obtained at the bottom layer, which are the outputs of the lowest layers. Figure 3 shows the architecture of Bi-LSTM model. The context-aware representation of the word C_i is made up of the two hidden states of the two-component Bi-LSTM, $h^i{}_{\mathrm{e}}$ and $h^i{}_{\mathrm{c}}$. The softmax layer, which forecasts emotions and causes, then receives these hidden states as input. As a result, the loss of the model can be determined as the weighted sum of two components. The word2vec vectors are adjusted during training. We have a group of emotion clauses

$$\mathrm{EG} = \left(C_1^{\mathrm{e}}, C_2^{\mathrm{e}}, \ldots, C_i^{\mathrm{e}}\right) \tag{1}$$

and a group of cause clauses at the conclusion of this step

$$\mathrm{CG} = \left(C_1^{\mathrm{c}}, C_2^{\mathrm{c}}, \ldots, C_j^{\mathrm{c}}\right) \tag{2}$$

Step 2: Pairing of Emotion Cause

In the second stage, we apply a non-equi-join method to EG and CG groups which are obtained in the previous stage. A non-equi-join is a join with conditions other than exact match. First, we create a data frame for EG and CG groups with corresponding column name which needs to be joined, then define a non_equi_inner_join function, and then join two data frames by arbitrary conditions and optionally, by exact match of some columns.

$$\begin{aligned}&\text{non\_equi\_inner\_join(EG, CG, non\_equi\_conditions}\\&\quad = [\text{`emotion\_clauses'} = \text{`cause\_clauses'}], \text{equi\_on} = \text{`emotion\_seeds'})\end{aligned} \tag{3}$$

Equi-join keys are used to partition the matching number of first keys, followed by equi_join + filtering for each partition and concatenation of the results. The default setting uses all equi-join keys for partitioning, but this may cause slower execution and less memory use for large datasets.

$$\begin{aligned} \text{result\_joined} = \text{pd.merge(EG\_lef,CG\_right,left\_on} \\ = \text{equi\_left\_on,right\_on} = \text{equi\_right\_on)} \end{aligned} \tag{4}$$

Algorithm for Two-Step Approach

Input: English Emotion Cause Document

Output: Emotion and Cause Pairs

Begin

- Importing Dataset and Initializing Variables

```
Emotion_Seeds = pickle.load(F)
Document_Clauses = []
Emotion_Labels = []
Cause_Labels = []
Emotion_Cause_Pairs = []
Emotion_Cause_Pairs_Label = []
```

- Dataset Pre-processing

```
Clauses = re.split("[.,!;:\"]+", line)
Cleaned_Clause = remove_punctuation_from_clause
clause_words = cleaned_clause.split()
```

- Emotion and Cause Classification

```
emotion_probs = self.emotion_model(lower_output)
cause_probs = self.cause_model(lower_output)
batch_clauses = train_clauses[start_index:end_index]
batch_emotion_labels=train_emotion_labels[start_index:end_
index]
```

- Emotion and Cause Pairing

```
EG=pd.DataFrame([[1,2]],columns=['emotion_seeds','emotion_
clauses'])
CG=pd.DataFrame([[1,2]],columns=['emotion_seeds','cause_
clauses'])
non_equi_inner_join(EG,CG,non_equi_conditions=['emotion
_clauses'='cause_clauses'],equi_on='emotion_ seeds')
result_joined= pd.merge( EG_left, CG_right,left_on=equi
_left_on, right_on=equi_right_on)
```

- Bi-LSTM Model Training

```
Bilstm_Hidden_Units = 300
s = get_s(inputs, name='cause_word_encode')
w1= get_weight_varible('word_att_w1' + name, [sh2,sh2])
w2= get_weight_varible('word_att_w2' + name, [sh2, 1])
s = att_var(inputs,sen_len,w1,b1,w2)
s = tf.reshape(s, [-1, FLAGS.max_doc_len, 2 * FLAGS.n_hidden])
```

End

5 Experiments

5.1 Dataset Description

FrameNet's emotion-driven framework is used for emotion causal pair extraction due to the lack of English datasets. XML tags and sentiment roots are used to represent various emotions in the dataset. The collection includes representations of the feelings (Table 1).

The task will be evaluated using precision, recall, and $F1$ score.

- **Precision**: Precision is defined as the ratio of correct pairs to proposed pairs. Precision can be calculated as follows:

$$\text{Precision} = \text{TP}/\text{TP} + \text{FP} \tag{5}$$

- **Recall**: The recall ratio is the number of correct pairs divided by the number of annotated pairs. Recall can be calculated as follows:

$$\text{Recall} = \text{TP}/\text{TP} + \text{FN} \tag{6}$$

- ***F*1 score**: $F1$ score is a harmonic mean of precision and recall. $F1$ score can be calculated as follows:

$$F1 = 2 \times \text{TP}/\text{TP} + \text{FP} + \text{FN} \tag{7}$$

Table 1 Count of different emotions in the dataset

Emotion	Happy	Sad	Surprise	Disgust	Anger	Fear	Shame
Total	211	107	53	38	199	144	68

Table 2 Experimental results for emotion and cause classification

Classification type	Emotion			Cause		
Method	*P*	*R*	*F*1	*P*	*R*	*F*1
W2V + LR	0.8954	0.8942	0.8946	0.8954	0.8942	0.8946
W2V + SVM	0.9118	0.9112	0.9130	0.9118	0.9112	0.9099
TF-IDF + SVM	0.9298	0.9292	0.9279	0.9298	0.9292	0.9279
Bi-LSTM (proposed model)	0.9510	0.9489	0.9489	0.9527	0.9528	0.9519
Bi-LSTM + attention	0.9609	0.9614	0.9609	0.9640	0.9645	0.9639

Table 3 Experimental results for emotion and cause pairing

Method	*P*	*R*	*F*1
W2V + LR	1.0	0.3236	0.4890
W2V + SVM	1.0	0.5491	0.7089
TF-IDF + SVM	1.0	0.1734	0.3409
Bi-LSTM (proposed model)	1.0	0.6184	0.7642
Bi-LSTM + attention	1.0	0.5953	0.7463

6 Results and Discussions

The complete dataset is split into training data (70%), validation data (10%), and test data (20%) to ensure variance. A skip-gram model was trained on the dataset's texts with 70 iterations to classify emotions and causes. Word2vec vectors with a 300-dimensional pre-trained range are used in the embedding layer for deep learning models. The Bi-LSTM model was trained and tested using train–test splits of the dataset, and the outcomes are reported. The results of Step 1 and Step 2 are displayed in Tables 2 and 3.

Table 2 shows that the Bi-LSTM model is best for emotion and cause classification using pre-trained word2vec vectors. Attention in combination can slightly improve the *F*1 score. Our proposed model provides the best performance compared to the other models.

Table 3 shows that the TF-IDF representation produces the worst results with an *F*1 score for emotion-cause pairing. By observing the results, we are concluding that the Bi-LSTM model produces the best results as compared to the other models.

7 Conclusion

In this work, we used a two-step process to create an emotion-cause pair extraction task. When compared to other models, the Bi-LSTM model and Bi-LSTM + attention offer the best results. Instead of a Cartesian product, we used the non-equi-join

function for emotion-cause pairing, which can remove unrelated emotion-cause pairs without training a separate filter. In the future, to finish this task in a single step, additional effort must be made.

References

1. Xia R, Ding Z (2019) Emotion-cause pair extraction: a new task to emotion analysis in texts. In: Proceedings of the 57th annual meeting of the association for computational linguistics, Florence, Italy. Association for Computational Linguistics, pp 1003–1012
2. Gui L, Xu R, Lu Q, Wu D, Zhou Y (2016) Emotion cause extraction, a challenging task with corpus construction. In: Li Y, Xiang G, Lin H, Wang M (eds) Social media processing. SMP 2016. Communications in computer and information science, vol 669. Springer, Singapore. https://doi.org/10.1007/978-981-10-2993-6_8
3. Yu J, Liu W, He Y, Zhang C (2021) A mutually auxiliary multitask model with self-distillation for emotion-cause pair extraction. IEEE Access 9:26811–26821
4. Qiu B, Shang L (2022) Learning emotion-aware contextual representations for emotion-cause pair extraction. In: Lu W, Huang S, Hong Y, Zhou X (eds) Natural language processing and Chinese computing. NLPCC 2022. Lecture notes in computer science, vol 13551. Springer, Cham. https://doi.org/10.1007/978-3-031-17120-8_45
5. Li M, Zhao H, Su H et al (2021) Emotion-cause span extraction: a new task to emotion cause identification in texts. Appl Intell 51:7109–7121. https://doi.org/10.1007/s10489-021-02188-7
6. Bi H, Liu P (2020) ECSP: a new task for emotion-cause span-pair extraction and classification. arXiv:2003.03507
7. Song H, Zhang C, Li Q, Song D (2020) End-to-end emotion-cause pair extraction via learning to link
8. Gui L, Yuan L, Xu R, Liu B, Lu Q, Zhou Y (2014) Emotion cause detection with linguistic construction in Chinese Weibo text. In: Natural language processing and Chinese computing
9. Cheng Z, Jiang Z, Yin Y, Li N, Gu Q (2021) A unified target-oriented sequence-to-sequence model for emotion-cause pair extraction. IEEE/ACM Trans Audio Speech Lang Process 29:2779–2791. https://doi.org/10.1109/TASLP.2021.3102194
10. Diao Y, Lin H, Yang L et al (2020) Multi-granularity bidirectional attention stream machine comprehension method for emotion cause extraction. Neural Comput Appl 32:8401–8413. https://doi.org/10.1007/s00521-019-04308-4
11. Xiao X, Wang L, Kong Q, Mao W (2020) Social emotion cause extraction from online texts. In: 2020 IEEE international conference on intelligence and security informatics (ISI), Arlington, VA, pp 1–6. https://doi.org/10.1109/ISI49825.2020.9280532
12. Cheng Z, Jiang Z, Yin Y, Wang C, Ge S, Gu Q (2023) A consistent dual-MRC framework for emotion-cause pair extraction. ACM Trans Inf Syst 41(4):1–27. Article 105. https://doi.org/10.1145/3558548
13. Huang W, Yang Y, Peng Z, Xiong L, Huang X (2022) Deep neural networks based on span association prediction for emotion-cause pair extraction. Sensors (Basel) 22(10):3637. https://doi.org/10.3390/s22103637. PMID: 35632043; PMCID: PMC9146116
14. Bao Y, Ma Q, Wei L, Zhou W, Hu S (2022) Multi-granularity semantic aware graph model for reducing position bias in emotion cause pair extraction. In: Findings of the association for computational linguistics: ACL 2022, Dublin, Ireland. Association for Computational Linguistics, pp 1203–1213
15. Sun Q, Yin Y, Yu H (2021) A dual-questioning attention network for emotion-cause pair extraction with context awareness. In: 2021 international joint conference on neural networks (IJCNN), Shenzhen, China, pp 1–8. https://doi.org/10.1109/IJCNN52387.2021.9533767

16. Cao W et al (2022) Causal narrative comprehension: a new perspective for emotion cause extraction. IEEE Trans Affect Comput 13(4):1743–1758. https://doi.org/10.1109/TAFFC.2022.3206960
17. Lin P, Yang M, Gu Y (2021) A hierarchical inter-clause interaction network for emotion cause extraction. In: 2021 international joint conference on neural networks (IJCNN), Shenzhen, China, pp 1–8. https://doi.org/10.1109/IJCNN52387.2021.9534291
18. Wang F, Ding Z, Xia R, Li Z, Yu J (2022) Multimodal emotion-cause pair extraction in conversations. IEEE Trans Affect Comput. https://doi.org/10.1109/TAFFC.2022.3226559
19. Fan C, Li D, Wu Y (2022) Combining BERT with Bi-LSTM for emotion-cause pair extraction. In: 2022 4th international conference on computer communication and the internet (ICCCI), Chiba, Japan, pp 1–6. https://doi.org/10.1109/ICCCI55554.2022.9850274
20. Qie P, Shuang K (2022) DCN-ECPE: dual-channel network for emotion-cause pair extraction. In: 2022 10th international conference on intelligent computing and wireless optical communications (ICWOC), Chongqing, China, pp 87–91. https://doi.org/10.1109/ICWOC55996.2022.9809891
21. Xu R, Hu J, Lu Q, Wu D, Gui L (2017) An ensemble approach for emotion cause detection with event extraction and multi-kernel SVMs. Tsinghua Sci Technol 22(6):646–659. https://doi.org/10.23919/TST.2017.8195347
22. Khunteta A, Singh P (2021) Emotion cause extraction—a review of various methods and corpora. In: 2021 2nd international conference on secure cyber computing and communications (ICSCCC), Jalandhar, India, pp 314–319. https://doi.org/10.1109/ICSCCC51823.2021.9478079
23. Xu B, Lin H, Lin Y, Diao Y, Yang L, Xu K (2019) Extracting emotion causes using learning to rank methods from an information retrieval perspective. IEEE Access 7:15573–15583. https://doi.org/10.1109/ACCESS.2019.2894701
24. Liang L, Ji X, Ren F (2020) Attention-based bi-LSTM-CRF network for emotion cause extraction in texts. In: 2020 IEEE international conference on mechatronics and automation (ICMA), Beijing, China, pp 1670–1675. https://doi.org/10.1109/ICMA49215.2020.9233688

Automatic Room Light Control

Shital Dongare, Swapnil Gawali, Prasad Rathod, Shivprasad Kavathe, and Amit Dolas

Abstract Due to the limited nature of natural resources and the effects of energy usage on the environment, energy conservation has gained importance in recent years. Lighting is a major factor in how much energy is used in buildings. Hence, it is essential to put policies in place that can cut down on energy used for unneeded lighting. Using an automatic light turn-off system that can recognize when there are people present or not in a room and turn the lights out appropriately is one such measure. The system is made to be both economical and energy-efficient, and it is simple to integrate into already installed lighting systems. This initiative has the potential to lower energy use and enhance sustainability in both residential and commercial structures. We implemented model and noted observations for 7 days. As an experimental result, we could save 283 w/h electricity in 7 days, which was nearly about 26% of total energy consumption.

Keywords Automatic light turn on–off · Ultrasonic sensor · Presence detection

1 Introduction

This project aims to develop an economical and energy-efficient system for automatic light switching using an ultrasonic sensor. By detecting occupancy in a space, the system automatically turns off the lights when no one is present. This approach has the potential to reduce energy usage in both residential and commercial buildings, promoting sustainable energy practices.

S. Dongare (✉) · S. Gawali · P. Rathod · S. Kavathe · A. Dolas
Vishwakarma Institute of Technology, Pune, India
e-mail: shital.dongare@vit.edu

S. Gawali
e-mail: swapnil.gawali22@vit.edu

P. Rathod
e-mail: prasad.rathod22@vit.edu

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_42

To understand the current landscape, we will review previous research on automatic light turn-off systems, including sensor types, technologies, and occupancy detection methods. We will explore the advantages and disadvantages of these systems, considering factors like user acceptance, cost-effectiveness, and energy efficiency.

Next, we will describe the process of creating and integrating our system, highlighting the hardware components and their connections. The software code developed to control the microcontroller will also be discussed. Testing strategies will be detailed, assessing the system's precision and efficiency under different lighting conditions, distances, and occupancy scenarios.

The data analysis and results' section will present the outcomes of our testing, comparing the performance of our system with existing systems. In the implications and recommendations section, we will discuss the benefits of implementing our system, including reduced energy consumption, enhanced sustainability, and potential cost savings. We will also suggest future studies to further improve the precision and efficiency of automatic light turn-off systems.

2 Literature Review

Building energy usage can be decreased by installing automatic light turn-off systems. These systems have employed infrared, ultrasonic, and microwave sensors, among other types of sensors, to detect occupancy. Ultrasonic sensors are favored among them because they can detect motion and occupancy regardless of lighting, making them useful in a variety of contexts.

A business building has an automatic light turn-off system using ultrasonic sensors installed as part of a study by Lee et al. [1]. Accurate occupancy detection by the system led to significant energy savings. The study did discover, however, that the position and orientation of the sensors, the height of the ceiling, and the presence of items that could obstruct the ultrasonic signals all had an impact on the system's performance [1].

De Almeida et al. [2] conducted another investigation to examine the efficacy of an automatic light turn-off system that included the usage of infrared and ultrasonic sensors. The technology proved effective in detecting occupancy and lowering energy use, according to the study, with an average energy savings of 23%. The study did, however, point up several drawbacks, such as false triggers brought on by the presence of objects and passers-by outside the room [2].

An automatic light turn-off system was created utilizing a mix of infrared and ultrasonic sensors in a study by Kim et al. [3]. The technology correctly and successfully reduced energy consumption by detecting occupancy. The study also discovered that the placement and calibration of the sensors had an impact on the performance of the system [3].

In this study, a passive infrared (PIR) and ultrasonic sensor combination was used to design and construct an autonomous lighting management system. The

system's accuracy in detecting human presence and motion allowed it to alter lighting levels accordingly, saving a lot of energy in the process. The study discovered that while ultrasonic sensors were better at detecting motion, PIR sensors were better at detecting human presence. A more reliable and accurate system was produced by combining the usage of the two sensors [4].

Zhang et al. studied an infrared, ultrasonic, and image sensor-based intelligent lighting control system which was created. The system's accuracy in detecting human presence and motion allowed it to alter lighting levels accordingly, saving a lot of energy in the process. According to the study, using a hybrid occupancy sensor—a device that integrates numerous sensors—provided a more reliable and precise system, enabling the accurate identification of human presence and mobility independent of the surrounding circumstances. While the infrared and ultrasonic sensors were better at detecting objects, the image sensor was particularly good at detecting people in the room [5].

Meena et al., and team Author Created GSM-based home security system. Each object trespassing within the motion sensor's range is detected by this motion-based security system. When an incursion is discovered, a message alert will be provided to the user depending on the mobility of the object. If the PIR motion sensor detects movement, this GSM-based home security system has the option to send an SMS warning. A wonderful safety feature is provided to the user by this SMS sending capability [6].

Automatic light turn-off systems effectively reduce building energy usage. Ultrasonic sensors are recommended due to their ability to detect motion and occupancy in any lighting. However, factors like sensor placement, signal obstruction, and false triggers can impact system performance. Further research is needed to enhance precision and efficiency.

3 Methodology

3.1 Architecture

See Fig. 1.

3.2 Proposed System

The proposed system for the "Automatic Room Light Control" project consists of the following components (Fig. 2):

(1) Ultrasonic sensors: Two ultrasonic sensors are placed at the entryway of the room to detect the presence of people. These sensors send and receive ultrasonic

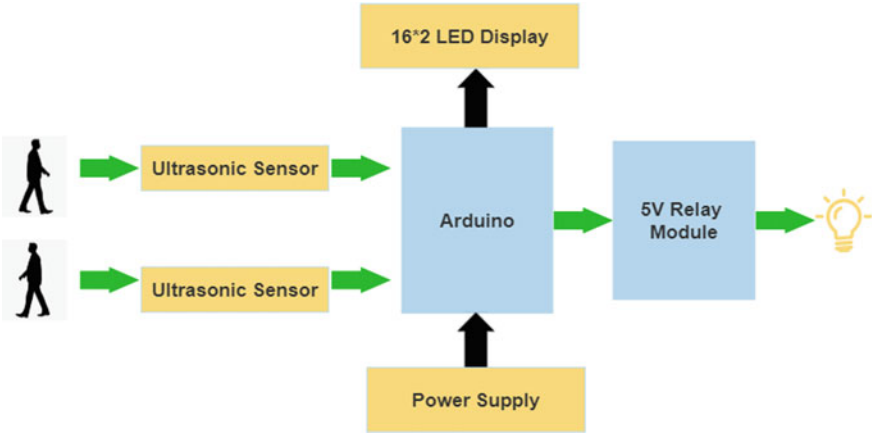


Fig. 1 Architecture of automation room light control

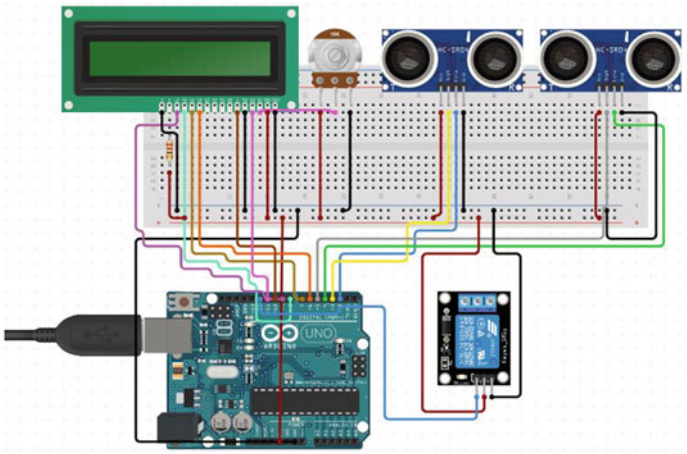


Fig. 2 Connection diagram

waves to determine the distance from any object in their path, allowing them to detect the presence of people.

(2) Arduino microcontroller: The Arduino microcontroller is the brain of the system, responsible for processing the sensor data and controlling the light. The microcontroller reads the sensor data and determines if a person has entered or left the room. Based on this data, it controls the relay module to turn the light on or off.

(3) Relay module: The relay module is used to control the light in the room. The microcontroller sends a signal to the relay module to turn the light on or off based on the sensor data.

(4) LCD screen: An LCD screen is used to display the number of people in the room. The screen displays the information in real-time, allowing users to see the occupancy status of the room briefly.

The proposed system works by using the ultrasonic sensors to detect the presence of people in the room. When a person enters the room, the sensor data is sent to the microcontroller, which turns on the light using the relay module. When a person leaves the room, the sensor data is again sent to the microcontroller, which turns off the light automatically after a short delay.

The proposed system works by using the ultrasonic sensors to detect the presence of people in the room. When a person enters the room, the sensor data is sent to the microcontroller, which turns on the light using the relay module. When a person leaves the room, the sensor data is again sent to the microcontroller, which turns off the light automatically after a short delay.

The system also displays the number of people in the room on the LCD screen, providing real-time information on the occupancy status of the room. This feature can be useful in managing the use of the room and making informed decisions on the use of energy resources.

Overall, the proposed system provides a user-friendly and automated solution to control the lighting of a room, help to conserve energy, and provide convenience to users.

4 Result and Discussion

The "Automatic Room Light Control" project provides a practical and efficient solution to controlling lights in a room based on occupancy. The project detects the presence of people using two ultrasonic sensors and updates the light accordingly. The system turns the light on when someone enters the room and turns it off when they leave, thus saving energy (Figs. 3, 4, 5, and 6).

The project has a simple circuit with two ultrasonic sensors, an LCD screen, and a relay to control the light. The program code uses the data from the ultrasonic sensors to detect the presence of people in the room and updates the person count accordingly. The program then controls the light based on the number of people detected.

The project has the potential to save energy and reduce energy costs by ensuring that lights are only turned on when there are people in the room. It also provides a convenient way to automatically control the lighting without the need for manual intervention.

Fig. 3 Output 1

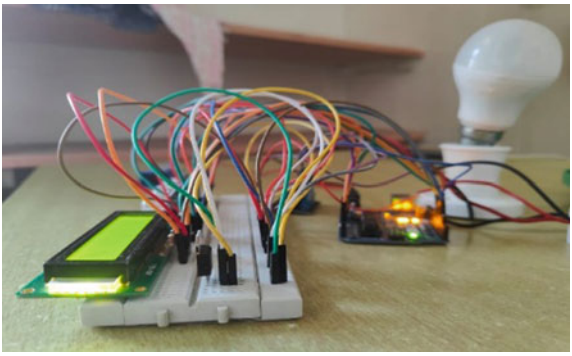


Fig. 4 Output 2

Fig. 5 Output 3

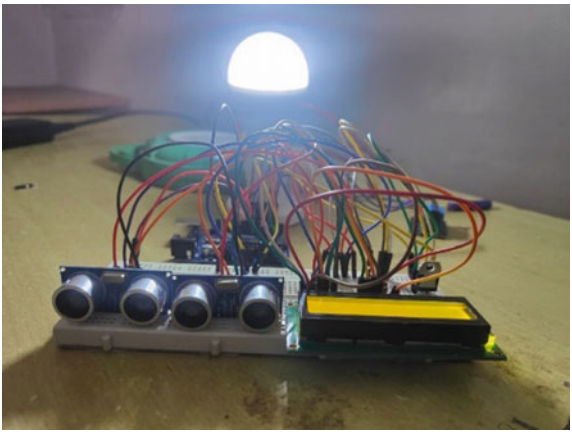


Fig. 6 Output 4

Overall, the "Automatic Room Light Control" project is an excellent example of using technology to create a practical and efficient solution to a real-world problem. It can be implemented in homes, offices, and other places where energy savings and convenience are required.

4.1 Test Cases

4.1.1 Introduction

The Automatic Room Light Control System is designed to detect the presence of people in a room and control the lighting accordingly. This system aims to improve energy efficiency and provide a more convenient user experience. The purpose of this test case report is to document the testing process and results of the Automatic Room Light Control System.

4.1.2 Test Case 1: Functional Testing

Objective: To verify that the system can detect the presence of a person in the room and turn on the lights automatically.

Steps:

- Enter the room and wait for the system to detect your presence.
- Observe if the lights turn on automatically.
- Exit the room and wait for the system to turn off the lights automatically.
- Result: The system was able to detect the presence of a person in the room and turn on/off the lights automatically.

4.1.3 Test Case 2: Usability Testing

Objective: To evaluate the user interface of the system and its response time.

Steps:

- Observe the user interface of the system.
- Enter the room and observe how quickly the system turns on the lights.
- Try to manually turn the lights on/off using the system.
- Result: The user interface of the system was easy to use and understand. The system's response time was fast enough to turn on the lights before the user enters the room. The manual override functionality worked correctly.

4.1.4 Test Case 3: Performance Testing

Objective: To verify the reliability and energy efficiency of the system.

Steps:

- Test the system under different environmental conditions (e.g., different lighting levels, temperature, and humidity).
- Test the system's ability to handle multiple users.

- Monitor the energy consumption of the system.
- Result: The system was able to control the lights reliably and consistently under different environmental conditions. The system was able to handle multiple users and adjust the lighting accordingly. The energy consumption of the system was minimal, and it did not waste energy by keeping the lights on unnecessarily.

4.1.5 Test Case 4: Security Testing

Objective: To verify the security of the system.

Steps:

- Attempt to access the system without proper authorization.
- Attempt to tamper with the system.
- Monitor the system for any unauthorized data collection or transmission.
- Result: The system was secure and could not be accessed or controlled by unauthorized users. The system was able to detect and prevent tampering or hacking attempts. The system did not collect or transmit any sensitive user data without their consent.

4.1.6 Result

After creating model, we experimentally tested it in the room for 7 days. Here, we used Philips Steller bright led bulb which consumes 20 w/h electricity. After the 7 days of observation, we got very good results. We were able to save energy and billing of electricity also. After implementation, we observed following statistics. Electricity consumption without using module for 8 h electricity consumption in one day = 20*8 = 160 w/h.

$$\text{Electricity consumption in seven days} = 20 * 8 * 7 = 1120\,\text{w/h}.$$

Here, we observed that though switch is on for 8 h still, we find that bulb was on for 5/6 h according to presence or absence of person in the room. Observation of 7 days has been given in Table 1.

5 Drawback and Limitations

While our proposed automatic light turn-off system utilizing ultrasonic sensors presents numerous benefits, there are some limitations and drawbacks that need to be acknowledged. These include:

Table 1 Observation of 7 days

Days	Hours	Power consumption (20 w/h)
1	5.20	105.6
2	5.00	100
3	6.32	129.36
4	5.47	114.51
5	5.50	115.5
6	7.00	138.6
7	6.45	133.65
Total		837.22

Total energy saving for 7 days = 1120–837.22 = 282.78w/h.
Percentage of energy saved for 7 days = (282.78/1120) × 100 = 25.24%.

- False triggers: False triggers can occur when the ultrasonic sensor mistakenly detects motion or occupancy, leading to unnecessary light activation or deactivation. This can be caused by factors such as sudden air movements or vibrations, resulting in inaccurate readings. To address this, advanced algorithms need to be developed to enhance the system's ability to differentiate between human presence and other sources of movement.
- Signal obstruction: Objects or obstacles in the environment can obstruct the ultrasonic signals, leading to inaccurate motion detection. This can occur when there are furniture, partitions, or other physical barriers present in the monitored space. Employing multiple sensors and adjusting their positions can help to mitigate signal obstruction and enhance the system's accuracy.
- User acceptance: The success of any automated system relies on user acceptance and satisfaction. It is essential to consider the user's comfort and convenience, as well as their perception of the system's reliability and effectiveness. Factors such as ease of use, intuitive interface, and user feedback should be considered during the system's design and implementation.

6 Overcoming Drawback

To address the limitations and drawbacks of our proposed automatic light turn-off system, the following strategies can be implemented:

- False triggers: Implement advanced signal processing algorithms that can distinguish between human presence and other sources of motion accurately. By analyzing the patterns and characteristics of the detected signals, the system can reduce false triggers and improve overall accuracy.

- Signal Obstruction: Employ multiple sensors and adjust their positions to ensure better signal reception and minimize signal obstruction caused by objects or barriers in the environment. This can involve placing sensors at different heights or using reflective surfaces to redirect signals and enhance detection.
- User acceptance: Prioritize user-centric design by incorporating user feedback during the system development phase. Provide intuitive interfaces, clear instructions, and customizable settings to enhance user experience and increase user acceptance of the technology.

By implementing these strategies, we can overcome the drawbacks associated with the automatic light turn-off system, improving its accuracy, efficiency, and user satisfaction. Continued research and development efforts in these areas will further enhance the system's performance and contribute to its successful implementation in various residential and commercial settings.

7 Comparative Analysis

To evaluate the distinctiveness of our proposed automatic light turn-off system, a comparative analysis with existing frameworks can be conducted. This analysis aims to highlight the advantages and uniqueness of our approach compared to traditional systems.

Existing frameworks typically rely on passive infrared (PIR) sensors for occupancy detection, which have limitations in accurately differentiating between human presence and other heat sources. In contrast, our system utilizes ultrasonic sensors, which can detect motion and occupancy regardless of lighting conditions, making it more reliable and versatile.

Moreover, PIR-based systems often suffer from false triggers due to changes in ambient temperature or small animals moving within the detection range. Our proposed system, with advanced signal processing algorithms, effectively mitigates false triggers and improves overall precision.

Additionally, our system emphasizes energy efficiency by integrating a microcontroller that enables automatic light switching based on occupancy detection. This feature distinguishes our approach from conventional systems that require manual intervention or fixed time schedules for light control.

By comparing our system's performance with existing frameworks, we can demonstrate its superior accuracy, adaptability to various environments, and energy-saving potential. This analysis reinforces the distinctive features and advantages of our proposed automatic light turn-off system, positioning it as a valuable contribution to the field of energy-efficient building management.

8 Conclusion

An effective and dependable method for automatic light management, in conclusion, is the employment of an ultrasonic sensor. The sensor determines whether there are people or objects in the space and turns on or off the lights in response, saving energy and enhancing user convenience. The system is simple to set up and can be altered to suit various lighting conditions and environments. Overall, an ultrasonic sensor-based automatic light management system is a useful and efficient way to improve indoor comfort and sustainability.

References

1. Lee, Kim, Park (2017) Published paper development of an ultrasonic sensor-based automatic light-turn-off system for energy saving in business buildings published in the Journal of Sensors
2. De Almeida et al (2016) Investigation to examine the efficacy of an automatic light turn-off system
3. Kim SH, Jung HJ, Lee JH, Kang JH (2018) Development of an automatic light turn-off system utilizing a combination of infrared and ultrasonic sensors
4. Bhatia PS, Rana PS, Patheja PS (2015) An automated lighting control system utilizing passive infrared and ultrasonic sensors
5. Zhang H, Shen L, Wang X (2020) Hybrid occupancy sensors are used in development and performance assessment of an intelligent lighting control system
6. Meena NH, Parashar N, Bhephalawat N, Jain P, Prayag, IV B. Tech, Department of EE, SKIT Jaipur, India paper on automatic light control and home security using PIR sensor, GSM and Arduino system

EthFor: Forensic Investigation Framework for Ethereum Blockchain

Radhika Moondra, Vikas Sihag, and Gaurav Choudhary

Abstract The use of blockchain technology has become increasingly popular for protecting data storage and transmission through decentralised, peer-to-peer networks. This technology is essentially a synchronised decentralised database managed by Distributed Ledger Technology (DLT), which shifts trust from a centralised network to a completely decentralised network of nodes. As a result, information is securely preserved and transmitted across numerous nodes spread across wide areas, ensuring that transactions remain private, secure, and immutable. Blockchain offers various benefits, including traceability, accountability, transparency, and provenance of recorded transactions on the ledger. Therefore, this study proposes a forensic investigation framework for Ethereum blockchain.

Keywords Blockchain · Ethereum · Digital forensics

1 Introduction

Blockchain is a decentralised digital ledger technology that allows transactions to be recorded and verified without the need for intermediaries. The use of blockchain technology and cryptocurrency has grown significantly in recent years, with various industries exploring its potential uses beyond finance. According to a report by Deloitte, more than 80% of executives view blockchain as a strategic priority for their organisations. As of January 2022, the total market capitalization of all cryptocurrencies exceeded $2.5 trillion USD. The number of unique cryptocurrency users is

R. Moondra (✉) · V. Sihag
Department of Science, Technology and Forensics, Sardar Patel University of Police, Security and Criminal Justice, Jodhpur, India
e-mail: mtcs20rm@policeuniversity.ac.in

V. Sihag
e-mail: vikas.sihag@policeuniversity.ac.in

G. Choudhary
DTU Compute, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
e-mail: gauch@dtu.dk

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_43

estimated to be between 101 million and 200 million globally, as per a study by the Cambridge Centre for Alternative Finance.
Cryptocurrency transactions are recorded on the blockchain, making them secure, transparent, and irreversible [1]. Cryptocurrencies use cryptography to ensure the authenticity and security of transactions [2]. Consistency, transparency, equal rights, immutability, and availability are qualities that are fundamental to blockchain architecture and design [3].
In this paper, we present EthFor, a digital forensic framework for Ethereum transactions to locate artefacts regarding smart contracts and transactions for investigation purposes. *Structure of the paper*: Sect. 2 covers related work. Section 3 and 4 present proposed methodology and experimental results, respectively. Section 5 covers conclusion as well as future work.

2 Literature Review

Using technology and science to investigate and prove facts in criminal and civil courts of law is known as blockchain forensics, to be more specific. In blockchain forensics, latent evidence that transaction events have left on a blockchain digital ledger is recovered and examined [4].

In order to help out investigators with clear guidance during the investigative process, forensic systems have been provided [5] over the years. At the Digital Forensic Research Workshops (DFRWS) groups of academics and professionals listed seven phases or processes for performing digital forensics such as identification, preservation, collection, examination, analysis, presentation, and decision-making (Table 1).

3 Methodology

The forensics study of any data source involves some basic conditions. All these conditions vary based upon demands and data that is available. Throughout this section blockchain-based forensic framework will be explained. This section also specifies how the experiment was done and corresponding results were obtained as well.

3.1 The Procedure

Our system must be empowered with features like integrity, traceability, authentication, verifiability, and security in order to maintain the chain of custody and the trail of events. All forensic procedures must be verifiable in order for an investigation to be considered valid, and outcomes from evidence analysis must be repeatable. The blockchain-based forensic architecture that is detailed in the next section can help with the above mentioned needs.

Table 1 A brief of literature surveyed

Paper	Challenges addressed	Proposed solution	Artefacts
[6]	Heterogeneity and lack of transparency	Internet-of-Forensics (IoF)	Mobiles,laptops, smart devices, evice/Hardware MAC or identification, number, IP address registration, domain names, routers, switches, gateways
[7]	CoC, presevation of integrity	Security, data integrity, multi-party verification	PC, laptop, external hard drive, USB, mobile devices
[8]	Proof of exixtance, privacy preservation	IoT forensic chain (IoTFC)	Date, time stamps, transactional evidences file signature
[9]	Identity privacy concerns	BIFF Blockchain-based IoT Forensecis Framework	time stamp, identity of a device, geolocation, personal ID, UID of device
[10]	collection and preservation of DF evidence	Cyber-Trust blockchain (CTB) for CoC	Metadata of evidences, IP addresses, IoT devices, IP packets, device image
[11]	Roles, access control, privacy	Distributed CoC, health data privacy	Time stamps, IoMT devices, history, public key, SCL, IP and MAC address
[12]	Privacy, data security	Efficiency, accuracy, security and privacy of the data	Sensors, wearable devices
[13]	Privacy, approval and honesty	Auditability, interoperability, and accessibility of health data	Hash values, eHealth records
[14]	Data privacy mechanism for IoT models	Integrity, authenticity, non-repudiation and identity privacy	Time stamps, UserID, speed of the car, road, traffic jams
[15]	Feasibility of the communication and computing in the mobile device	HACIT 2	Vehicle ID, timestamps, GPS position, speed
[16]	Privacy and preservation of evence data	Block4Forensic (B4F) framework	Black box, EDR, Basic safety message, OBD ports, time stamps, vehicle ID,
[17]	Security, trust and privacy of data and processes in ITS	Deployment of blockchain in ITS	delays, location, travel time, and surveillance
[18]	centralised evidence collection, Log integrity, multi-party collusion	secure logging as-a-service for cloud environment, integrity, confidentiality	Signature, data hash, CSP, Network transaction, IP address, Time stamps, accessing time, virtual machine logs, deletion of file
[19]	Integrity, confidentiality of logs	BlockSLaaS-secure logging as-a-service for cloud environment	VM logs, network transactions,
[20]	Distributed storage of forensic data	Data encryption, distributes storage	hexadecimal contract number, farmer ID, signature, payment destination, metadata
[21]	Process provenance	Proof of existence, privacy preservation for records	Hash value, time stamps, public key,

4 Blockchain-Based Forensic Framework

4.1 Environment Setup

In this setup, the framework of the technique is shown in Fig. 1. At each level, different rules and policies will be used. The creation of the case is the first stage where a case can be opened by a prosecutor or an investigator. Then, the necessary forensic technologies are used to gather and analyse relevant evidence. To facilitate future searches and categorisations, the description of each activity (such as saving and analysing evidence) could come with a descriptive file attached in JSON format. It is presumed that safe and private storage is utilised to preserve the evidence like Inter-Planetary File System (IPFS). The evidence's hashes may, more precisely, store the SHA-256 hash.

All of the material may be collected and presented in court after the inquiry is complete. The interactions described above are documented in a smart contract and so preserved on the blockchain forever. The chain of custody is maintained and the immutability of the blockchain assures the verifiability of the investigation.

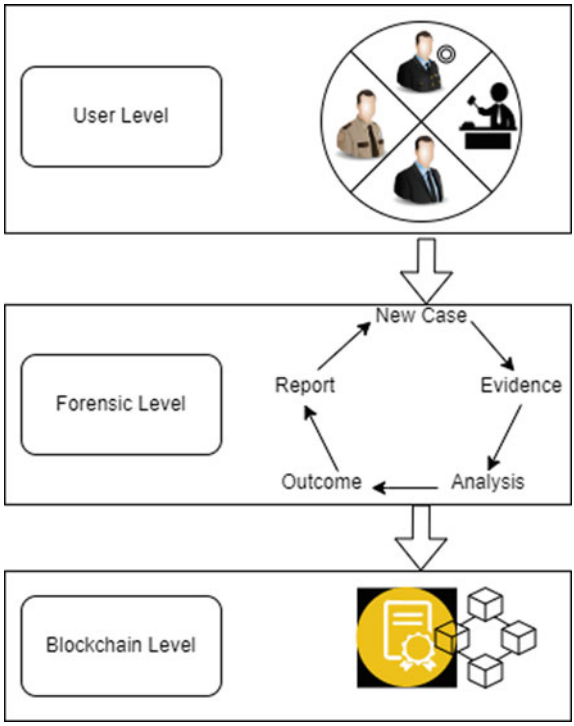


Fig. 1 Levels of proposed framework

Other blockchain-based identity management systems may be incorporated in this layer. For instance, the MetaMask, a chrome extension, is being used for testing settings to control the wallets and execute the smart contracts via the web interface.

4.2 Experimental and Integration Details

Relationships between the various framework participants and the forensic incidents have been built with the use of a smart contract, and a number of tests have been conducted on a local private blockchain to show the viability and efficacy of the suggested method. To be more explicit, truffle 5.6 was used to construct and deploy a fully functional smart contract, while Node 18.10 and Ganache-cli 7.5 were used to instal an Ethereum-based blockchain. Moreover, a graphical interface was developed to access and upload data stored in the blockchain using the Node Package Manager NPM 9.1. This interface also delivers the correct hash of the directory for a certain inquiry, along with information on how it connects to IPFS and other unique information.

Ethereum's dependable consensus system and smart contract architecture are the main reasons we picked it. As a result, the investigators will be able to evaluate the information pertaining to a specific inquiry in real time basis for management or verification goals using the get/retrieve functionalities incorporated into the smart contract.

Smart contracts may be compiled and deployed using Truffle, which leverages the Solidity compiler. Since doing so costs money and needs waiting for the deployment and execution of transactions to be processed, we may be unable to publish our continuing smart contract improvements to the public blockchain. We need some kind of testnet since anything we enter into the blockchain network is irreversible and cannot be deleted. At this point, Ganache and Ganache-CLI are quite helpful.

There will be 10 virtual accounts altogether, each with an initial balance of 100 ether. Because these virtual accounts will be required to purchase gas while processing transactions on the blockchain. Ten private keys comes with ten virtual accounts are included. The keys listed below are used to sign a transaction before it is added to the blockchain. The port that Ganache listens to is the next item. By default, that port's number is 8545. We receive an output from the Ganache terminal window 6 that is the transaction receipt when you develop and deploy a smart contract or conduct any transaction against Ganache [22] (Figs. 2 and 3).

It shows that there are 6 blocks mined in this account. Each with its mining time stamp and amount of Gas used is shown. Amount of Gas used, time stamp, Blockhash, Transaction hash, Sender account address, and so on. In Fig. 4, all the transactional detailed associated with Block 5 of account 0x20A9a9a316288Aa5f97eB1C1f85d2 847FD4982a0 are shown.

In addition to the advantages that smart contracts make possible, it is crucial to protect the confidentiality of the transactions and the participants. As a result, only individuals who fulfil certain duties can alter the contents.

```
Available Accounts
==================
(0) 0x03AA97f1a1A3dbD4D640a0cD3983298b371d8296 (100 ETH)
(1) 0x6442f743a2fF53635752daBE30FE4f4Cd69D1026 (100 ETH)
(2) 0x74Df25d89C662c7676997bb518604403d047553C (100 ETH)
(3) 0xd3AaDa9Da82c2582e187Ed18B13C1B4056b4D7D8 (100 ETH)
(4) 0x111050d9686D1568eE9CeEc15aB8fB3976f3aEd4 (100 ETH)
(5) 0xBB1EE14e0320619287645e6AdFd27C3302159986 (100 ETH)
(6) 0xEc15f7517044d873105a5067B2477D340BcFAb90 (100 ETH)
(7) 0x3673607Ed713eA29e2A2d2EbA9455e554D871B3F (100 ETH)
(8) 0x1290E1755c2a529b5f20adf3968f4Fa1798ae0de (100 ETH)
(9) 0x52c4394eE7a643DD479DC328d46C977419081a67 (100 ETH)

Private Keys
==================
(0) 0xdbefbb7277eae113f1e70f0e705e7814d4f98f0187d62ac5afc3c1ebbb02d51c
(1) 0x6bc3bced3eebae2151e4e21e32d5d6d50575ffefcd3b561afc3086976057ff38
(2) 0x129ed4a683e80e66e14dabb53ac4dcc213a7b89571b49f084f67218f290e2d9c
(3) 0xb2786d1bdbacb28adbef6584095d35e094ef90887d06edb6fa78b262143abd5a
(4) 0x4a5fa9714bbdace1a54c46afc804de4574d6a47ef19fb3f4ef450d546505a968
(5) 0x9440d0c84699dcde0b18f0953a605b540438d8eed3798d6fc8e488b2a90821fd
(6) 0x2faa588cf59e13275257b198771da7d14da699aa9e45f0d9e6e8177d78d9a08f
(7) 0x3ea5207ed3cc418d513677f21949d63ad96b1e858273fd557154379ee6071974
(8) 0x20b28994831a0779a17a9d62f6e3ef4dd7da50a501641ea3089c4499dcb6547c
(9) 0x0b33bcac17e907a44cf14211d5479b260d835bd1895cad6b822475464a599036

HD Wallet
==================
Mnemonic:      manual steak upgrade hero junk boost eight hub scissors predict area rescue
Base HD Path:  m/44'/60'/0'/0/{account_index}

Gas Price
==================
20000000000

Gas Limit
==================
46000000
```

Fig. 2 Available accounts on Ganache Testnet with private keys and mnemonic

5 Conclusion

Digital evidence is essential to cybercrime investigations because it is used to connect people to illegal activities. Because of this, it's crucial to guarantee the accuracy, reliability, and auditability of digital forensic evidence as it moves through several tiers of the chain of custody throughout a cybercrime investigation.

As they are latent, unpredictable, and fragile, they may quickly and easily cross jurisdictional boundaries, and they are often time- and machine-dependent, digital evidences present unique challenges. As a consequence, it may be difficult to ensure the legitimacy and legality of the processes and techniques used to gather and convey evidence in the digital age. For the forensic community, the ability of blockchain technology to provide a complete picture of events (transactions/actions) to their source offers enormous potential.

```
Compiling your contracts...
===========================
> Everything is up to date, there is nothing to compile.


Starting migrations...
======================
> Network name:    'development'
> Network id:      1670178979887
> Block gas limit: 46000000 (0x2bde780)


1_initial_migration.js
======================

   Deploying 'Migrations'
   ----------------------
   > transaction hash:    0xb8b3b3e41956eef65fa586c1fd52d7f171dce37cc5283d1103fda5ed2b25c717
   > Blocks: 0            Seconds: 0
   > contract address:    0x7f22A84CDE22e875D30Ccf0497099d821d614cca
   > block number:        1
   > block timestamp:     1670179026
   > account:             0x03AA97f1a1A3dbD4D640a0cD3983298b371d8296
   > balance:             99.99549526
   > gas used:            225237 (0x36fd5)
   > gas price:           20 gwei
   > value sent:          0 ETH
   > total cost:          0.00450474 ETH

   > Saving migration to chain.
   > Saving artifacts
   -------------------------------------
   > Total cost:          0.00450474 ETH


2_deploy_contracts.js
=====================

   Deploying 'Financial_forensics'
   -------------------------------
   > transaction hash:    0xaa8e7ece658b08544d41896e36160863cc86547b9de2962ee26dbc75c6a532f3
```

Fig. 3 Smart contract deployment

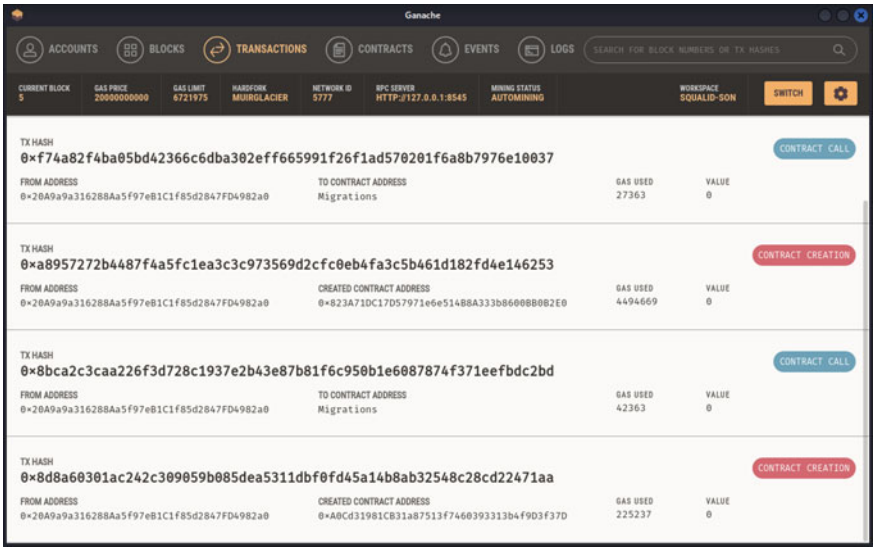


Fig. 4 Transactional details associated with account 0x20A9a9a316288Aa5f97eB1C1f85d2847FD4982a0

References

1. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. Decentralized Bus Rev 21260
2. Blockchain explained: what is blockchain? | Euromoney learning
3. Casino Fran, Dasaklis Thomas K, Patsakis Constantinos (2019) A systematic literature review of blockchain-based applications: current status, classification and open issues. Telematics Inform 36:55–81 March
4. Exploring Blockchain Forensics
5. Mas'ud MZ, Hassan A, Shah WM, Abdul-Latip SF, Ahmad R, Ariffin A, Yunos Z (2021) A review of digital forensics framework for blockchain in cryptocurrency technology. In: 2021 3rd international cyber resilience conference (CRC). IEEE, pp 1–6
6. Kumar G, Saha R, Lal C, Conti M (2021) Internet-of-forensic (IoF): a blockchain based digital forensics framework for IoT applications. Future Gener Comput Syst 120:13–25
7. Ryu JH, Sharma PK, Jo JH, Park JH (2019) A blockchain-based decentralized efficient investigation framework for IoT digital forensics. J Supercomput 75(8):4372–4387
8. Li Shancang, Qin Tao, Min Geyong (2019) Blockchain-based digital forensics investigation framework in the internet of things and social systems. IEEE Trans Comput Soc Syst 6(6):1433–1441
9. Le DP, Meng H, Su L, Yeo SL (2018) Biff: A blockchain-based IoT forensics framework with identity privacy. In: TENCON 2018-2018 IEEE region 10 conference. IEEE, pp 2372–2377
10. Brotsis S, Kolokotronis N, Limniotis K, Shiaeles S, Kavallieros D, Bellini E, Pavué C (2019) Blockchain solutions for forensic evidence preservation in iot environments. In: 2019 IEEE conference on network softwarization (NetSoft). IEEE, pp 110–114
11. Malamas V, Dasaklis T, Kotzanikolaou P, Burmester M, Katsikas S (2019) A forensics-by-design management framework for medical devices based on blockchain. In: 2019 IEEE world congress on services (SERVICES), vol 2642. IEEE, pp 35–40
12. Chakraborty S, Aich S, Kim HC (2019) A secure healthcare system design framework using blockchain technology. In: 2019 21st international conference on advanced communication technology (ICACT). IEEE, pp 260–264
13. Mousa Mohammed Khubrani (2021) A framework for blockchain-based smart health system. Turkish J Comput Math Edu (TURCOMAT) 12(9):2609–2614
14. Billard D, Bartolomei B (2019) Digital forensics and privacy-by-design: example in a blockchain-based dynamic navigation system. In: Annual privacy forum. Springer, pp 151–160
15. Kevin D, David B (2018) Hacit2: a privacy preserving, region based and blockchain application for dynamic navigation and forensics in VANET. In: International conference on ad hoc networks. Springer, pp 225–236
16. Cebe M, Erdin E, Akkaya K, Aksu H, Uluagac S (2018) Block4forensic: an integrated lightweight blockchain framework for forensics applications of connected vehicles. IEEE Commun Mag 56(10):50–57
17. Waseem M, Ahmed KA, Azeem MT (2021) Blockchain based intelligent transport system. Mod Innov Syst Technol 1(3):70–88
18. Pourvahab M, Ekbatanifard G (2019) Digital forensics architecture for evidence collection and provenance preservation in iaas cloud environment using sdn and blockchain technology. IEEE Access 7:153349–153364
19. Rane S, Dixit A (2019) Blockslaas: blockchain assisted secure logging-as-a-service for cloud forensics. In: International conference on security and privacy. Springer, pp 77–88
20. Ricci J, Baggili I, Breitinger F (2019) Blockchain-based distributed cloud storage digital forensics: where's the beef? IEEE Secur Priv 17(1):34–42
21. Zhang Y, Wu S, Jin B, Du J (2017) A blockchain-based process provenance for cloud forensics. In: 2017 3rd IEEE international conference on computer and communications (ICCC). IEEE, pp 2470–2473
22. Truffle quickstart—Truffle Suite

Classification of Knee Osteoarthritis Using Deep Learning a Rigorous Analysis

Punita Panwar, Sandeep Chaurasia, and Jayesh Gangrade

Abstract Knee Osteoarthritis (K-OA) is a form of degenerative joint disorder, which expressively affects the people above 45 years. K-OA is one of the common type of arthritis causing discomfort, mobility restriction and affecting quality of lifecycle and independency of people. K-OA is a chronic disability that causes sustainable impact on economy and society. Generally, radiologists review MRI images or X-ray of knee and assign KL-grades but Manual diagnosis of K-OA is very time consuming and tedious task so, to overcome this problem deep learning methodologies like convolutional neural network have been established to increase the efficiency and workflow in clinical department. We reviewed the state of art methodologies for classification of K-OA based on deep learning techniques. The aim of current review paper is to provide a brief overview to the researcher about the currently existing method of classification in Knee Osteoarthritis.

Keywords Knee Osteoarthritis (K-OA) · Classification · Convolutional neural network (CNN) · Deep learning

1 Introduction

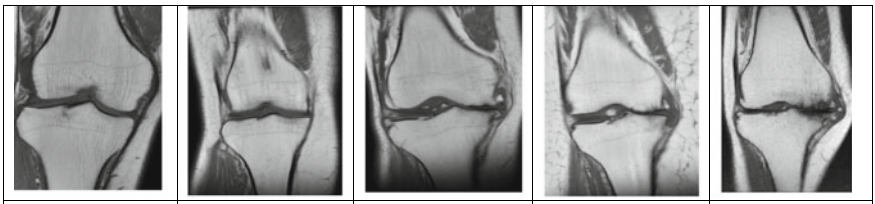
Knee osteoarthritis (K-OA) is the most common type of musculoskeletal disease that includes inflammation and major changes in structure of knee joint. In human body, K-OA affects specially the whole knee joint and it mainly happens when degeneration starts in articular cartilage and generation of osteophytes near joints [1]. K-OA, also known as degenerative joint disease because of wear and tear of the articular cartilage

P. Panwar (✉) · S. Chaurasia · J. Gangrade
Manipal University Jaipur, Jaipur, Rajasthan, India
e-mail: punita.202303038@muj.manipal.edu

S. Chaurasia
e-mail: sandeep.chaurasia@jaipur.manipal.edu

J. Gangrade
e-mail: jayesh.gangrade@jaipur.manipal.edu

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_44



G0	G1	G2	G3	G4
Grade 0 No Radigraphic	Grade 1 Doubtful Joint Space Narrowing (JSN) and Possible Osteophytic lipping	Grade 2 Possible JSN and Definite Osteophytes	Grade 3 Definite JSN Sclerosis and Multiple osteophytes	Grade 4 Marked JSN Severe sclerosis and Large osteophytes

Fig. 1 KL-grades distribution of right knee MRI

[2]. Due to this disorder, major causes such as activity limitation, pain and physical disability has been seen in older adult. Major risk factors that triggers K-OA include gender, age, obesity, trauma, heredity, bone deformities, and life style. It has been exposed that K-OA has a more severe effect on women, due to knee anatomy, prior knee injury, and hormonal impacts can play a major role [3]. K-OA majorly affects medial joint, lateral joint, and patellofemoral joint in knee and generally develops gradually over 10–15 years. Radiographs of knee Osteoarthritis patients probably show a joint space narrowing (JSN), due to formation of osteophytes (bone spurs) and cartilage loss [2].

In 2010–2012, Doctors diagnosed around 52.5 million persons affected by arthritis and 22.7 million persons has impact of arthritis attributes. Hootman et al. [4] and in 2013–2015, it is increased by 1.9 million which is 54.4 million persons affected by arthritis and by 2040, it is predicted to affect 78.4 million people [5]. There are many medical imaging techniques are available for measuring the knee joints such as X-ray, Magnetic Resonance Imaging (MRI), and Computed Tomography (CT) [6]. X-ray radiographs has high resolution and highest reasonable imaging method to detect joint abnormalities [7]. However MRI is additional sensitive than radiographs images because it generates high soft tissue structures that is easily understandable by the radiologist for Measuring the Severity of K-OA (shown in Fig. 1) and it can help in categorizing injuries associated to bones, ligament, meniscus, and cartilage [8].

2 Literature Study Selection

The systematic search was carried out through PubMed, IEEE Xplore, Science Direct, Google Scholar, and Web of Science until Dec, 2022. Total 542 articles searched via mentioned databases shown in Fig. 2. After the search procedure, 45 articles were selected.

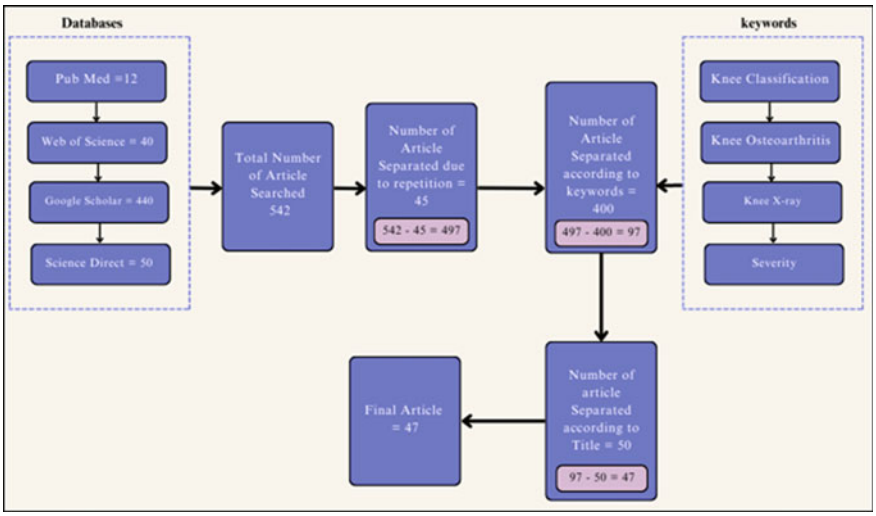


Fig. 2 Literature search process

3 Deep Learning

Artificial neural network is used in deep learning because they are developed to resemble the way human mind processes information as well as learns from them. It comprises layers of interconnected nodes, called neurons, which process and transmit information. In deep learning, ANNs are used to build deep neural networks with multiple layers of neurons stacked upon one another, allowing the network to learn complex representations of the input data. The layers in the neural network hierarchy hierarchically define peculiar features of the input data, and as the data passes through each layer, the network can extract more abstract and high-level representations of the input. This ability to extract useful features automatically is one of the primary strengths of deep learning models. Models of deep learning depicted remarkable performance through various applications. This section aims to provide a brief overview of numerous algorithms of deep learning that have been performed and studied through the examined papers.

3.1 AlexNet

AlexNet is an architecture for deep neural networks that specifically uses convolutional layers for image processing developed in the year 2012 that outshined the procedures used for classical image object recognition. AlexNet comprises eight layers, divided into 3 layers that are fully connected and 5 that are convolutional and it includes 650,000 Neurons with sixty Million of trainable criterions. AlexNet's

achievement in ImageNet competition pushed researchers to inspect the potential that deep learning has for solving categorization on the basis of image problems [9].

3.2 BVLC CaffeNet

In 2014, Berkeley Vision and Learning Center (BVLC) developed a single GPU version called 'CaffeNet'. It is more computationally efficient than AlexNet as it does not require data expansion and pooling layers to be allocated in advance to local response normalization layers. In the latest released model of CaffeNet, the sequence of local normalization layers and pooling is similar to AlexNet. Moreover, along with individual NVIDIA K40 GPU, the modified version of CaffeNet is able to operate over 60 million images every single day [10].

3.3 VGGNet Group

In 2014, VGG was released by a Group called Visual Geometry at Oxford University, a deeper neural network when compared with AlexNet, has an edge in using minor 3×3 convolution filters. VGG became the earliest network to have weight layers beyond 16 or 19, enabling researchers to include additional weight layers due to smaller size convolution filters used in the architecture. There are three fully connected layers and various weight layers inside VGGNet according to its structure, with VGG-16 having 3 layers that are fully connected and 13 are convolutional, and VGG-19 having 3 fully connected layers and 16 convolution layers. In this, we use ReLU which is an Activation Function. However, because of a large amount of weight layers, the VGGNet is tedious to instruct and has a heavy network size of above 550 mb, resulting in long inference times.

3.4 ResNet

In 2012, the invention of deep convolutional neural networks (DCNN) has given an increase in neural network's depth. However, simply increasing the layers did not always result in higher accuracy. Researchers found that there is an utmost threshold in the amount of layers in a neural network, beyond which, the % of error starts to expand, known as the exploding gradient problem [11]. To address this situation, the latest structure was introduced in the year 2015 that incorporated the notion of "Skip Connections". Skip Connection leaves training starting from small amount of layers and connects straight with the output. Multiple versions were created, including ResNet-50, ResNet-101, and ResNet-152. This has been observed ResNet that with 50/101/152 layers possess a lower rate of error for the

classification of image compared with a thirty four-layer plain net [11]. ResNet was constructed by placing separate blocks containing two to three convolution layers, which are called Residual Blocks.

3.5 ResNeXt

ResNeXt is a deep learning architecture that was invented as an advancement to ResNet, with a focus on image classification tasks. One of the main features of ResNeXt is the concept of 'cardinality', which enables the network to reduce the parameters as well as increase the layers. The ideal value of cardinality was observed which came out to be 32 [12]. The architecture of ResNeXt comprises the parallel placing of layers, as opposed to the sequential stacking used in traditional ResNet models. This is achieved through a 'split-transform-merge' strategy that was obtained from the Inception model. Features of the image are transmitted to the upsampling path from the downsampling path through skip connections [13].

3.6 SENets

SENets have gained significant attention in recent years due to their effectiveness in learning complex feature dependencies in a computationally efficient manner. The fundamental constituent of SEnets is the 'squeeze and excitation' block, which enables the network to improvise the channel interdependencies without extending the computational costs [14]. This block operates by squeezing the feature map of all channels into a unique numeric value then feeding it along two-layer neural network that gives a vector of the similar size (n) as the original feature map, i.e., where n is Number of Convolution Channels. These vector values are then used as weights to scale every channel on the basis of its significance [14].

3.7 Siamese Networks

These networks are a type of neural network that consists of two or more symmetrical subnets. The key characteristic of Siamese networks is that the weights between the identical subnets are shared [15]. It works by taking input along every identical subnet, comparing the result from the subnets, then deciding about the class to whom the given data (image) associates [16]. The architecture of a Siamese network is close to that of a CNN, with convolutional layers as well as pooling layers. However, unlike CNNs, Siamese networks do not have a softmax layer. Instead, the contrast of its result from the dense layers is sent through one neuron with a 'monotonic' activation function, which outputs either 0 or 1. To better understand how a Siamese network works, beginners can refer to [17].

4 Literature Review

See Table 1.

Table 1 Comparative analysis of classification method and techniques

References	Classification technique	Dataset	Source	Result
Tiulpin et al. [18]	Deep Siamese convolutional neural network	MOST = 18,376 Scans, OAI = 2957 scans	OAI, X-ray	Accuracy = 0.93
Chen et al. [19]	VGG-19 with proposed ordinal loss	4130 Scans	OAI, X-ray	Accuracy = 0.697, MAE = 0.344 (Mean Absolute Error)
Pedoia et al. [20]	DenseNet	4,384 T2 sequence subjects	OAI, X-ray	Accuracy = 0.834, sensitivity = 0.769 and specificity = 0.779
Norman et al. [21]	Ensemble-DenceNet	Total Scans = 39,593, training = 25,873,validation = 7779 testing = 5941	OAI, X-ray	Sensitivity normal = 0.83 mild = 0.70 moderate = 0.68 severe = 0.86, Specificity normal = 0.86 mild = 0.83 moderate = 0.97 severe = 0.99
Kim et al. [22]	CNN six SE-ResNet	Total = 4366 knee scans training = 3464 validation = 386 testing = 516	Hospital, X-ray	Accuracy in Kl0 = 0.97, Kl1 = = 0.85, Kl2 = = 0.75, Kl3 = 0.86, Kl4 = 0.95
Tiulpin and Saarakkala [23]	SE-ResNet50 and SE-ResNext50-32 × 4d	4796 participants training = 19,704, validation = 386 testing = 11,743	OAI and MOST	Accuracy = 0.98
Zhang et al. [24]	ResNet with convolutional block attention module	Total = 54,640 scans, training = 38,232, testing = 10,986, validation = 5422	OAI,X-ray	Accuracy = 0.748, MSE = 0.36, quadratic kappa score = 0.88
Leung et al. [25]	ResNet-34	4796 participants training = 19,704, validation = 386 testing = 11,743	OAI	Accuracy = 0.98

(continued)

Table 1 (continued)

References	Classification technique	Dataset	Source	Result
Liu et al. [26]	Faster R-CNN (Region proposal netwrol + Fast R-CNN)	Total = 54,640 images, training = 38,232, testing = 10,986, validation = 5422	From Hospital	Accuracy = 0.748, MSE = 0.36, Quadratic kappa score = 0.88
Nguyen et al. [27]	Deep Siamese CNN with pi model	Total = 728 Patients	OAI and MOST	Accuracy = 0.87
Swiecicki et al. [28]	Faster R-CNN and VGG-16	2770 Scans	MOST-MRI Posterior-Anterior	Accuracy = 0.825, sensitivity = 0.782 and specificity = 0.948
Dalia et al. [29]	YOLOv5 and VGG16, ResNet	4796 patients	OAI, X-ray	Accuracy = 0.698
Sarvamangala et al. [30]	MCB ResNet50	7424 graded knee	OAI, X-ray	Accuracy = 0.95
Olsson et al. [31]	ResNet	Total = 6103 scans, training = 5700, test = 300, validation = 403	Hospital-X-ray	Accuracy (K0 = 0.97),(K1 = 0.88),(K2 = 0.80),(K3 = 0.87),(K4 = 0.87)
Abdullah et al. [32]	Faster R-CNN + Modified ResNet-50 and AlexNet	3172 X-ray scans	LAB-X-ray	Knee JSW accuracy = 0.985, knee OA severity classification accuracy = 0.989, sensitivity = 0.99, specificity = 0.997
Ribas et al. [33]	SVM with a linear kernel, KNN with the Euclidean distance k = 9 and the Linear Discriminant Analysis (LDA)	688 knee images	OAI, X-ray	Accuracy = 0.816, sensitivity = 0.776,specificity = 0857
Wang et al. [34]	YOLOv2 for Knee joint segment and ResNet34 and DenseNet121 for classification	4796 participants	OAI, X-ray	Accuracy = 0.701(DenceNet121), 0.683 (ResNet34)
Cueva et al. [31]	Deep Siamese convolutional neural networks and a fine-tuned ResNet-34	4796 participants	OAI (Training) hospital (validation and test)	Average Multi class Accuracy = 0.617 and (K0 = 0.73),(K1 = 0.27),(K2 = 0.50),(K3 = 0.73),(K4 = 0.73)
	RefineDet	2579 images	Hospital, X-ray	Overall Accuracy = 0.977,

5 Conclusion

In conclusion, deep learning studies for knee images have accomplished incredible development. In the current study, we presented the summary of recent deep learning related research in the way of classification of diseases. In future deep learning-based methods may play a major part in the analysis, treatment, prognosis, imaging of knee osteoarthritis diseases. As the Artificial Intelligence technology develops and doctors become familiarized to Artificial Intelligence.

References

1. Brenner SS et al (2004) Osteoarthritis of the knee–clinical assessments and inflammatory markers. Osteoarthritis Cartilage 12(6):469–475
2. Lespasio MJ et al (2017) Knee osteoarthritis: a primer. Permanente J 21
3. Hame SL, Alexander RA (2013) Knee osteoarthritis in women. Curr Rev Musculoskelet Med 6:182–187
4. Hootman JM et al (2016) Updated projected prevalence of self-reported doctor-diagnosed arthritis and arthritis-attributable activity limitation among US adults, 2015–2040. Arthritis Rheumatol 68(7):1582–1587
5. Barbour KE et al (2017) Vital signs: prevalence of doctor-diagnosed arthritis and arthritis-attributable activity limitation—United States, 2013–2015. Morb Mortal Weekly Rep 66(9):246
6. Chan WP et al (1991) Osteoarthritis of the knee: comparison of radiography, CT, and MR imaging to assess extent and severity. AJR Am J Roentgenol 157(4):799–806
7. Mollenhauer J et al (2002) Diffraction-enhanced X-ray imaging of articular cartilage. Osteoarthritis Cartilage 10(3):163–171
8. Li X, Majumdar S (2013) Quantitative MRI of articular cartilage and its clinical applications. J Magn Reson Imaging 38(5):991–1008
9. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks.Commun ACM 60(6):84–90
10. jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM international conference on multimedia, pp 675–678
11. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–8
12. Xie S, Girshick R, Dollár P, Tu Z, He K (2017) Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 1492–500
13. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D,Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
14. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7132–7141
15. Chicco D (2020) Siamese neural networks: an overview. Artif Neural Netw 73–94
16. Li MD, Chang K, Bearce B, Chang CY, Huang AJ, Campbell JP, Brown JM, Singh P, Hoebel KV, Erdoğmuş D, Ioannidis S (2020) Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging. NPJ Digital Med 3(1):1–9
17. Siamese networks. line by line explanation for beginners—by krishna prasad—towards data science. https://towardsdatascience.com/siamese-networks-line-byline-explanation-for-beginners-55b8be1d2fc6. Accessed on 18 Dec 2020

18. Tiulpin A et al (2018) Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. Sci Rep 8(1):1–10
19. Chen P et al (2019) Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. Comput Med Imaging Graph 75:84–92
20. Pedoia V et al (2019) Diagnosing osteoarthritis from T2 maps using deep learning: an analysis of the entire Osteoarthritis Initiative baseline cohort. Osteoarthritis Cartilage 27(7):1002–1010
21. Norman B et al (2019) Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs. J Digital Imaging 32(3):471–477
22. Kim DH et al (2020) Can additional patient information improve the diagnostic performance of deep learning for the interpretation of knee osteoarthritis severity. J Clin Med 9(10):3341
23. Tiulpin A, Saarakkala S (2020) Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks. Diagnostics 10(11):932
24. Zhang B et al (2020) Attention-based cnn for kl grade classification: Data from the osteoarthritis initiative. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI). IEEE
25. Leung K et al (2020) Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: data from the osteoarthritis initiative. Radiology 296(3):584–593
26. Liu B, Luo J, Huang H (2020) Toward automatic quantification of knee osteoarthritis severity using improved Faster R-CNN. Int J Comput Assist Radiol Surg 15:457–466
27. Nguyen HH, Saarakkala S, Tiulpin A (2020) Deep semi-supervised learning for knee osteoarthritis severity assessment from plain radiographs. Osteoarthritis Cartilage 28:S311–S312
28. Swiecicki A et al (2021) Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists. Comput Biol Med 133:104334
29. Dalia Y et al (2021) deepoa: clinical decision support system for early detection and severity grading of knee osteoarthritis. In: 2021 5th international conference on computer, communication and signal processing (ICCCSP). IEEE
30. Sarvamangala DR, Kulkarni RV (2021) Grading of knee osteoarthritis using convolutional neural networks. Neural Process Lett 53(4):2985–3009
31. Olsson S, Akbarian E, Lind A et al (2021) Automating classification of osteoarthritis according to Kellgren-Lawrence in the knee using deep learning in an unfiltered adult population. BMC Musculoskelet Disord 22:844. https://doi.org/10.1186/s12891-021-04722-7
32. Abdullah SS, Rajasekaran MP (2022) Automatic detection and classification of knee osteoarthritis using deep learning approach. La radiologia medica 127(4):398–406
33. Ribas LC et al (2022) A complex network based approach for knee Osteoarthritis detection: data from the Osteoarthritis initiative. Biomed Signal Process Control 71:103133
34. Wang Y et al (2021) Learning from highly confident samples for automatic knee osteoarthritis severity assessment: data from the osteoarthritis initiative. IEEE J Biomed Health Inf 26(3):1239–1250

Meeting Manager

Sonal Sharma, Mahi Rastogi, P. Nagasundar, Sai Kaustubh, and Spurthi Srinivas

Abstract In recent years, the world has encountered numerous challenging situations that have made in-person meetings cumbersome. Therefore, the use of video conferencing apps has transitioned from being a luxury to a necessity. Nowadays, video conferencing is the primary means of communication for school, university, and work- related discussions. With the ease of scheduling online meetings, individuals tend to schedule multiple meetings in a single day, sometimes on short notice, leading to mismanagement of time and difficulty in keeping track of links and groups. Therefore, it is essential to manage time effectively and prioritize meetings to avoid burnout and confusion. Setting clear guidelines for scheduling meetings and prioritizing tasks can help individuals better manage their time and prevent overlap or double-booking. While some video conferencing apps offer online calendars that provide reminders for scheduled meetings, others send email reminders, which can cause confusion and result in missed meetings or overlapping schedules. This can cause significant disruptions in the workday and impact productivity. To overcome this challenge, individuals can take certain steps to ensure effective communication and time management. One such approach is to identify the preferred video conferencing apps used by their organization and prioritize their usage accordingly.

Keywords APIs · Node.js · Virtual conferencing platforms

S. Sharma (✉) · M. Rastogi · P. Nagasundar · S. Kaustubh · S. Srinivas
Jain (Deemed-to-Be University), Bengaluru, India
e-mail: s.sonal@jainuniversity.ac.in

M. Rastogi
e-mail: 19btrct024@jainuniversity.ac.in

P. Nagasundar
e-mail: 19btrct027@jainuniversity.ac.in

S. Kaustubh
e-mail: 19btrct036@jainuniversity.ac.in

S. Srinivas
e-mail: 19btrct039@jainuniversity.ac.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_45

1 Introduction

1.1 Background of the Project

Meetings are an integral aspect of any business, with virtual meetings quickly gaining traction as they provide a platform for discussing and resolving issues, updating team members and management on project progress, and aligning everyone's efforts to achieve optimal outcomes or to discuss future prospects [1]. In any organization, each member is accountable to two pools of individuals: the upper pool or higher-ups who require constant updates and the team members or subordinates who work alongside them or in different teams. When scheduling meetings, it is important to consider both of these pools, as individuals may receive invitations to attend meetings hosted by either pool. To effectively manage meetings and ensure that they are productive, individuals need to carefully consider the purpose and goals of the meeting, the necessary attendees, and the best approach to facilitate effective communication. This can involve using appropriate video conferencing apps or platforms, providing clear and concise agendas, and allocating sufficient time for discussion and decision-making [2].

By prioritizing effective communication and collaboration, organizations can ensure that meetings are productive and contribute to achieving the desired outcomes. This can help teams to remain aligned and focused on their goals, while also providing a platform for addressing challenges or opportunities that may arise along the way. In today's fast-paced world, it is common for individuals to forget their planned tasks, especially online meetings. Despite setting up multiple reminders, overlapping meetings, unavailability of members, and various other reasons can cause people to miss important meetings.

To address this challenge, we propose a "Meetings Management System," which allows users to efficiently create, update, and manage their meeting schedules on a single platform. Additionally, the system sends out links to members in the group, making it easier to schedule meetings with the required participants. The platform simplifies the process of scheduling meetings, allowing users to select their preferred time slots and choose from available options such as Zoom and Google Meet. This flexibility ensures that individuals can select the best-suited option based on their needs and preferences. By using this Meetings Management System, users can streamline their meeting schedules and ensure that they do not miss any important meetings. The system provides a comprehensive view of all scheduled meetings and alerts users about upcoming meetings to keep them on track. Furthermore, it eliminates the need for multiple reminders and helps users to manage their meetings efficiently.

The Meetings Management System is a comprehensive solution for individuals to schedule, manage, and stay on top of their online meetings, improving productivity and reducing the risk of missing important meetings. In today's world, video conferencing has become a vital tool for individuals to communicate with one another, especially during the pandemic when physical meetings are restricted. As a result,

there has been a surge in the popularity of video conferencing apps, with many people using them for work-related meetings, online classes, and even casual group video calls. However, with the plethora of video conferencing apps available, users often find it challenging to decide which one to use. Currently, popular options include Google Meet, Zoom, and Microsoft Teams.

Our proposed app eliminates the need for users to switch between multiple apps to manage their scheduled meetings. Our app provides a single platform for users to manage their meetings across all these apps. This feature eliminates any confusion about where the scheduled meeting is taking place and enables users to access their meeting details conveniently in one place. By using our app, users can efficiently manage all their meetings occurring in various apps on a single platform, making it easier to keep track of schedules and preventing conflicts between overlapping meetings. The app provides a streamlined solution to help users to manage their online meetings effectively and increase productivity. In conclusion, the proposed app offers a unique solution to the challenge of managing multiple video conferencing apps by providing a comprehensive platform for users to manage their meetings effectively. With this app, users can manage all their meetings seamlessly, thus increasing productivity and reducing confusion.

1.2 Problem Statement

Organizing online meetings and keeping track of them on different apps can be a daunting task that requires constant attention to avoid missing or being late to important events. Moreover, with the availability of multiple apps for video conferencing, it becomes a hassle to manage everyday life. People interact with different groups regularly on various apps, and each group has a different preference when it comes to video conferencing apps. This results in the need to download and manage multiple applications, which can be frustrating and time-consuming. To address this issue, the proposed app in this paper provides a single platform to manage all the meetings occurring on various apps, eliminating the need to constantly switch between different apps, and simplifying the process of managing meetings.

2 Literature Review

2.1 Existing Methods

Virtual video conferencing software has become an integral part of organizational communication, allowing businesses to conduct meetings with attendees from around the world, increase productivity, reduce expenses, save time and resources, and enhance accessibility. Various platforms, such as Google Meet, Microsoft Teams,

and Zoom Meetings, offer these services to users. Attendees are usually provided with a meeting link in advance, which they can use to join the meeting at the scheduled time. These video conferencing tools provide an experience similar to an in-person meeting, with features such as screen sharing, virtual backgrounds, and chat options [3]. Different video conferencing platforms offer a variety of features to suit the needs of the users. The unique features of different platforms can help users choose the one that best suits their needs and preferences.

(1) Google Meet allows users to have one-on-one or group calls that can last up to 24 h. Additionally, it provides full integration with other Google products such as the option to add the scheduled event to your Google Calendar for easy access. It has a time limit of 60 min on calls and supports up to 100 participants in the free account and also has live captioning available. Only one screen can be shared at a time, there is not a waiting room for participants and it needs an improvement privacy wise [4]. Google Meet allows screen recording only when joined through a laptop.
(2) Zoom offers the ability to record meetings from the attendees' perspective, if enabled by the meeting host. Moreover, it also provides breakout rooms that are available with all meeting subscriptions. It supports multiple screens' sharing and co-annotation. It supports up to 100 participants and has a 40-min time limit on free calls. Has a meeting room for participants to join prior a meeting, allows up to 49 persons on screen and has commendable security features [5].
(3) Microsoft Teams allows seamless integration with Microsoft 365, Skype, OneDrive, and other Microsoft products. Its app bar allows pinning of personal apps. It has a good security system with URL scanning, data encryption of meetings, chats, and files. It supports up to 100 participants for a set time of 60 min. It allows screen recording and the recordings will be deleted after a set period of time [6].

With the availability of multiple videos conferencing platforms, keeping track of meeting schedules and links can be quite challenging. Attendees are required to constantly monitor their schedules to avoid any overlap or conflicts. Moreover, accessing different platforms for different meetings can be a hassle, and it may be challenging to remember which meeting is taking place on which platform. To manage these challenges, attendees often manually mark the event in their calendars and take note of the links to prevent any mix-ups. Additionally, attendees may have difficulty attending meetings that are missed or unavailable, and they may have to request meeting recordings from the admin, which can be inconvenient.

2.2 Proposed Method

We propose the development of a platform that allows users to sign up and use various video conferencing applications, such as Google Meet and Zoom, without any limitations. Users will be able to create and manage their schedules on a dashboard

and add them to their personal calendars. In addition to this, the platform will enable users to send out meeting links in a unique way, where the participants can vote for their preferred time slot. Based on the majority, the system will automatically schedule the meeting at that time and generate a video conferencing link through the chosen platform. The system will then share the invites with the participants via email. This process is illustrated below (see Fig. 1). The aim of this platform is to provide a seamless and convenient way for users to manage their meetings and avoid any scheduling conflicts.

2.3 Proposed Workflow

See Fig. 1.

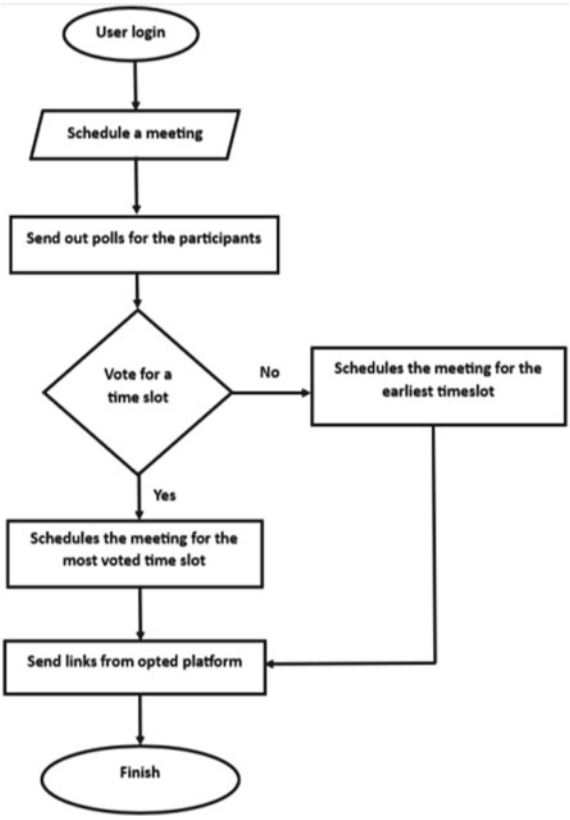


Fig. 1 Platform workflow [1]

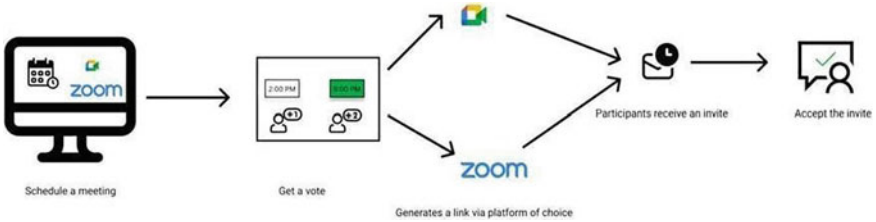


Fig. 2 Proposed workflow [2]

2.4 APIs

An Application Programming Interface (API) is a set of protocols, tools, and routines that allow different software applications to communicate and share data with each other [7]. APIs act as a bridge between different applications, allowing them to interact with each other in a standardized way without requiring access to each other's code or data. APIs can be used for a wide range of purposes, including accessing data from remote servers, integrating with third- party services, building custom applications, and automating tasks. They are commonly used in web development, mobile app development, and software integration projects. APIs typically use a standard set of communication protocols, such as Representational State Transfer (REST), Simple Object Access Protocol (SOAP), or GraphQL [8], which define how data are transmitted between different applications. APIs also often require authentication, which helps to ensure the security of the data being shared. APIs are typically offered by software vendors or service providers, who make them available to developers to integrate with their own applications. The documentation and code samples provided by the API vendor help developers to understand how to use the API and integrate it into their own projects (Fig. 2).

2.5 Google APIs

Google APIs are application programming interfaces provided by Google that allow developers to access data and functionality from various Google services such as Google Maps, Google Drive, Google Calendar, and more. Using Google APIs, developers can integrate Google services into their own applications, create custom interfaces, automate tasks, and retrieve data from Google's vast database. These APIs are available in a variety of programming languages, including Java, Python, PHP, and.NET, making them accessible to developers of all skill levels. Some examples of how Google APIs are used include displaying maps on websites, analyzing data from Google Analytics, retrieving search results from Google Search, and managing files in Google Drive. To use a Google API, developers must first register their application and obtain an API key, which provides access to the specific Google service.

Google provides comprehensive documentation and code samples to help developers get started with using their APIs [9].

2.6 Google Meet APIs

Google Meet API is a programming interface that enables developers to access Google Meet's video conferencing and collaboration capabilities within their own applications. The API allows developers to integrate video conferencing features like joining a meeting, starting a new meeting, and managing participant access from their own applications. Using the Google Meet API, developers can create custom interfaces, automate tasks, and retrieve data from Google Meet. This API is particularly useful for organizations that want to create their own custom video conferencing solutions or integrate Google Meet into their existing workflow or application. Some use cases of Google Meet API include creating custom meeting scheduling applications, developing virtual classroom solutions, and automating meeting attendance tracking for enterprises. To use Google Meet API, developers must first register their application and obtain an API key, which provides access to the Google Meet service. Google provides comprehensive documentation, code samples, and libraries in various programming languages such as Python, Java, and JavaScript to help developers get started with using their API [10].

2.7 Google Calendar APIs

The Google Calendar API is a programming interface that allows developers to access, manage, and integrate Google Calendar functionality within their own applications. With the API, developers can create, modify, and delete events, as well as manage calendars, attendees, and notifications programmatically. Using the Google Calendar API, developers can create custom interfaces, automate tasks, and retrieve data from Google Calendar. For example, developers can create applications that automatically schedule appointments, manage meeting rooms, and synchronize events across multiple calendars. The Google Calendar API is particularly useful for businesses, organizations, and individuals who rely on Google Calendar for managing their schedules and events. The API is available in a variety of programming languages, including Java, Python, PHP, and.NET, making it accessible to developers of all skill levels. To use the Google Calendar API, developers must first register their application and obtain an API key, which provides access to the Google Calendar service. Google provides comprehensive documentation, code samples, and libraries in various programming languages to help developers get started with using their API [11].

2.8 Zoom Meet APIs

The Zoom Meeting API is a programming interface that allows developers to access and integrate Zoom's video conferencing and collaboration capabilities within their own applications. The API enables developers to create custom video conferencing solutions, automate tasks, and retrieve data from Zoom. With the Zoom Meeting API, developers can create, modify, and delete meetings, manage participants, and schedule and start meetings programmatically. Additionally, the API provides access to a wide range of features, including screen sharing, recording, virtual backgrounds, and more. The Zoom Meeting API is particularly useful for businesses, organizations, and individuals who rely on Zoom for their video conferencing and collaboration needs. It is available in a variety of programming languages, including Python, Java, PHP, and.NET. To use the Zoom Meeting API, developers must first register their application and obtain an API key and secret, which provides access to the Zoom service. Zoom provides comprehensive documentation, code samples, and libraries in various programming languages to help developers get started with using their API. Additionally, the Zoom developer community provides a range of resources and support to help developers build their own custom video conferencing solutions [12].

2.9 Authentication

Authentication is the process of verifying the identity of a user or system attempting to access a particular resource or service. In the context of computer security, authentication is essential for verifying the identity of the user and ensuring that only authorized users can access sensitive information or perform certain actions. Authentication typically involves a username and password, which are used to verify the identity of the user. Other forms of authentication can include biometric data, such as fingerprint, iris scan or facial recognition, or the use of security tokens, such as smart cards or USB devices. Once a user's identity has been authenticated, they are granted access to the resource or service they are trying to access, depending on whether they have the authority. This process is known as authorization, which involves determining what actions a user is allowed to perform and what resources they are allowed to access. Authentication and authorization are essential components of computer security and are used in a wide range of applications and systems, including web applications, mobile apps, and operating systems. They help to ensure that only authorized users can access sensitive information or perform certain actions, which helps to protect against unauthorized access, data breaches, and other security threats [13].

Types of Authentications There are several types of authentication methods that can be used to verify the identity of a user or system. The choice of authentication method depends on the specific requirements of the system or application, as well

as the level of security needed to protect sensitive information or resources. Some of the most common types of authentications include:

(1) Password-based authentication: This is the most commonly used form of authentication, which involves the user entering a password that is verified by the system.
(2) Multi-factor authentication (MFA): This method requires users to provide two or more pieces of information to verify their identity, such as a password and a security token, or a fingerprint and a PIN code.
(3) Biometric authentication: This method involves using unique physical characteristics of a user, such as fingerprints, facial recognition, or iris scans, to verify their identity.
(4) Certificate-based authentication: This method uses digital certificates to verify the identity of a user or system. Certificates are issued by a trusted third party, such as a Certificate Authority (CA), and can be used to verify the identity of a user or system.
(5) Token-based authentication: This method involves using a security token, such as a smart card or USB device, to verify the identity of a user or system.
(6) Social media authentication: This method involves using social media platforms, such as Facebook or Google, to authenticate a user's identity.

2.10 Hardware and Software Requirements

Hardware and software requirements are:

(1) React.
(2) TypeScript.
(3) Graphql.
(4) Mongodb.
(5) Nodejs.
(6) Docker and Kubernetes (Microservices Architecture).
(7) CircleCi/Jenkins (CI/CD pipeline).
(8) GCP (Google Meet API).
(9) ZOOM (API).
(10) Amazon SES (Mail Service).

3 Implementation

3.1 Front-End UI

Login page In Fig. 3, we are using Google authentication flow which is designed to handle both existing and new users in a seamless and secure manner. This is implemented using authorization code grant, which is a type of OAuth flow used for web server applications. The flow works as follows: When a user attempts to log in to our application, they are redirected to the authorization server where they are prompted to enter their login credentials. Once the user is authenticated, the authorization server generates an authorization code and sends it back to our application's server. Our application then exchanges this authorization code for an access token, which is used to access the user's resources on the authorization server. If the user already exists in our database, we use the access token to log them in to the application. If the user is new, we create a new user record in our database using the information obtained from the authorization server, and then log them in. By using the authorization code grant, we ensure that user credentials are never shared with our application, and instead, the authorization server handles authentication and issues to access tokens. This adds an extra layer of security to our authentication flow and helps to protect user data. Overall, our authentication flow using the authorization code grant is designed to be secure, efficient, and user-friendly, making it easier for users to access our application while maintaining the integrity of their data.

OAuth authentication In Sect. 2.9, we learned about authentication and its importance in verifying the identity of users accessing a system or application. We also learned about different types of authentication methods, including password-based authentication, biometric authentication, and multi-factor authentication. In this section, we will focus on OAuth and social media authentication, which is a type of authentication that has gained popularity in recent years. OAuth is an open standard

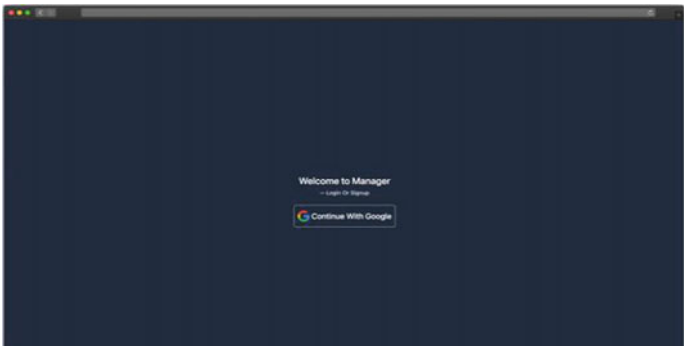


Fig. 3 Login page [3]

for authorization that allows users to grant access to their resources on one site (such as Facebook or Google) to another site (such as a mobile app or a third-party service) without revealing their username and password. Social media authentication is a specific implementation of OAuth that allows users to authenticate themselves using their social media accounts, such as Facebook, Twitter, or Google. Instead of creating a new account or entering login credentials, users can simply click a button that says "Sign in with Facebook" or "Sign in with Google" and grant permission for the application to access their profile information. The benefits of social media authentication are numerous. For users, it eliminates the need to create and remember multiple login credentials for different sites and applications. For developers, it simplifies the authentication process and can increase user engagement by reducing the barriers to entry. For end users, removes the need to remember the credentials to multiple accounts. However, it is important to note that social media authentication does have some drawbacks, such as the potential for security vulnerabilities if not implemented properly, as well as concerns around user privacy and data sharing. It is crucial for developers to carefully consider the trade-offs and ensure that their authentication systems are secure and protect user data.

Open Authorization (OAuth) is an authorization protocol that allows third-party applications to access a user's resources without requiring the user to disclose their login credentials. This protocol is widely used by companies such as Google, Facebook, and Twitter to allow users to access third-party applications with their existing account credentials. OAuth works by separating the roles of the client application, the user, and the resource server. When a user attempts to access a client application, the client application redirects the user to a login page provided by the resource server. The user enters their login credentials on the resource server's page, which verifies their identity and provides the client application with an access token. This access token allows the client application to access the user's resources on the resource server without requiring the user to disclose their login credentials to the client application. The access token is typically time-limited and can be revoked by the user at any time. OAuth is a secure way to authenticate users and allow third-party applications to access their resources [14]. However, it is important to note that OAuth is not a replacement for secure communication between the client application and the resource server. It is still necessary to use encryption and secure communication protocols to protect the user's data from interception and tampering.

Authorization code grant The authorization code grant is a type of OAuth 2.0 flow used for web server applications to obtain an access token on behalf of a user. It is considered to be the most secure and widely used OAuth flow for server-side applications. The flow works as follows:

(1) The user attempts to access a protected resource on the application server and is redirected to the authorization server to authenticate.
(2) The user provides their credentials to the authorization server and, if successful, the authorization server issues an authorization code to the application server.

(3) The application server then sends a request to the authorization server, including the authorization code and the client secret, to exchange the authorization code for an access token.
(4) The authorization server verifies the authorization code and client secret and issues an access token to the application server.
(5) The application server can then use the access token to access protected resources on behalf of the user.

The key advantage of the authorization code grant is that it separates the authorization process from the resource owner's credentials, which means that user credentials are never shared with the application server. This adds an extra layer of security to the authentication process and helps to protect user data. In addition, the authorization code grant allows the application server to obtain an access token with a longer lifespan than other OAuth flows, which is especially useful for applications that require long-lived access to user resources. Overall, the authorization code grant is a secure and efficient way for web server applications to obtain access tokens on behalf of users, making it an essential component of modern authentication and authorization protocols [15].

Figure 4 (see Fig. 4) [16] provides a visual representation of the authorization code grant flow, illustrating how an application communicates with an authorization server to obtain an access token on behalf of a user. By examining this Fig. 4, we can gain a deeper understanding of how the authorization code grant works and how it was implemented in our application.

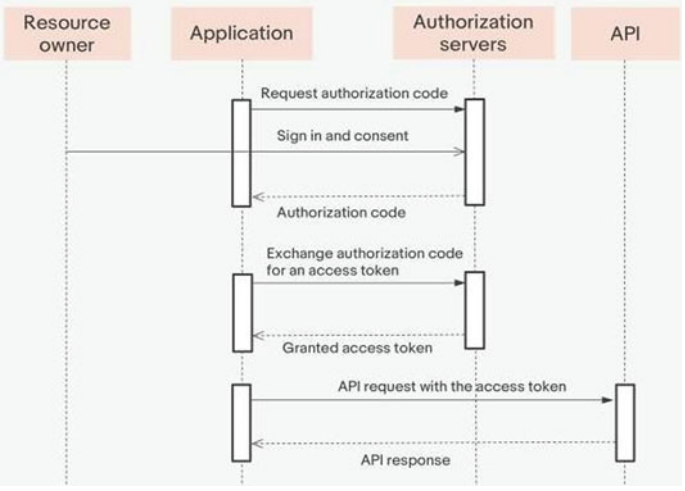


Fig. 4 Authorization process sequence diagram [4]

3.2 Dashboard

A meeting scheduler dashboard allows users to manage their meetings effectively. It comprises several components that enable users to schedule, track, and monitor meetings efficiently at just a glance:

(1) The first component is "Upcoming Meetings." It provides a comprehensive overview of all scheduled meetings in the upcoming weeks or months. This component should display the meeting date, time, duration, and attendee names. This feature allows users to plan their schedules and ensure that all necessary attendees are available for the meeting.
(2) The second component is "Finished Meeting Details." It offers statistics on meetings that have already occurred. These statistics could include details such as meeting duration, attendance rate, and participant engagement levels. This feature enables users to evaluate the success of their meetings and identify areas for improvement.
(3) The third component is "Groups." It helps users to manage multiple attendees efficiently by grouping them into specific categories. Users can create groups for different projects, departments, or teams, making it easier to schedule meetings for multiple attendees simultaneously.
(4) The fourth component is "Settings." It allows users to customize their dashboard view, set up meeting reminders, and configure notifications. Users can adjust the dashboard's layout to suit their preferences and create notifications to remind attendees about upcoming meetings.

Overall, a meeting scheduler dashboard provides users with a comprehensive and organized approach to scheduling and managing meetings. It offers an easy-to-use interface, enabling users to create, track, and evaluate meetings effectively (Fig. 5).

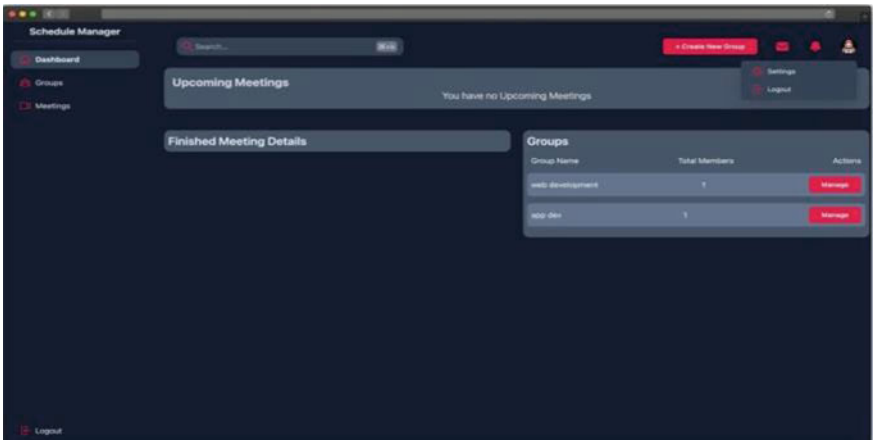


Fig. 5 Dashboard [5]

Creating a group To create a group, the user is required to provide certain information to ensure that the group is set up correctly. These details include the group name, the organization it belongs to, a brief description of its main purpose, and the participants that will be added to the group.

(1) The first detail required is the "Group Name." It should be a descriptive and straightforward name that accurately represents the group's purpose. This name should be unique to avoid confusion with other groups in the organization.
(2) The second detail is the "Organization it belongs to." This is important to ensure that the group is associated with the correct organization, especially for large corporations or organizations with multiple subgroups.
(3) The third detail is the "Description of the group." It should provide a clear and concise summary of what the group aims to achieve. This description helps potential members understand the group's purpose and decide whether to join.
(4) Finally, the user needs to "Add Participants" to the group. The user can add members manually by inputting their names or email addresses, or by importing them from an existing database. It is essential to ensure that all necessary participants are added to the group to ensure the success of any scheduled meetings or events (Fig. 6).

Group listing The Group listing page includes several essential elements to provide potential members with the necessary information about the group.

(1) Firstly, the "Group Name" should be clear, concise, and accurately reflect the group's purpose. This name should be easy to understand and memorable to help potential members find the group easily.

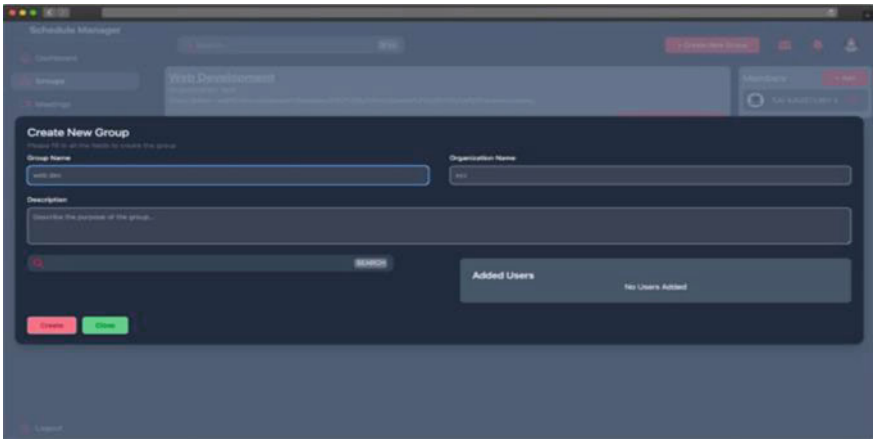


Fig. 6 Creating groups [6]

(2) Secondly, the "Organization Name" should be displayed to help potential members find groups that align with their interests or affiliations. This information allows users to identify groups within their organization that they may be interested in joining.
(3) Thirdly, the "Group Size" should be displayed to help potential members understand the level of engagement and interaction within the group. For example, a smaller group may offer more intimate discussions, while a larger group may provide more diverse perspectives.
(4) Fourthly, the "Group Members" should be displayed to allow potential members to see who is currently in the group. This information provides insight into the group's demographics and can help potential members decide whether the group is the right fit for them.
(5) Lastly, the "Manage Group" option should be included to help group administrators manage the group's settings and restrictions. This feature allows administrators to control who can join the group, the level of access members have, and other essential details needed to ensure the group's smooth functioning (Fig. 7).

Manage group The meeting overview component of the group dashboard provides a concise summary of all scheduled meetings for that particular group. It includes the meeting title, number of attendees, and date and time when the meeting is scheduled to be held. This information allows group members to prepare and plan accordingly, ensuring that they attend the meeting and are well prepared.

(1) Additionally, this component provides an option to create a new meeting, enabling group administrators to schedule future meetings with ease. The meeting creation feature allows administrators to input the meeting title, select the date and time, and specify the number of attendees required for the meeting.
(2) Moreover, the "Add Users" option is also provided to ensure that all necessary group members are invited to attend the meeting. Group administrators can add

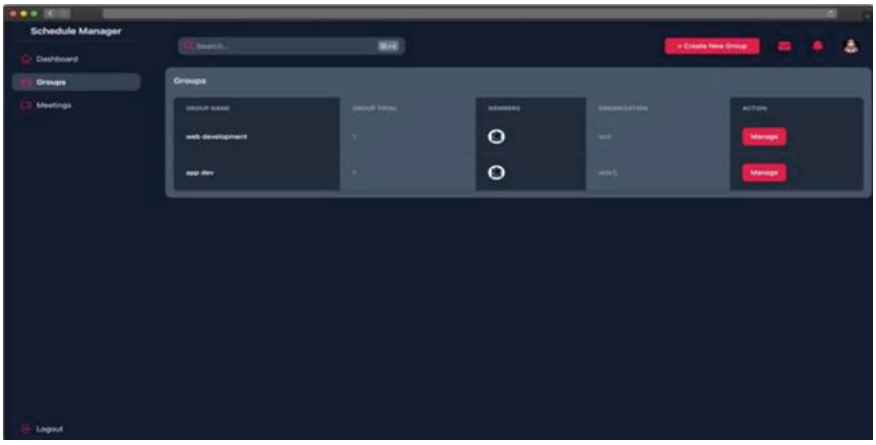


Fig. 7 View all groups [7]

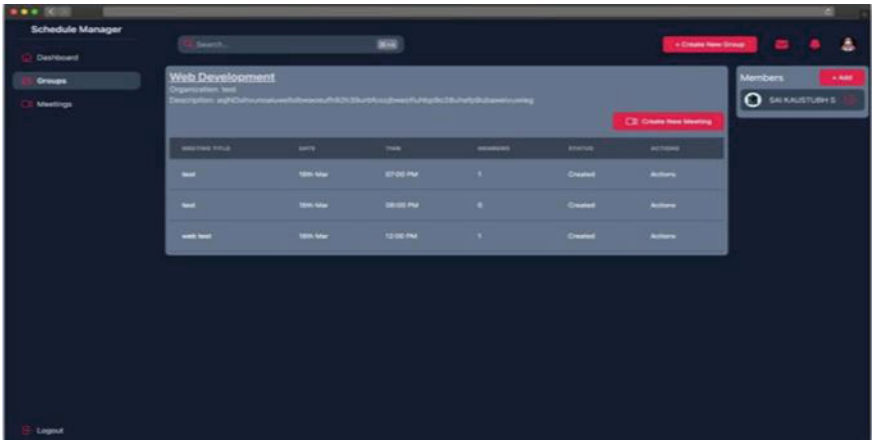


Fig. 8 Manage groups [8]

new members or invite existing members to attend the meeting, ensuring that all relevant stakeholders are present (Fig. 8).

Creating a meeting When creating a new meeting, the user is required to provide essential details to ensure that the meeting is properly organized and runs smoothly.

(1) Firstly, the user must provide a clear and concise "Title" that summarizes the meeting's scope. This title should be descriptive and easily understood by all attendees, allowing them to prepare and plan accordingly.
(2) Secondly, the user needs to provide a brief "Description" that explains the meeting's scope and purpose. This description should be detailed enough to give attendees a clear understanding of what will be discussed during the meeting.
(3) Thirdly, the user must specify the "Timing" and "Duration" of the meeting. This information is crucial as it enables attendees to plan their schedules accordingly and ensures that the meeting stays within the allotted time frame.
(4) Fourthly, the user must "Add Participants" to invite to the meeting. This step allows the user to invite all relevant stakeholders and ensures that all necessary parties are present during the meeting.
(5) Additionally, "Polling" is an optional feature that may be used based on the meeting's requirements. Polling allows users to gather feedback and opinions from attendees, which can be useful in making informed decisions during the meeting.
(6) Finally, the user must select the platform they would like to use, such as Zoom or Google Meet.

This selection ensures that all attendees have the necessary information to join the meeting and helps avoid any technical issues during the meeting (Fig. 9).

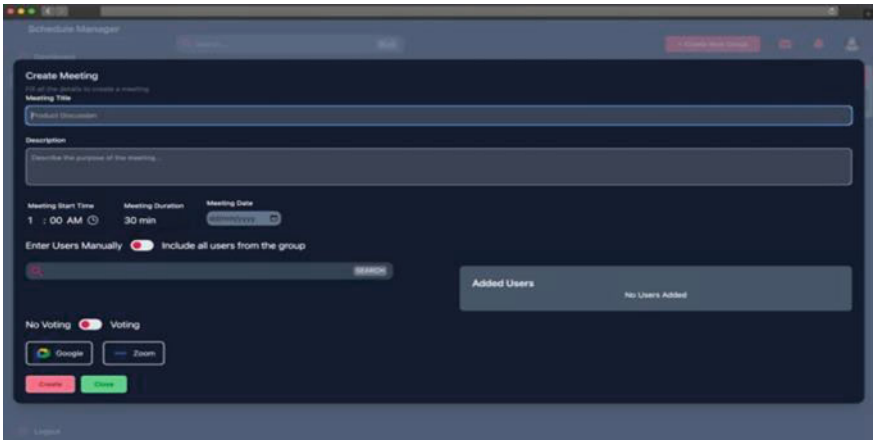


Fig. 9 Create a new meeting [9]

Google Meet NPM package This is a Node.js function named meet that creates a Google Meet event in the user's primary Google Calendar. The function takes an options object as a parameter, which should include the following properties:

(1) date: A string in YYYY-MM-DD format representing the date of the event.
(2) startTime: A string in HH:MM format representing the start time of the event.
(3) endTime: A string in HH:MM format representing the end time of the event.
(4) summary: A string representing the title of the event.
(5) location: A string representing the location of the event.
(6) description: A string representing the description of the event.
(7) clientId: A string representing the client ID of the Google API project.
(8) clientSecret: A string representing the client secret of the Google API project.
(9) refreshToken: A string representing the refresh token of the user's Google account.
(10) attendees: An array of objects representing the attendees of the event. Each attendee object should have an email property representing the email address of the attendee.
(11) alert: A number representing the number of minutes before the event start time that a reminder should be sent.

The function first checks if the user is busy at the specified time slot by sending a query to the Google Calendar API using the events.list method. If the user is busy, the function throws an error. If the user is not busy, the function constructs an event object and sends a request to the Google Calendar API using the events.insert method to create the event. The function returns the hangout link of the created event [17].

Zoom Meeting NPM package This is a Node.js module that provides a function zoom (options) which can be used to schedule a Zoom meeting and add it to a Google Calendar.

(1) The function takes an options' object as input which contains the necessary configuration parameters for both the Zoom and Google Calendar APIs, including the Zoom API key, Zoom API secret, developer email, Google OAuth2 credentials (client ID, client secret, refresh token), meeting details (title, start time, end time, attendees, description, location), and reminder settings.
(2) The function uses the JSON Web Token package to generate a JSON Web Token (JWT) based on the provided Zoom API key and secret and sends a POST request to the Zoom API to create a new meeting using the generated token. If the API call is successful, the function extracts the join_url, start_url, and password from the API response.
(3) Next, the function uses the Google API's package to authenticate with the Google Calendar API using the provided OAuth2 credentials, and checks if there are any conflicting events scheduled during the specified time slot. If there are conflicting events, the function throws an error and stops execution.
(4) Finally, if there are no conflicting events, the function constructs a new event object with the meeting details and inserts it into the user's Google Calendar using the calendar.events.insert method. The function then returns an object containing the join_url, start_url, and password for the scheduled Zoom meeting [18].

Meeting listing The meeting listing page is a crucial feature of our application that facilitates the organization and scheduling of meetings. This page enables users to view all their upcoming meetings in one place, providing them with a quick overview of their schedule. Additionally, this page provides users with information about the platform on which the meeting will be held, allowing them to prepare themselves accordingly.

(1) The meeting listing page typically provides users with a list of their upcoming meetings. This list usually includes the date, time, and duration of the meeting, as well as the name of the meeting's host. Additionally, the page may display the meeting's topic or agenda, which can help participants prepare for the discussion.
(2) One important feature of the meeting listing page is the ability to display information about the platform on which the meeting will be held. This information is typically presented as an icon or logo, which users can click on to access the platform directly. This feature can be particularly helpful for users who are not familiar with the platform, as it allows them to learn more about it before the meeting.
(3) If the host enables the voting feature, participants can vote on certain aspects of the meeting, such as the meeting time or agenda. Users who are not the host of the meeting can vote on these aspects by clicking on the vote button next to the relevant item. This feature can be particularly helpful for hosts who want to ensure that the meeting time and agenda are convenient for all participants. Upon clicking yes, the user will be added to the accepted list and on clicking no he would be added to the declined list. This would be handled through a dialog popup as shown (see Fig. 10).

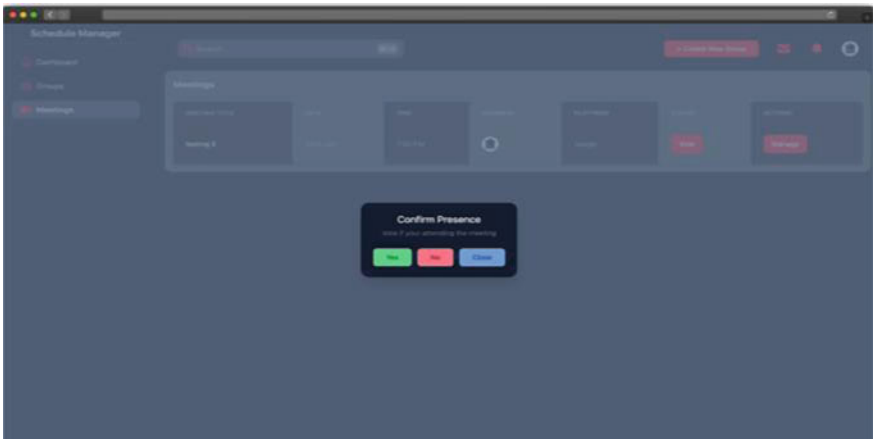


Fig. 10 Sending RSVP [10]

(4) The manage button allows users to access the meeting's functionality and view it in detail. The manage page typically provides users with information about the meeting's attendees, as well as any files or documents that have been shared in the meeting. Additionally, users can use the manage page to modify the meeting's settings, such as the meeting time, agenda, or platform.

Meeting details screen This screen provides users with detailed information about a particular meeting, such as the meeting title, description or agenda, meeting time, and platform information. Additionally, users can access a direct link to the meeting platform, making it easy for them to join the meeting.

(1) The meeting detail screen typically provides users with detailed information about a particular meeting. This screen includes the meeting's title, description, and agenda, which can help participants prepare for the discussion. Additionally, the screen displays the meeting time, including the start and end times, and the time zone. This feature can be particularly helpful for users who are in different time zones and need to adjust their schedule accordingly.
(2) Another critical feature of the meeting detail screen is the platform information. This feature provides users with information about the platform on which the meeting will be held. The screen displays the platform's name or logo, as well as any relevant information about the platform, such as the system requirements or the recommended browser. Additionally, users can access a direct link to the meeting platform, allowing them to join the meeting with just one click.
(3) The meeting detail screen typically includes a button or link that users can copy and share with others. This feature is particularly helpful for hosts who want to invite additional participants to the meeting or for participants who want to share the meeting details with their colleagues. Additionally, users can customize the link sharing options, such as setting permissions for sharing or restricting access to certain users (Fig. 11).

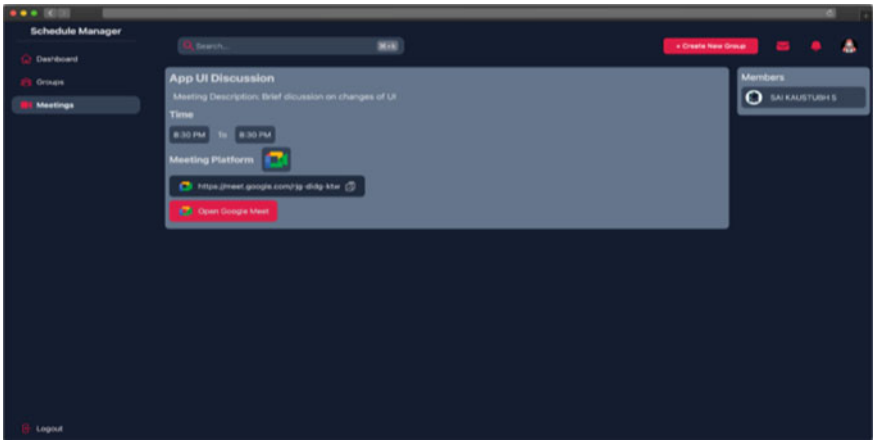


Fig. 11 Meeting details [11]

4 Conclusion and Future Scope

4.1 Conclusion

In today's fast-paced world, video conferencing has become an essential tool for businesses and individuals to connect with people around the globe. With the rise in demand for virtual communication, there has been an increase in the number of video conferencing platforms such as Google Meet, Zoom, and Microsoft Teams. While these platforms offer unique features and advantages, it can be difficult to manage multiple apps for different groups and schedules [19].

To address this issue, a Meetings Management System has been proposed that allows users to seamlessly create, update, and manage their meeting schedules on a single platform. This system sends out links to group members and allows users to easily schedule meetings based on their time preference, choosing from options such as Google Meet and Zoom. With the Meetings Management System, attendees no longer need to keep track of multiple links and schedules, and meetings will not overlap without the user realizing it. The platform also offers a unique feature where participants can vote for the time slot of their choice, and based on the majority, the system can automatically schedule the meeting and generate a video conferencing link through the desired platform. This will also eliminate the need for attendees to personally request meeting recordings from the admin. To use the platform, users can sign up with an email and password, and the system uses the bcrypt library to securely hash the password and store it in a NoSql database called MongoDB. With this, users can access all the aforementioned applications freely, schedule their tasks on a dashboard, and add them to their calendar. Overall, the Meetings Management System offers an efficient and user-friendly solution for managing multiple video

conferencing platforms and schedules, and can greatly enhance productivity and communication for businesses and individuals alike.

4.2 Future Scope

The future scope for applications such as these holds endless possibilities with more percentage of people spending a lot more time online, in the virtual realm. It can be better a better security system or user interface, and there are lots of scope for further improvement.

(1) With remote working and schooling online growing day by day, more popular meeting platforms such as Microsoft Teams, Cisco Webex, and other up and coming meeting platforms such as Whereby an Dialpad Meetings can be integrated to make the experience hassle free.
(2) It is easy to get lost with multiple tabs running at any given point. To avoid any confusion like this, the meeting will automatically be opened in the new tab if our application is open or running in the background.
(3) Meeting recordings and minutes can be sent to attendees who missed the meeting, allowing them to catch up without having to conduct another meeting.

References

1. https://www.investopedia.com/terms/v/video-conferencing.asp
2. https://www.researchgate.net/publication/239773432_Videoconferencing_as_a_Mode_of_Communication_A_Comparative_Study_of_the_Use_of_Videoconferencing_and_Face-to-Face_Meetings
3. https: //www. sylamtechgroup.com%2Fwp-content%2Fuploads%2F2019%2F04%2Fintro-video-conferencing-wp- engb.pdf&usg=AOvVaw2G1bpURRMhsNYvr2GiSO NQ
4. Comparing Zoom, Microsoft Teams and Google Meet
5. https://explore.zoom.us/docs/en-us/free-video-conferencing.html
6. https://www.stanfieldit.com/microsoft-teams/
7. https://en.wikipedia.org/wiki/API
8. https://stoplight.io/api-types
9. https://en.wikipedia.org/wiki/Google_APIs
10. https://cloud.google.com/blog/products/g-suite/how-google-meet-keeps-videoconferences-secure
11. https://developers.google.com/calendar/api/guides/overview
12. https://zoom.github.io/api/
13. https://www.sailpoint.com/identity-library/authentication-methods-used-for-network-security/
14. https://auth0.com/intro-to-iam/what-is-oauth-2
15. https://www.oauth.com/oauth2-servers/server-side-apps/authorization-code/
16. https://www.google.com/url?sa=i&url=https%3A%2F%2Fdeveloper.ebay.com%2Fapi-docs%2Fstatic%2Foauth-authorization-code-grant.html&psig=AOvVaw2l62OhkmOzHektS-U4FBJc&ust=1683651294623000&source=images&cd=vfe&ved=0CBEQjRxqFwoTCMiclIOY5v4CFQAAAA AdAAAAABA

17. https://www.npmjs.com/package/google-cal-meet-api
18. https://www.npmjs.com/package/google-cal-zoom-api
19. https://www.researchgate.net/publication/224218879_Augmented_reality_for_immersive_remote_collaboration

Microlearning: Sustainable Learning for Businesses

Suruchi Pandey and Sanjay Pandey

Abstract Corporate training is seeing a rise in the popularity of microlearning and learning more about it can help us understand why and how it is employed in the workplace. Understanding the advantages and applications of microlearning in training will help create and deliver training initiatives that are more efficient and interesting for participants. More research on microlearning applications can lead to enhance overall job performance and assist employees in acquiring the knowledge and abilities they need to do their job. Organizations can sustain learning for the long term and survive in a competitive business environment. Understanding microlearning and how the latest trend and technology can help professionals deliver training programs more engaging and successful. This article presents a discussion on microlearning leading sustainable learning for growing businesses. The article discusses pointers from corporate leaders, learners, and practitioners.

Keywords Microlearning · Training · Sustainable · Businesses · Employees

1 Introduction

To prevent training from being ineffective, it is imperative in the area of learning and development to deal with people's dwindling attention spans. One tactic for addressing this problem and helping learners better absorb knowledge is microlearning. The adaptability of microlearning is one of the key reasons corporate training functions more effectively with this approach. The smaller knowledge bits can be used in numerous contexts. Additionally, there are several ways to use microlearning nuggets [1]. The necessary variation can be provided by using a single microlearning course across multiple departments. Similar microlearning modules

S. Pandey (✉)
Symbiosis Institute of Management Studies, Symbiosis International Deemed University, Pune, India
e-mail: Suruchi.p@sims.edu

S. Pandey
Career Launcher, Pune, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_46

can be connected to construct longer training programs. Because they are virtual, microlearning programs are a more affordable method of employee training. Due to modem rapid authoring tools with templates and easily downloadable multimedia resources, the cost to build microlearning courses is also significantly lower. Because the content is simple, learners may acquire and remember more information [2].

Microlearning aids memory retention by presenting knowledge in manageable chunks. Employees can obtain the information they need whenever and however, they need it by using microlearning modules. It is just-in-time (JIT) learning, as the name suggests. By loading a microlearning module, learners can get immediate learning support rather than spending time and energy reading through a lot of material to tackle a straightforward problem. It improves their efficiency and broadens their knowledge of the issues and solutions pertinent to their field of work. JIT with microlearning is growing in favor among both learners and L&D professionals as a result of these adaptable delivery choices [3].

Information is accessible from anywhere thanks to this digital infrastructure of bite-sized modules, allowing learners the control to direct own professional development. Millennials and Gen Z, who make up the majority of the workforce today, have short attention spans and may become disinterested in courses if they are drawn out and tiresome. Microlearning increases learning outcomes since it actively engages learners. Microlearning is gaining more and more traction with learners since it can be done at anytime, anywhere. Utilizing microlearning tools like games, movies, case studies, and animations may make training more engaging and immersive [4].

The building of microlearning courses doesn't take very long. Consequently, they are easily customizable for each person and can be created to meet business specific demands. They make a major difference in their applicability and usage by corporate personnel, who want to access such modules while traveling or performing other duties because they were built using a mobile-first strategy.

2 Microlearning in Literature

Learning is the activity or process of acquiring knowledge or skills through study, practice, instruction, or first-hand experience. Microlearning is the practice of acquiring knowledge through brief informational bursts that are employed to achieve a particular learning objective. Videos, essays, ebooks, audio clips, and any other kind of content that is narrowly targeted enough to address an instant need can all be used to convey short informational snippets. A logical response in a quick-paced world that values learning effectiveness is to learn as much as you can in a short period. More information must be retained. Also, it is needed to keep data for longer periods. Additionally, this information is applied to other difficulties in life not just training. To scale and control the quality of learning for various types of learners and their background microlearning can go a long way [5, 6].

Bites/Bite-sized learning, chunks, micro-content, snacks, and other phrases are frequently used in conjunction with microlearning. As learning is best retained when it is applied in specific circumstances, microlearning typically includes practicing. Microlearning as known refers to breaking large chunks of information into smaller pieces which enable learners to better assimilate the information [4]. Traditional learning theories also proposed chunking of information for better learning and retention of learning. Here microlearning adoption has provided a solution to corporates that deals with multi-facet aspects. It is a go thing for learners and training administrators [7].

2.1 Where is Microlearning Exactly Used?

2.1.1 Compliance Education

The process of informing and educating employees about the various laws, rules, and corporate policies that regulate the workplace and directly affect their daily tasks and behaviors is known as compliance training. Microlearning will dramatically increase retention rates for compliance training. Popular subjects covered in compliance training are Anti-Harassment Education, Occupational Safety (OSHA), Training in Information Security, Substance Abuse at Work, Diversity and Workplace Violence Training, Medical and Family Leave, HR Laws and Procedures.

Compliance training may seem laborious, time-consuming, and excruciatingly repetitious. However, this is the reason why employers must act quickly. Microlearning design, which offers several benefits in terms of training time flexibility, cost effectiveness, and convenience of use, is the outcome of the need for a novel method that produces successful results [8].

2.1.2 Employee Education

Offering benefits to the human element can make up a sizeable portion of the benefits a business provides. Companies nowadays are aiming to increase employee performance as well as their happiness, health, and safety [9]. As a result, they devote a great deal of time and resources to employee education. These classes might cover Financial and Accounting, Programs for Corporate Sales Training, Consumer Assistance, Leadership and management in information technology for health and wellness, among many others. Used in online collaborative education too [10].

2.1.3 Customer Education

Often known as customer training, is any intentional, structured learning activity (such as onboarding) created by a business or industry to teach attitudes, information,

or abilities to consumers. Customers who are well-informed about your products are more likely to appreciate their full potential and progressively transition to new or more sophisticated uses for them. Customers are given a thorough how to manual that can address the vast majority, if not all, of their questions once they begin using the device. Any remaining sophisticated queries are then best handled by a customer success manager or member of customer support. Customer education is especially advantageous for businesses that have a complicated item, a product that demands customers modify their current behavior, products that routinely receive new features updates, products that need extensive support [11].

3 Methodology

Discussions with an expert who is practicing learning and development functions were done to understand the practitioner's perspective. Video and Audio interviews were taken from 30 employees across the industry. These semi-structured interviews lasted for 20–30 min. The selection of the sample was purposive. Thirty employees views were sought to map the practices. Few questions have quantitative representation in this study. Excel was used to make the graph and populate the data. Qualitative analysis was done of the Interview transcripts.

3.1 Data Collection

Questionnaire was used to capture informed on microlearning perspectives of learners. The working professionals who are either currently working or have had previous experience working in varied sectors were chosen as respondents. This helped in gathering genuine responses. The spectrum as can be seen from the Google form responses was huge constituting 14 distinct sectors–IT, manufacturing, consulting, hospitality, marketing, ecommerce, medical, education, banking and finance to list a few. This gave a much broader and clearer perspective as to how prevalent microlearning is in the industries and in what ways it's benefiting the employees, and in what ways it is being implemented across various industries.

Hygiene practiced during the interviews was kept by taking prior permission, arranging an appointment and informing objectives. Recordings were done with approval and transcripts were prepared for further analysis. The Interview Questions comprised of:

- To what extent is microlearning implementable in corporate training, particularly in the context of your company
- What are the advantages of microlearning to learners
- What are the benefits of micro-learning for business?
- What are your views on microlearning and different generations at workplace

- What are the formats in which microlearning can be deployed
- Share in brief your experience from the microlearning coursed taken
- What is a limitation of microlearning compared to traditional methods.

4 Analysis

The interview data is presented in a discussion based on the views expressed by the practitioners and learners. Responses comprising the major themes are quoted below. These were also echoed by major other respondents. R is used to indicate respondent.

R1. Many respondents informed that their company implemented microlearning training modules in their induction program and for technical skills courses like python [R1]. The implementation process was discussed with foreign university professors and experts in the Indian market. [R1, R3, R4].

Another respondent informed that his current employer uses microlearning training modules to teach drivers about driving safety and security. Also that microlearning entails learning in little doses. It is a quick learning exercise that takes about 2–5 min to complete and has a clear learning goal in mind. It was highlighted how microlearning empowers learners by giving them the freedom to learn whenever they choose, on any platform, and by giving them access to learning materials while they are working. [R6] Respondents were a strong proponent of microlearning and noted its advantages for learners, including learner-centricity, just-in-time instruction, accessibility, and reduced time commitment. [3] Respondent claims that microlearning is advantageous for every organization because it is inexpensive, agile, has a quick development cycle, and is simple to update. It also concurred that microlearning applies to all business sectors and unquestionably enhances performance [12].

IT enthusiasts mentioned that microlearning is more about self-development than just focusing on the job and what the company may require. She believes that it is very helpful to put something meaningful on her resume, especially when being a youngster, getting to learn more simultaneously working long shifts. She also informed that it is useful when it comes to being in touch and revising the longer session after a week or two without one's ability to grasp things being a hurdle in the process of learning. For her, microlearning is cost-effective and also time-saving. One can be engaged, involved, and learned at the same time without worrying much about the pace of learning things [R10].

The mode of communication generally is short audio and visual clips, which the learners can take up at their convenience and own suitable pace. In the interview, it was seen that microlearning is quite prominently implemented in organizations, particularly in IT companies like ours [R4].

Microlearning generates interest in individuals. See how millennials who are always keen on picking up new things and keen to learn, they constitute part of the workforce where microlearning has penetrated the maximum [R5].

R9. Coming to the benefits of microlearning, it has replaced the old hours-long training sessions and classroom sessions. It's easier to go the microlearning way. If one has to learn a new concept or get acquainted with a new technology all that he/she has to do is click on a link and the lessons are ready to be taken. It is good for refresher.

R17. As it is said nothing comes with only positives there are downsides to microlearning as well. It cannot be used as an alternative to deep and thorough knowledge about a particular topic which one gain when they attend formal classroom training and lessons.

R18, R23. Microlearning can help employees develop their soft skills. Soft skills are the character traits and conduct required for success in the workplace, such as leadership, teamwork, and communication. Employees who participate in microlearning receive brief, targeted material that they may utilize to swiftly upgrade their knowledge and abilities in these areas. For instance, a microlearning course on effective communication could give staff members pointers and methods for speaking more clearly at work. Employees could get knowledge on how to function well in a team, including techniques for resolving disagreements and enhancing collaboration, through a microlearning course on teamwork.

R20. Not just Millennials or Gen Z, associates like microlearning courses. Although it's common to associate these generations with a penchant for digital and mobile learning, microlearning courses can be useful for associates of all ages and generations.

R25. The flexibility, accessibility, and effectiveness of microlearning courses make them an easy and fast option for staff members to upgrade their knowledge and abilities.

Additionally, because of its potential to enhance employee performance and satisfaction, many firms are incorporating microlearning into their overall training and development strategy. As a result, regardless of the age or generation of the workforce, microlearning courses are growing in acceptance and popularity.

R11. The IT/ITES business is not the only one that uses microlearning as a training technique. While microlearning can be used in a variety of industries and organizations, it is most effective in the technology industry due to its fast-paced environment and demand for flexible, just-intime training and development possibilities.

R16. Microlearning, for instance, can be utilized in the healthcare sector to give staff members instant access to knowledge about safety protocols, patient care, and medical procedures. Microlearning can be utilized in the retail sector to train staff members on customer service, product knowledge, offers, and sales strategies. Additionally, microlearning can be utilized in the financial sector to teach staff members about rules, compliance, and financial analysis.

5 Results from Responses

This segment presents a few quantitative responses from the study.

1. Respondents comprised both the genders

From the data, it can be analyzed that only 11 women out of 33 respondents are being provided microlearning courses after joining the organization (Fig. 1).

The Author takes to the fact that all the buzz of the concept of DEI is only limited to recruitment. Afterward, that parity should be maintained. This is a common fact that the more courses an employee does increases the chance of him being promoted and also increasing his value in the organization. Therefore companies should give equal opportunities even after the recruitment is done. Therefore equality should not be limited to just reaching your company's fixed quota for women, but also companies should try to bring women to the same level by providing an equal number of microlearning courses. Women may not have the same access to mentors and networks as men, which might restrict their possibilities for growth.

2. Industry-wise respondents

From the data collected from the respondents, it was found out the maximum number of microlearning courses have been provided to the employees from IT sector (Fig. 2).

Because there is a significant need for technology workers with the necessary skills and knowledge to work in this field, the IT industry offers a lot of microlearning courses. Companies and organizations are continually looking for staff that are knowledgeable about the most recent technologies, strategies, and platforms because technology is always evolving. By assisting staff members in growing and broadening their skill sets, these businesses can respond to this demand by offering microlearning IT courses. Plausible explanations on why the IT industry offers so many microlearning courses. An industry that moves quickly: The information technology sector moves quickly and is continually changing. The newest trends and technology can be kept up with quickly with the help of microlearning courses.

- Demand for Upskilling: Due to the fierce competition in the IT sector, there is a constant need for upskilling and ongoing education. IT professionals can update their skills flexibly and conveniently using microlearning courses. After it sector, it can be observed that the second most prominent sector for microlearning courses

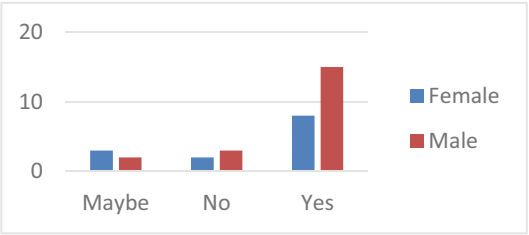


Fig. 1 Gender

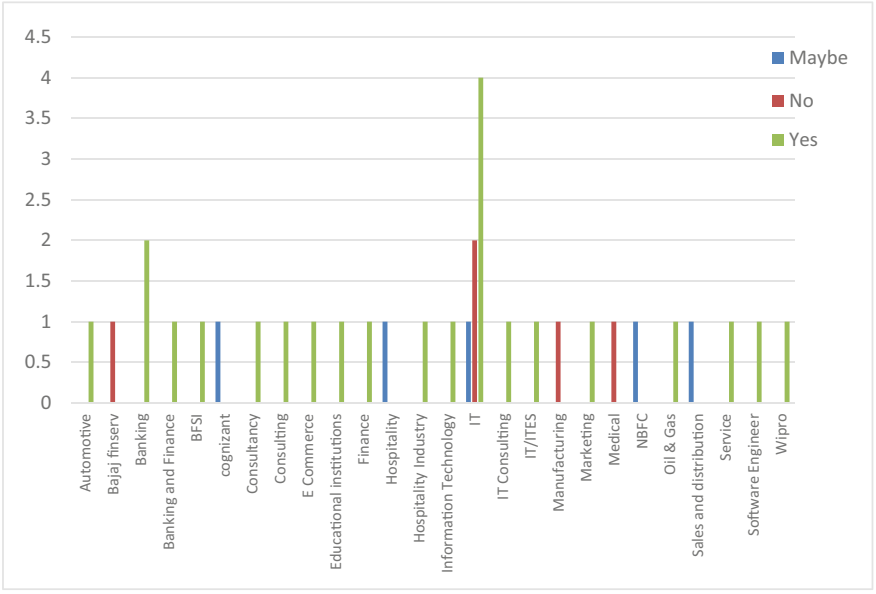


Fig. 2 Industry-wise response of microlearning

is the banking sector reasons why banks provide their staff with a wide variety of microlearning courses include:

- Increasing efficiency and productivity: Microlearning programs give workers the opportunity to learn new knowledge and skills quickly and effectively, which can boost their overall output.
- Enhancing job performance: Banks can assist their employees to enhance their job performance by offering microlearning courses that keep them abreast of the most recent best practices and strategies in their particular fields of expertise.

There are still several industries in India without microlearning programs. These sectors include, among others:

- Agriculture: Traditional methods still dominate the Indian agricultural sector, and there are few chances for technology-based training and growth.
- Small and Medium Enterprises (SMEs): Many SMEs in India lack the infrastructure and resources necessary to set up microlearning programs for their staff.
- Government offices and employees still lack access to microlearning courses and training programs, despite the government sector's efforts to modernize and digitize.
- Healthcare: The Indian healthcare industry is struggling with several issues, including a lack of qualified staff and inadequate training facilities. Microlearning programs could be able to help with some of these issues.
- Area and Department of Microlearning

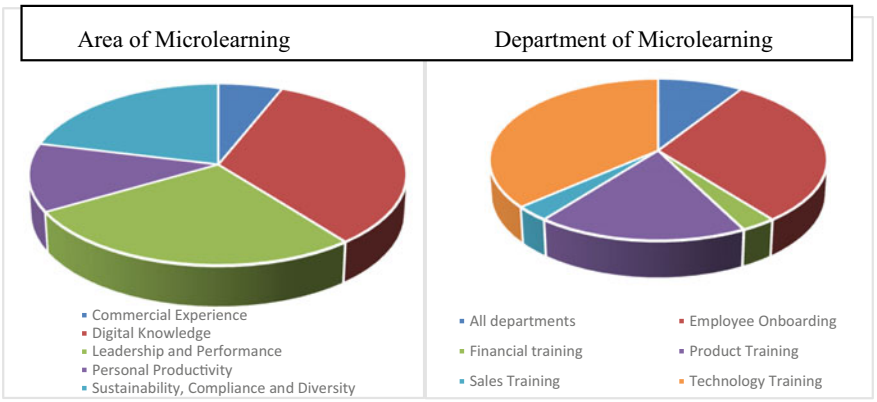


Fig. 3 Response on microlearning application in areas/department

From the data, it can be seen that there are some particular departments for which microlearning courses are used. During the employee onboarding the employee is new to the company therefore company tries to invest in that employee as much it can so that the employee believes that the company is doing something for him and in turn his efficiency increases Also, the addition of various skill sets and knowledge of the employees will indirectly help the company as it will give the employee a chance to increase his domain and not stick with a particular job role, that is the company can expand his or her job description. From the data, it can be analyzed that the technology department is also very much active when talking about microlearning courses. The IT industry spans a vast range of topics, from software development and programming languages to cybersecurity and data management. IT professionals can more easily learn what they need to know in a short period with the help of microlearning courses because they enable a deeper focus on particular topics (Fig. 3).

- Cost-Effective—Microlearning programs are frequently less expensive than conventional training techniques, opening them out to a wider spectrum of IT specialists.
- Technology Adoption—The IT industry is leading the way in terms of innovation and technology adoption. Technology can be used in microlearning courses to create fun and dynamic learning environments.
- Satisfaction with the microlearning opportunities

The majority 80% of the respondents were satisfied with the content and modules of microlearning. IT sector employees rated greater satisfaction than other industry.

The greatest microlearning courses are offered to employees in the IT industry for several reasons. The rapid rate of change: Since the technology sector is continually developing, workers must stay abreast of these advancements. The content in microlearning courses is just-in-time, bitesized, and can help workers swiftly upgrade their knowledge and abilities. Many IT workers are on the go and want access to training and development options that can fit into their hectic schedules.

Employees can access training materials whenever and wherever they need to by taking advantage of microlearning courses, which are made to be delivered through mobile devices.

6 Discussion

The study concludes that microlearning is a method of training that presents brief, precise chunks of knowledge in a way that is simple to comprehend and retain. Because it enables employees to swiftly pick up new knowledge and abilities in a way that is simple to consume and retain, microlearning is advantageous for businesses. It is a perfect fit for employee training programs due to this fact. Employee completion rates are higher when they can complete the training at their own pace and on their schedule, which is made possible through microlearning in the workplace. Overall, organizations can benefit from investing in microlearning since it can enhance productivity and learning results.

This is simply a business solution that works for multiple sectors. Microlearning can be applied in a wide range of training settings, including sales, marketing, production procedures, and shop-floor activity. Video courses with better content design, for instance, are fantastic for marketing education. To introduce new industrial workers to a company's production process, create a quick eBook or quiz. This demonstrates that no matter the situation, microlearning may be an effective training method. For training managers, microlearning can be a useful substitute for conventional classroom instruction. Comparing microlearning to standard classroom instruction has various benefits.

- Flexibility: Microlearning courses may be accessed whenever and from wherever making them a flexible choice for busy training managers who might not be able to attend conventional classroom sessions.
- Just-in-time learning: Microlearning courses are made to deliver knowledge in brief, precise bursts, making it simpler for training managers to get the information they require at the precise moment they need it.
- Learning that is more specifically customized to the demands and learning preferences of each learner is made possible by the use of microlearning courses.

7 Conclusion

Organizations are aware of the advantages of this method of development and training. Because businesses in this area are seeking ways to keep up with the rapid developments in technology and to give staff flexible, just-in-time training and development opportunities, microlearning has become particularly popular in the IT/ITES sector. To retain sustainable learning in businesses microlearning is the way.

Microlearning is also being used in across sectors, including healthcare, banking, and retail, to give workers instant access to knowledge and training. This is crucial in highly regulated businesses where workers must be knowledgeable about compliance and safety requirements.

The authors suggest that future work in this area can be done by scaling up the employee's responses. The use of gamification and AI in this area and clubbing technology with microlearning will be the next level to suit future learning spaces. It is also suggested that other industries will adopt it in line with IT/ ITES.

References

1. Skalka J, Drlik M, Benko L, Kapusta J, Rodriguez del Pino JC, Smyrnova-Trybulska E, Stolinska A, Svec P, Turcinek P (2021) A conceptual framework for programming skills development based on microlearning and automated source code evaluation in virtual learning environment. Sustainability 13(6):3293
2. Skalka J, Drlik M, Benko L, Kapusta J, Rodríguez del Pino JC, Smyrnova-Trybulska E, Turcinek P (2021) Conceptual framework for programming skills development based on microlearning and automated source code evaluation in virtual learning environment. Sustainability 2021(13):3293
3. Gaikwad P, Pandey S (2022) A review on special skill sets from industry 4.0 perspective. In: Proceedings—2022 2nd international conference on electronic and electrical engineering and intelligent system, ICE3IS, pp 276–281
4. Hug T (2021) Sound pedagogy practices for designing and implementing microlearning objects. In: Microlearning in the digital age. Routledge, pp 33–57
5. Dolasinski MJ, Reynolds J (2020) Microlearning: a new learning model. J Hosp Tour Res 44(3):551–561
6. Leong K, Sung A, Au D, Blanchard C (2021) A review of the trend of microlearning. J Work-Appl Manage 13(1):88–102
7. Andriotis N (2018) What is microlearning: a complete guide for beginners—elearning industry. www.elearning.com
8. Pandey A. Blog on elearning.adobe.com/2022/02/5-examples-of-microlearning-in-corporate-training-that-boost-learner-engagement-and-performance/
9. Ivins TZ (2008) Microfranchising microlearning centers: a sustainable model for expanding the right to education in developing countries. J Asynchronous Learn Netw 12(1):27–35
10. Romero-Rodríguez JM, Ramirez-Montoya MS, Glasserman-Morales LD, Ramos Navas-Parejo M (2023) Collaborative online international learning between Spain and Mexico: a microlearning experience to enhance creativity in complexity. Edu+ Training 65(2):340–354
11. Is microlearning good for business? Benefits and outcomes (getrapl.com)

Video Dehazing Using Dark Channel Prior and Type-2 Fuzzy Sets

Nisha S. Amin, Geeta Babusingh, G. G. Rajput, and R. L. Raibagkar

Abstract Dehazing refers to the methods designed for haze removal from the videos to increase visibility of objects. Dark Channel Prior (DCP) approach and haze veil analysis approach are two popular methods in the literature for haze removal. The DCP approach to remove haze from video frames is employed in his paper. The quality of the haze removed frames is increased by sharpening the edges of the objects in the image by using type-2 fuzzy sets. Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) parameters are used to test the effectiveness of the proposed approach. Video samples are gathered from the Kaggle dataset, with a particular emphasis on the vehicle objects in the scene. Results obtained after applying the proposed approach to several video frames show the efficacy of the method compared to other dehazing approaches.

Keywords Video dehazing · Dark channel prior · Type-2 fuzzy set · Canny edge detection · PSNR · SSIM

1 Introduction

The presence of haze or fog in images degrades its quality and reduces its visibility. Due to the existence of haze in the environment, videos shot in poor lighting conditions frequently have low contrast. Dehazing is the procedure used to remove haze from images and improve visibility of the image contents. To meet the growing need for applications in television, remote surveillance, intelligent transportation, remote sensing, satellite imaging, and underwater pictures, image processing for dehazing images and videos is a promising technology. Dark Channel Prior (DCP) is a popular method used to estimate haze in an image [1]. The method being fast and effective

N. S. Amin (✉) · G. G. Rajput
Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India
e-mail: skanda.nsa@gmail.com

G. Babusingh · R. L. Raibagkar
Department of Applied Electronics, Gulbarga University, Kalaburgi, Karnataka, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_47

can be applied to videos. A video, essentially made up of frames, is haze removed by applying the DCP approach and the quality of the videos is enhanced using image enhancement techniques. Different variations of the methods available in the literature include single scale dark channel prior, multi-scale dark channel prior, guided filter-based dark channel prior, and hybrid methods that combine the dark channel prior with other dehazing methods such as image fusion [2]. In this paper we employ dark channel prior approach for video dehazing, suggested by He et al. [1], and apply type-2 fuzzy set for enhancing the quality of dehazed video frames.

2 Literature Survey

A rapid execution technique for outdoor video dehazing in accordance with dark channel prior was suggested by Park et al. [3] in 2018. It achieves the rapid implementation of the dark channel prior approach aimed at the dehazing of outdoor videos. By maintaining the original method's dehazing quality, the overall execution time is decreased by up to 49%. In hazy image, every pixel in the image has a different level of haze density. Hence, finding the black pixel in an image clears out any haze. Camera records the blurry image and locates the airlight. A cost function made up of a term for the standard deviation and a term for the histogram uniformity is developed to assess the contrast. Lastly, testing results show that the suggested method can recreate the features in the original sceneries more clearly while also removing haze.

A unique quad-tree approach was put out by Luan et al. [4] to estimate the atmospheric light. At the video dehazing stage, a straightforward but effective parameter that specifies the similarity of the inter frame visual content maintains the ambient light in the same scene without changing it. This technique efficiently gets rid of unexpected flickers. Results from experiments demonstrate that the suggested technique significantly increased video dehazing efficiency and prevented halos and block effect.

A super pixel-based technique to eliminate haze from a single night-time haze image was put out by Yang [5]. The glow in the input night image is separated into glow and glow-free images by resolving a quadratic optimization issue. The night-time glow-free images are processed using the super pixel-based method. The estimation of ambient light and the transmission map are the two elements that need to be distinguished in the proposed approach. Using the SLIC technique, the night-time glow-free haze image is separated into super-pixels. Each super pixel's brightest pixel intensity is referred to as the super-pixel's ambient light. Using a preceding dark channel, the transmission map is approximated. The base layer is used to construct the transmission map. The WGIF is used to deconstruct the image and to remove morphological artefacts. The resulting transmission map is then applied with a threshold to prevent noise in the sky region after the haze has been eliminated. Yet the technique becomes more sophisticated as glow-free night-time haze photos are segmented into super-pixels.

An image and video dehazing using view-based cluster segmentation was proposed by Yu et al. [6]. In order to prevent distortion in sky areas and make the sky and white objects clear, a view-based cluster segmentation approach is used. Here, the distant view to estimate the sky region is changed to prevent distortion and the depth is first clustered using a Gaussian Mixture Model (GMM). Second, a single hazy image is split into K-classifications, and third, video dehazing is done using an online GMM cluster. This method, which is improved by view-based cluster segmentation to lessen colour distortion, is used to estimate atmospheric light and transmission independently for the depth map and haze image. Using an online GMM cluster, a video dehazing technique is described to enhance overall contrast. The overall brightness of the dehazed sequence image is raised, and ultimately a video dehazing technique is suggested that can restore hazy video while saving a lot of time for the cluster of dehazed sequence videos.

Fast execution technique for dark channel prior-based outdoor video dehazing was proposed [7] by Park et al. This enables the quick implementation of the outdoor video dehazing dark channel prior approach. While maintaining the same dehazing quality as the previous approach, the overall execution time is decreased by up to 49%. Dehazing is the signal processing technique used to eliminate haze. Every pixel in the picture has a different level of haze density. Therefore, finding the black pixel in a picture clears out any haze. Camera records the blurry image and locates the airlight. A cost function made up of a histogram uniform term and a standard deviation term is developed to assess contrast. Final experimental results show that the suggested method can recreate the features in the original scenes with greater clarity while also removing haze.

Dong et al., inquired about [8] in single outline infrared target recovery beneath terrible climate conditions particularly solid sea waves and oceanhaze. The particular creators changed the Visual Consideration Demonstrate (VAM) and coordinates with anti-vibration pipeline sifting calculation (a multi outline clutter expulsion strategy) to identify the frail targets seen on picture borders. A programmed extraction of Saliency Outline (SM) taken after by foundation clutter concealment might abdicate superior visual encounter than the existing comes about in several climate conditions, particularly for oceanic targets.

He et al. in [1] concluded that predicting the best transmission map just from hazy images could produce better results than applying the ground truth images. The essential stable properties of the initial transmission map were preserved by using a difference structure-preservation dictionary when it was altered.

A single image defogging approach using dark channel technique has been proposed by Wu et al. [9] to calculate the light in the atmosphere, which signifies the illumination present in the image. Using a transmission map and this ambient light, a fog-free picture reconstruction is performed. The transmission map demonstrates how fog affects an image's perspective. Four transmission maps are suggested in this work to recreate pictures with various chromatic contrasts. According to the proposed method, a transmission map is selected adaptively depending on the density of the

fog in order to reconstruct an image with an optimal colour contrast. As a result of the experiments, the proposed method is able to reconstruct images more qualitatively than the previous methods.

Archana [10] claims that depth information may be recovered using a depth map, colour attenuation prior characteristics can be employed in an adaptive linear model, scene radiance can be recovered using a tropospheric dispersion model, and haze can be accurately eliminated from an image. The depth map may be used to simply retrieve the original image. Even if a way was found to use the brightness and saturation of the fuzzy image to create the scene depth, the issue still remained. Because it remains constant under identical air conditions, the scattering coefficient cannot be adjusted. Since the existing single image dehazing techniques rely on ongoing research, a particularly adaptable model is needed.

According to Manjunath [11], a simple yet efficient way for dehazing a single image is to modify the detail method. Because of this algorithm's use of multiple scattering, each input image seems a little bit fuzzy. The single picture dehazing paradigm makes dehazing incredibly simple and effective. Since the algorithm is based on local content and has access to a wider range of images, it is more sensitive than a colour-based algorithm. Several physical models are applied to this issue. Dispersion caused by air layer particles, such as fog, haze, often has an impact on imaging in rainy weather.

Overall, the choice of the method depends on the specific application and the computational resources available. Dehazing is a challenging task and especially video dehazing is crucial since, the haze removal need to be done in real-time environment. In this paper, we present DCP and type2 fuzzy set-based approach for haze removal and enhancing the quality of the dehazed frames in the video.

3 Methodology

In haze affected images, usually the haze is non-uniform across an image making it difficult to estimate it and remove. Further, the lighting conditions affect the appearance of the haze in an image, making the task more difficult. Despite these, several effective dehazing methods have been proposed including the DCP methods and its variations. The suggested approach in this paper is based on the DCP. The method is based on the observation that in images captured outdoor, the pixels with lower intensity values correspond to non-sky regions containing objects such as trees and other objects. These pixels represent the presence of shadows which are not present in the sky region due to scattering of light by the atmosphere. The method involves computation of the dark channel based on these low intensity values. The estimated haze or depth is then subtracted from the original image. Haze removed image is then enhanced by applying type-2 fuzzy set approach. Figure 1 presents the suggested methodology's block diagram. Input to the method consists of hazy images; separation of the colour components is carried out and transmission map and atmospheric

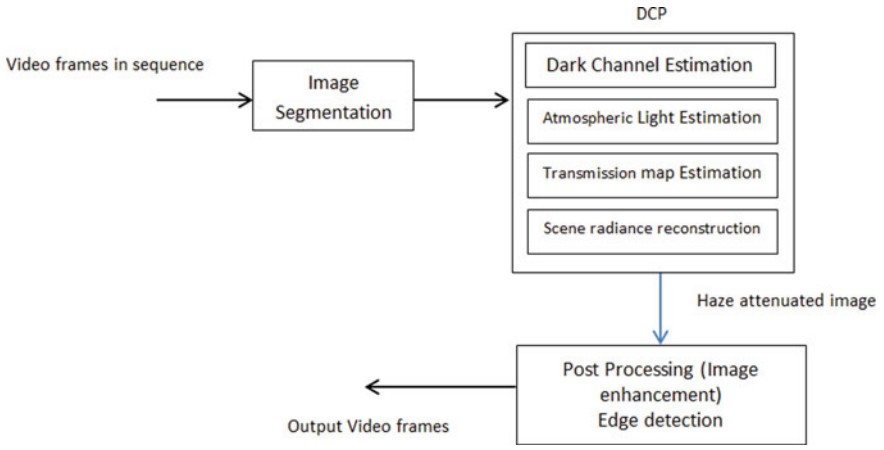


Fig. 1 Block diagram illustrating the suggested approach

lights are estimated from the components. Type-2 fuzzy set canny edge detection technique is used to improve the detection accuracy. Following is the detail presentation of the method.

3.1 Dark Channel Prior

The approach used to estimate the amount of haze present in an image is described below in terms of formulas used for computing the same.

The dark channel is computed as the minimum value of the pixel intensity across the RGB for every local window in the image. The dark channel of an image I at a pixel location x is defined as representing the window centred at pixel x.

$$\min_{y\in\omega(x)}\left(\min_{c\in\{r,g,b\}} I^{c}(y)\right), \omega(x) \tag{1}$$

where I^{c} is a colour channel (red, green, and blue) of I, J^{dark} is an approximated dark channel of I, and (x) is a local patch around the pixel at x. Similar to Eq. (2), J^{dark} is used to locate the pixels with the highest 0.1% values in order to determine the ambient light. These pixels are often opaque and fuzzy. Some haze removal approaches choose the pixel with the highest intensity to estimate the quantity of airlight, which is frequently incorrect (for instance, if there is an item with a greater intensity than the airlight). The top 0.1% of pixels' indices is kept in memory as follows:

$$m = \arg_{\tilde{x}} \max\left(J^{\text{dark}}(\tilde{x})\right) \tag{2}$$

$I(m)$ is then added together by adding the values of the three colour channels. The highest total is represented by the A pixel as follows:

$$A = \arg_{\tilde{I}} \max \left(\sum_{c \in \{r,g,b\}} \tilde{I}^{c}(m) \right) \tag{3}$$

The normalised dark channel is defined by using I and the estimated A as follows:

$$J_N^{\text{dark}}(x) = \min_{c \in \{r,g,b\}} \left(\frac{I^{c}(x)}{A^{c}} \right) \tag{4}$$

then it is presumed that ~ t is the coarse transmission

$$\tilde{t}(x) = 1 - \left(\omega \times \min_{y \in (x)} \left(J_N^{\text{dark}}(y) \right) \right) \tag{5}$$

Based on a dense field based on pixels, He et al. [5] solution uses a 0.95 aerial perspective factor, whereas Ke and Chen's [12] solution uses a moving average filter to smooth an incoherent, noisy transmission map. Both methods gives a nearly identical outcome, referred to in [5] as the refined (improved) gearbox map t. To retrieve J, the improved transmission t and A are applied as follows:

$$J(x) = \frac{I(x) - A}{\max(t(x), t_0)} + A \tag{6}$$

He et al. [5] when t_0 is set to 0.1, division by zero is avoided in places where the transmission map value is less than 0.1. Soft matting necessitates intensive calculations and memory, therefore Chen's [12] generated the dark channel by just choosing the value of the minimal colour channel for the pixel. Equations (2) and (3) were applied to estimate A, and Eqs. (1) and (5) in [13] were transformed into:

$$J^{\text{dark}}(x) = \min_{c \in \{r,g,b\}} (I^{c}(x)) \tag{7}$$

3.2 Atmospheric Light Estimation

The ambient light may be precisely calculated from the dark channel when a sizable local patch is employed to construct it. Employing a second dark channel with a bigger local patch size only for atmospheric light estimate is advised if the local patch size used to produce the dark channel is insufficient. It is demonstrated that using local entropy effectively improves estimation accuracy by allowing for the avoidance of ambient light estimation from bright objects (Fig. 2).

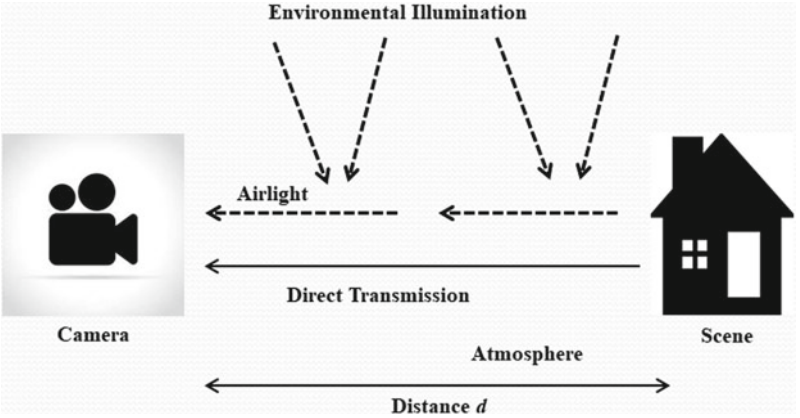


Fig. 2 Light reflection from atmospheric particles

The concepts that follow are used to describe the atmospheric light scattering model in the field of computer vision [14–16]:

$$I(p) = J(p)t(p) + A(1 - t(p)) \tag{8}$$

In this equation, (A) denotes ambient light, $R(p)$ indicates pixel p intensity, $J(p)$ indicates the output image, $I(p)$ indicates the input image, and $t(p)$ denotes the transmission map.

3.3 Type-2 Fuzzy Set

Fuzzy logic was developed to address issues with estimation that arise in decision-making. Many researchers used this function to assist them address various issues, and numerous implementations had already been made using the fuzzy concept. In the realm of fuzzy logic, an integral component of a collection or set has a degree of membership that is not always one or nothing, in conflict to what the concept of set theory predicts. The majority of fuzzy logic methods for identifying edges in images are built on the principle of dividing an image's grey levels range to three values [17]. The entropy maximum that could best characterise the image is represented by these three numbers. The fuzzy system's membership function is constructed using these values. To depict the conversion from and to crisp values, researchers employed several membership functions.

It is common to use the Gaussian membership function. However, triangle membership function is frequently used for practical purposes [13]. In order to get the most possible entropy during implementation, researchers develop their own fuzzy systems with this idea in mind.

The fuzzy system functions are used as a thresholding system in the edge detection issue to improve the ability to distinguish between neighbouring pixel points along edge lines or curves. In a type-2 fuzzy set, the membership function is defined as a fuzzy set with a range of possible membership values as appose to a single crisp membership value. The output membership functions will result in new ranges for the image's grey levels. In order to make the pixels' grey levels easier to detect, additional ranges have been added.

4 Experimental Results

Experiments are performed on hazy sample video frames obtained from the Kaggle dataset (https://www.kaggle.com/datasets/aalborguniversity/aau-rainsnow). Results of the suggested procedure on murky/hazy video frames are shown in Fig. 3. The suggested approach is implemented with a MATLAB R2020a (8.1.0.430) s/win 2.60 GHz Intel(R) Core(TM) i5-3230 CPU and 4 GB RAM system. Table 1 contains the computed execution times for example images. The finer borders of the images that were hidden in the hazy ones are retained in the rebuilt images. We note that the suggested method convincingly recovers the fine features of the objects while removing the haze present in the input images. The reconstructed output frame has less brightness.

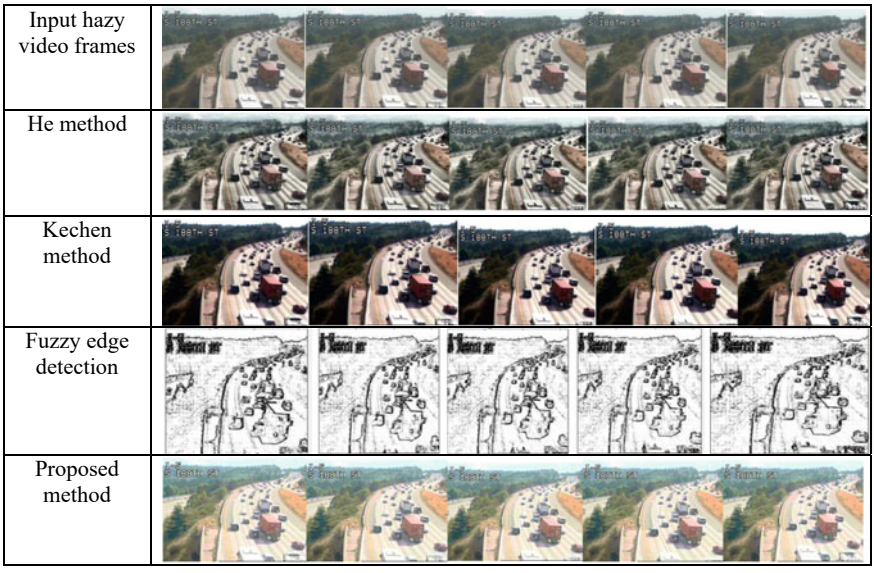


Fig. 3 Sample dehazed video frames using proposed method

Table 1 Comparison of proposed method with different methods

Method	Videos					
		1	2	3	4	5
He method	PSNR	20.18	20.74	19.48	18.59	20.30
	SSIM	0.83	0.80	0.77	0.81	0.79
Ke.Chen method	PSNR	13.12	12.84	12.17	11.63	12.33
	SSIM	0.70	0.68	0.65	0.71	0.66
Proposed algorithm	**PSNR**	**23.95**	**22.57**	**22.34**	**29.03**	**22.57**
	SSIM	**0.90**	**0.89**	**0.89**	**0.91**	**0.90**

To find edges in grayscale images, Canny's edge detection technique, which is widely used, is implemented. Insufficient illumination can cause the area boundaries in an illustration to become blurry, which causes ambiguity in the gradient image. Our technique for handling uncertainty uses type-2 fuzzy sets to automatically choose threshold values for segmenting gradient images using the traditional Canny's edge detection algorithm. The results show how our algorithm does a good job on video frames.

The outcomes are compared with outcomes from the literature using the techniques of He et al. (3) and Kechen (24). Figure 3 displays the outcomes for example video frames. Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) quality measures are used to compare the effectiveness of the approaches.

The PSNR is mostly utilised in image compression to rate the quality of the rebuilt picture. Typically, a greater PSNR would mean the reconstruction is of higher calibre. This equation is used to estimate PSNR.

Peak value is the highest value in the picture intensity values, and PSNR is defined as 10log10 (peakval2)/MSE. MSE stands for mean square error.

According to SSIM, picture degradation is defined as a shift in how structural information is perceived and highlights extremely interdependent or spatially confined pixels. The perceived quality of video frames is estimated by SSIM. SSIM $(x, y) = [l\ (x, y)]\ \alpha.[c\ (x, y)]\ \beta\ [s\ (x, y)]$ determines how similar the restored video frames are to the original video frames. Where I stands for brightness, C for contrast, S for structural information, and α, β, γ, are the positive constants [18]. In terms of PSNR and SSIM parameters, Table 1 compares the results of the suggested approaches with those of different techniques in the literature.

In comparison to the approaches in [5, 12], our algorithm obtains the fastest PC execution time which is shown in Table 2 and also the comparison of frame speed per second on different videos is shown in Table 3. Overall, in comparison with alternative methods, our technique gives improved appearances for removing haze in all circumstances (Fig. 4).

Table 2 Time comparison between several techniques

Time	Methods		
	He method	Ke.Chen method	Proposed algorithm
Video 1	13.848584	4.245117	**2.554671**
Video 2	15.393599	3.718119	**2.939792**
Video 3	16.150856	4.361065	**2.830060**
Video 4	45.211380	4.198889	**2.558005**
Video 5	16.081074	3.714851	**2.905553**

Table 3 Speed comparison between different techniques

Methods					
FPS	Video 1	Video 2	Video 3	Video 4	Video 5
He method	94.336986	81.978592	85.157217	28.160768	82.944635
Ke.Chen method	56.771108	64.817721	55.261733	57.396132	64.874745
Proposed algorithm	**17.402501**	**15.655858**	**14.921810**	**5.330516**	**14.986561**

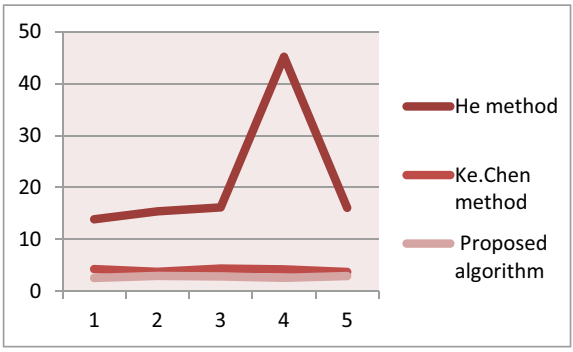


Fig. 4 Line graph of speed complexity of different methods

5 Conclusion

The importance of the haze reduction algorithms is growing rapidly. Due to pollution, mist, and haze in our surroundings, the videos captured suffer in visual appearance and make it difficult for other applications like object detection, it's crucial to eliminate the haze in-time for increasing its visual quality and assist in further processing. In order to get adequate and noise-free vision in applications like automobiles, the research has revealed the must-know aspects of fog removal algorithms. The traditional DCP approach along with type-2 fuzzy sets proposed in this paper has yielded excellent results. The efficiency of the method is described in terms of quantitative parameters namely, PSNR and SSIM.

.

References

1. Singh D, Kumar V (2017) Dehazing of remote sensing images using improved restoration model based dark channel prior. Imag Sci J 65(5):282–292
2. Chen SZ, Ren ZG, Lian QS (2016) Single image dehazing algorithm based on improved dark channel prior and guided filter. 42:455–465. https://doi.org/10.16383/j.aas.2016.c150212
3. Park Y, Kim T-H (2018) Fast execution schemes for dark-channel-prior-based outdoor video dehazing. IEEE Access 6:10003–10014. https://doi.org/10.1109/ACCESS.2018.2806378
4. Luan Z, Zeng H, Shang Y, Shao Z, Ding H (2018) Fast video dehazing using per-pixel minimum adjustment. Math Prob Eng 2018:8. Article ID 9241629. https://doi.org/10.1155/2018/9241629
5. Yang M, Liu J, Li Z (2018) Super-pixel based single nighttime image haze removal. In: IEEE transactions on multimedia
6. Yu F, Qing C, Xu X, Cai B (2016) Image and video dehazing using view-based cluster segmentation
7. Park Y, Kim TH (2018) Fast execution scheme for dark-channel-prior outdoor video dehazing. IEEE, Access 6:10003–10014
8. Dong L, Wang B, Zhao M, Xu W (2017) Robust infrared maritime target detection based on visual attention and spatiotemporal filtering. IEEE Trans Geosci Remote Sens 55(5):3037–3050
9. Wu X, Liu XY, Yuan F (2021) Fast dehazing method for improving the image quality in pellet size measurement. Comput Ind 128:103438. ISSN 0166-3615, https://doi.org/10.1016/j.compind.2021.103438
10. SArchana ME, Abiraha A (2016) An effective haze removal algorithm using color attenuation prior model. In: International conference on emerging engineering trends and science issue no 8, pp 387–392
11. Manjunath V, Phatate R (2016) A single image haze removal algorithm using color attenuation prior. Int J Sci Res Publ 6:291–297
12. Biswas R, Sil J (2012) An improved canny edge detection algorithm based on type-2 fuzzy sets. Procedia Technol 4:820–824. ISSN 2212-0173. https://doi.org/10.1016/j.protcy.2012.05.134
13. Alavala CR (2008) Fuzzy logic and neural networks basic concepts and applications. New Ege International Limited Publications
14. Kim W, Lee R, Park M, Lee S (2019) Low-light image enhancement based on maximal diffusion values. IEEE Access 7:129150–129163
15. Lee H, Sohn K, Min D (2020) Unsupervised low-light image enhancement using bright channel prior. IEEE Signal Process Lett 27:251–255
16. Ke N, Chen J (2013) Real-time visibility restoration from a single image. In: Proceedings on IEEE international conference on image processing, pp 923–927
17. Zhao M, Fu AM, Yan H (2001) A technique of three-level thresholding based on probability partition and fuzzy 3-partition. Fuzzy Syst IEEE Trans 469–479
18. Dong T et al (2019) Efficient traffic video dehazing using adaptive dark channel prior and Spatial–temporal correlations. Sensors 19(7):1593
19. Lee S, Yun S, Nam JH et al (2016) A review on dark channel prior based image dehazing algorithms. J Image Video Proc. 2016:4. https://doi.org/10.1186/s13640-016-0104-y
20. Gupta N, Jha RK, Mohanty SK (2014) Enhancement of dark images using dynamic stochastic resonance in combined DWT and DCT domain
21. Singh T (2017) Foggy image enhancement and object identification by extended maxima algorithm
22. Shi J, Yang KJ (2017) An improved method of removing fog and haze effect from images. In: 4th international conference on machinery, materials and information technology applications (ICMMITA 2016) Copyright © 2017. Published by Atlantis Press
23. Guo F, Tang J, Cai ZX (2014) Image dehazing based on haziness analysis. Int J Autom Comput 11(1):78–86. https://doi.org/10.1007/s11633014-0768-7
24. Gibson KB, Vo DT, Nguyen TQ (2012) An investigation of dehazing effects on image and video coding. IEEE Trans Image Process 21(2):662–673

25. Shi Z, Zhu MM, Guo B, Zhao M, Zhang C (2018) Nighttime low illumination image enhancement with single image using bright/dark channel prior. EURASIP J Image Video Process 13:1–15
26. Eshwarappa L, Rajput GG (2022) Type-2 fuzzy sets-canny edge detection algorithm-based text extraction from complex video scene. Harbin Gongye Daxu Xuebao/J Harbin Inst Technol 54(4):76–83

Systematic Review on Growth of E-Agriculture in Context of Android-Based Mobile Applications

Vidya Kumbhar, Ashwini Patil, Sneha Kumari, and Nisha Bharti

Abstract About one-third population of India is dependent directly or indirectly on agriculture or allied activities. India is a leading producer of several commodities, but still, we lag behind in terms of agriculture productivity. There is enough potential to enhance the production of agricultural commodities. There are several constraints that hinder the process of enhancing agriculture production. One of the key factors is the lack of availability as well as the use of information technology in agriculture. Farmers need user-friendly technology. In this view, this paper is an attempt to review the available application for effective agriculture management through e-agriculture in the context of Android-based mobile applications. The findings of the paper revealed that there are several e-agriculture applications for pest and disease management, weather advisory, Fertilizer management, etc. This shows that mobile application for agriculture provides growth opportunities for e-agriculture. The findings also revealed that there is a need for an app that can provide holistic services to farmers. This kind of technology will help farmers in managing agriculture practices more effectively through e agriculture.

Keywords e-Agriculture · Android · Mobile phone · Management

V. Kumbhar (✉) · A. Patil
Symbiosis Institute of Geoinformatics (SIG), Symbiosis International (Deemed University), Pune, India
e-mail: kumbharvidya@gmail.com

S. Kumari
Symbiosis School of Economics, Symbiosis International (Deemed University), Pune, India

N. Bharti
Symbiosis Institute of International Business, Symbiosis International (Deemed University), Pune, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_48

1 Introduction

Agriculture is the mainstay of the Indian economy. Out of total population, 54.6% population depends on agriculture and allied sectors [1]. In terms of area, India is the seventh largest country in the world. Despite its large area, only 43% of it lies under cultivation. As per the data from FAO (2020), the country has the largest production of agricultural commodities such as pulses, jute and second largest producer of wheat, sugarcane, vegetables, soil seeds such as groundnut, cotton, and fruits (FAO, 2020). India is also the second largest fruit producer in the world [2]. However, it is important to note that the cropping intensity of India is reported at 143%, and only 46% of its area is under irrigation [3]. That means, only a small portion of our land is under repeated sowing. One of the key constraints in increasing the cropping intensity is the lack of irrigation facilities and dependency on monsoon.

Post-green revolution, food production in India has increased, but the report suggests that India still has 25% of the total world's hungry population and has around 190 million people who are undernourished. Another important concern is the decreasing contribution of agriculture to total GDP. The data suggest that the contribution of agriculture to GDP has decreased from 50% in the 1950s to 15.4% in 2015–16 (at constant prices). The level of mechanization is also quite low as compared to many developed countries.

Artificial intelligence market is growing very fast. As per an estimate by IDC (2017), the investment in this sector is growing at a rate of 50.1% of the compound annual growth rate (CAGR) at global level and it is expected to reach up to reach USD57.6 billion in 2021. Recognizing the importance of Artificial intelligence (AI), NITI Aayog floated a national strategy in the year 2018, which was meant to identify the Goals and Priority sectors keeping agriculture research in the priority sector [3, 41]. The objective was to enhance farmer's income, increase farm's productivity, and reduce the wastage. Hence, this study tries to review the existing applications in this area and identify the research gaps for improving the applications.

2 Need of Mobile Application Development in Agriculture

From agriculture statistics of the year 2014, India is the world's largest producer of fresh fruits and the world's largest five producers of over 80% of agricultural-produced items. Due to lack of technology exposure, farmers lack information about the rates of the crops and yields. Hence, farmers sell their crops at low rates resulting in loss. Thus, they take loans from banks and other sources. Mobiles require only basic skills for handling. Through the SMS-based service, the information is sent in textual format limited to 160 characters, so it requires support to send the complete meaning of the information. SMS-based service study done in Kalburgi district KVK of Karnataka conducted by Sidramappa et al. (2013) reveals that farmers mainly depend on SMSs related to weather forecasting (83%), integrated pest management

(77%), disease management (72%), new varieties and seeds (68%), livestock management (65%), and organic practices (45%). Some recent apps are available in local languages, in audio, video format for easy access by farmers.

3 Review on Mobile Applications in agriculture

3.1 Mobile Applications for Crop Advisory

Crop advisory systems help farmers in different agriculture management activities such as seeding, soil management, fertilizer, irrigation, and other farm-related activities. "Padi2U" is a mobile application developed for paddy crop. The application provides advice to the famers in local Malay language. The data available in an application are on drone images, seeds information, paddy variety, diseases, fertilizers, and pests information [4]. Java-based Android application developed provides an advice to improve the crop yield focusing on soil water availability, average temperature, the soil properties such as the soil alkalinity, soil type, and crop growth cycle. The Artificial Neural Network-based decision support system helped farms with decision-making and using ANN with statistical analysis helped for decision-making [5, 6]. "Krushi Samriddhi" is an Android application developed using Eclipse software for decision support system in agriculture for farmers to get high crop advice on land details, soil details, crop history, crop planned, fertilizer used, pesticides, Harvesting, Pest and disease, etc. and guidance to the farmer [7]. "eAgro" is Android and SQLite-based application developed for farmer to advice crop. This application will receive the user landscape information through the Global Positing System (GPS) tracking and will suggest them the most suitable crop according to the given information [8]. "AgroMobile" is the Android-based application for crop suitability and weather forecasting [9]. "Agro-vision" is another Android-based application which provides information about crop protection and helps farmers for crop selection, and step-by-step guidance gives to the farmer through the crop video [10].

"AgriMaps" is another Android-based application which uses the open street map as map resource and uses GPS locator for identifying users' location. The application shows the center of map and land profiles including soil features with 500 m radius of the farm [11]. Android-based mobile application designed for advisory system for Thai rice plantation provides information about history of Thai rice, preparation of land for cultivation, and the advisory required at different stages of crop growth cycle. The application provides the timely advice and alerts for the crop protection [12]. "MahaFarm" is an another Android-based mobile application for farmers which comprehensively help them in different agriculture activities such as agro information, weather forecasts, daily market prices and agriculture newsletters, Kisan loans, and other government-related information updates [13]. "Pusa Krishi" is an Android application that gives information about new varieties of crops developed by (ICAR)

Indian Council of Agriculture Research and also gives information about resource-conserving cultivation practices, farm machinery, and their implementation to the farmers which helped farmers for crop production improvement [14].

3.2 Mobile Application for Crop Disease Detection

Crop disease identification and management are an important activity in agriculture management. "E-agree" is the Android application which provides the soil formation to increase the crop yield and using image processing technique gives the disease detection information on the basis of leaf image [15]. "AgroMobile" is the Android application developed particularly for the Indian farmers and help them in their farming needs. This app is available in multiple languages and it gives solutions to various problems like detection of pest of diseases in plants [31]. "AgroApp" is developed using as a programming language Java and Eclipse Juno integrated with ADT plugin-based application which provides information about type of climate, crop disease and provides solution for that [1]. Rattanachai et al. [12] have developed mobile application for rice crop disease identification and treatment for the Thai farmers of central zone of Thailand. "Mentha Mitra" is a Java and xml-based Android SDK bilingual (Hindi and English) application for menthol mint crop. This application provides scientific advisory to the farmers and help them by suggesting high yielding varieties and identification and control of disease and pests of mint [16]. "Plantix" is the mobile application developed by PEAT for plant disease diagnostics and monitoring. The features of plantix are real-time diagnosis, weather information system, community feature facilities' interactions, and smallholder farmers as end-users [17].

3.3 Mobile Application for Agriculture Marketing

The existing applications for agriculture marketing provide only basic information to the farmer such as 7/12 Extract, weather forecast, minimum cost price, agriculture related news and developments, Krushi Offices, Fertilizers, Pesticides, Insecticides, and even enquiries about the particular information [18, 19]. "Eagree" is the market rate guide for farmer to communicate about market rate across multiple markets of geographically distributed area with the help of GPRS [12, 14]. The application is developed using MaSMT and MYSQL-based Android application for communication among the agriculture stakeholders as farmer, buyer, seller, and instructor [20]. For building an Android application, researcher has used Android, SQLite, and Eclipse-based methodology where vegetable prices are updated periodically [10]. "Krishi Ville" is another Android-based application which helps to provide the updates on weather forecast, agricultural news, etc. for crops [21]. "Farmer Helping

Service" is Android-based mobile application developed for horticulture crops available in Gujrati language [22, 23]. Agro App is developed for farmers as well as common man and it helps in various decision making like which crop to grow in which season and also help in dealing with various pest and diseases on those crops. This mobile application also provides diversified videos of agriculture work [24]. "MAHAAGRI SMS" is a Government of India project. This application provides the link between farmers and different mandi. The represented five farmers are selected as "Krishi Mitra" and they represent the department in the village across the country. Out of total crops, 43 major crops were also identified to provide the crop advisory [25]. "IFFCO Kisan" is also an Android-based application which provides the knowledge database of different agriculture advisories as well as the latest mandi prices. The information about weather forecast and various tips on agriculture management activities are also provided through the application [25].

3.4 Mobile Application for Farm Monitoring

The mobile applications play important role to manage different activities under farm monitoring. "eFarm" is the Android application which makes under the Online MySQL, Microcontroller, and Webcam methodology. It helps in developing an automated system for irrigation as well as providing other inputs like pesticide in the farm. It helps farmers to schedule the daily farming activities and get reminders to farmer about their scheduled activity [18]. Performed field monitoring they used GPS and RFID tagging methodology, it uses web GIS systems for the GPS information. In this app, neural network and artificial intelligence are used to recommend decisions [26]. The feature of the "FarmManager" app is Annual structure, Farm customization, easy field management, Land field data, Physical field GPS addressing GGRS87, Easy job recording process, Employees and equipment. The methodology used in the Android application is GPS and Google Map live [27]. The administration of little homesteads, planned and created to prerequisites of the ranchers of Greece. It can store database, do cultivate customization, simple field the board, land field information, simple occupation recording procedure, representatives, and gear. Mohanraj et al., Rajalakshmi and Devi Mahalakshmi [28, 29] have developed the Android application for the farm monitoring and control of the farm. The software technology such as PHP and JSON format data storage was used in the application.

3.5 Mobile Application for Fertilizer Recommendations

"AgroNutrition" is a GPRS-based Android application designed for farmers to compute the available nutrient amount from the fertilizers which have applied to be on plants [30]. It consists of highly authentic and reliable database on agriculture [31]. "AgroApp" provides information to the farmers according to the type of soil

and suggest which fertilizer suits for which crop. App performs weather forecast for a specific period of time. Farmers can frequently ask questions to the expertise about farming activities [24]. Applications developed also suggest or provide element of specific farm with the limit value and recommendation for pest prevention and fertilizer based on observation of symptoms [32]. "Ecofert" is an Android application that has developed for solving the problem about reaching an ideal solution in what quantity and which fertilizers should be added to the fertilizer tanks of an irrigation system and minimizing the total cost of fertilizer. Ecofert models the fertilization selection as a Linear Programming (LP) problem and uses the Simplex algorithm for solving the fertilizer selection [33].

3.6 Mobile Application for Weed Management

Weeds are unwanted and undesirable plants which interfere with the utilization of land. Regional Agricultural Research Station, Tirupati, Andhra Pradesh, developed an Android application which includes the entire package of agriculture practices that are required for scientific cultivation of groundnut [34]. "Take Action" is the Android and iOS-based application which is designed to help farmers manage herbicide, fungicide, and insect resistance, it also provides an option to search suitable herbicide brand, and the goal of this application is to encourage farmers to adopt management practices that lessen the impacts of resistant pests and preserve current and future crop protection technology (Weed Science Society of America). University of Missouri USA developed "ID Weeds" application helps to identify weeds through images and the remedial measures [35]. "Kisan Yojana" application helps to bridge the gap between government and farmers. This application provides the timely information about the schemes and weed management studies, research.

3.7 Android Application to Connect with Farm Advisor

"mKRISHITM" is an agro advisory system based on wireless and sensor technology and built by Tata Consultancy Services available in regional languages too [36]. "Kheti" is an another mobile application which is designed to bridge the gap between the advisors and farmers. It was tested by the Sironj Crop Producer Company Ltd (SCPCL), a cooperative of small farmers. This application is consists of three components i.e. interactive voice recorder system (for recording of telephone conversation between farmer and advisor), a short dialogue strips (consists of 6 photographs) and an audio track. This Short Dialogue strip (SDS) is sent to a server and the advisors can see them and suggest the solutions [37]. Another android application is Krishi Video Advice tried to bridge the gap between farmers and advisors by providing video call facility for resolving the issues of farms. It also has video conference call

facility to contact the experts at the Kisan Call Center (KCC) to get the advice related to farm issues [38].

3.8 Mobile Application for Specific Crop

National Rice Research Institute (NRRI), Cuttack, developed a riceXpert bilingual Android application in Odia and English language for rice farmers to reach the latest rice technologies in real time. It has features like rice varieties, news, expert consultation, agricultural implements, weather information. Other feature of this application is customised solutions for rice crops. Farmers can send photos, audios and get quick solutions to their problems [39]. To provide detailed information to the farmers about groundnut cultivation Regional Agricultural Research Station, Tirupati, developed a Mana Verusanaga application, it is in offline mode and in Telugu language. To use this application, the farmers need to be registered and this registration can be done by sending even an SMS. The content of this application is related to suggestions on variety selection, seeds, pest and disease, nutrient management, fertiliser application etc. For cane growers and millers ICAR-Sugarcane Breeding Institute, Coimbatore developed a Cane Adviser mobile application for query handling. ICAR-National Research Centre on Pomegranate (NRCP), Solapur, has developed an Android application for pomegranate which includes publications for download from ICAR-NRCP website, updates of the weather forecast of the region, daily market rates of pomegranate in India, announcement from NRCP about upcoming programs such as seminar, training available in Hindi, Marathi, and English language [40].

4 Conclusion

Government has realized the importance of technology particularly Information and communication technology (ICT) in improving the productivity of agriculture produce. With the increased depth of mobile users even in rural users, mobile-based apps have become one of the key tools in bridging this knowledge gap for the farmers. In view of this, various government as well as non-government organizations have started working on development of mobile-based apps for supporting e-agriculture services. This paper revealed that there has been development of apps in various areas, but there is lack of an integrated application which provides holistic services at one platform. There is a need to develop a holistic service provider application that will be user-friendly for the farmers seeking all information in one place. This holistic app will be able to help in optimisation of inputs, thereby reducing the cost and improving the productivity and development of digitalisation in agriculture.

References

1. Department At A Glance (nd) Department of Agriculture, India. Retrieved August 30, 2020, from http://agricoop.nic.in/department-glance
2. Agriculture in India: information about Indian agriculture & its importance (2020, July) India Brand Equity Foundation. https://www.ibef.org/industry/agriculture-india.aspx
3. Annual Report (2019, March) Department of Agriculture, Cooperation, and Farmer's Welfare. http://agricoop.nic.in/annual-report
4. Athirah RN, Norasma CYN, Ismail MR (2020) Development of an android application for smart farming in crop management. IOP Conf Ser Earth Environ Sci 540:012074. https://doi.org/10.1088/1755-1315/540/1/012074
5. Koli RP, Jadhav VD (2015) Agriculture decision support system as android application. Int J Sci Res 4(4):903–906
6. Patel V, Thakkar R, Radadiya B (2014) An android application for farmers to disseminate horticulture information. Int J Comput Appl 88. https://doi.org/10.5120/15337-3671
7. Patil TR, Shamshuddin K, Patil R, Sadanand P (2016) Krushi Samriddhi: a decision support system for farmers to get high crop yield. In: 2016 international conference on computational techniques in information and communication technologies (ICCTICT). IEEE, pp 35–40
8. Patodkar V, Simant S, ShubhamSharma CS, Godse S (2015) E-agro android application (integrated farming management systems). Int J Eng Res General Sci 3(3):458–469
9. Prasad S, Peddoju SK, Ghosh D (2013) AgroMobile: a cloud-based framework for agriculturists on mobile platform. Int J Adv Sci Technol 59:41–52
10. Jagtap R, Patil V (2017) Agro-vision: android based solution for regional farming. Int J Technol Res Eng 4(11):2363–2367. http://www.ijtre.com/images/scripts/2017041125.pdf
11. Jordan R, Eudoxie G, Maharaj K, Belfon R, Bernard M (2016) AgriMaps: improving site-specific land management through mobile maps. Comput Electron Agric 123:292–296. https://doi.org/10.1016/j.compag.2016.02.009
12. Rattanachai R, Sreekaewin P, Sittichailapa T (2015) Development of Thai rice implantation recommend system based on android operating system. Procedia Soc Behav Sci 197:1048–1052
13. Bhave A, Joshi R, Fernandes R (2014) Mahafarm-an android based solution for remunerative agriculture. Int J Res Advent Technol 2(4):43–47
14. Mobile Apps Kisan Suvidha and Pusa Krishi launched (n.d.) Department of Agriculture Cooperation & Farmer's Welfare. Retrieved July 2020, from http://agricoop.nic.in/recentinitiatives/mobile-apps-kisan-suvidha-and-pusa-krishi-launched
15. Reddy S, Pawar A, Rasane S, Kadam S (2015) A survey on crop disease detection and prevention using android application. Int J Innov Sci Eng Technol 2(4):621–626
16. Singh PP, Pandey P, Singh D, Singh S, Khan MS, Semwal M (2020) 'Mentha Mitra'—an android app based advisory digital tool for menthol mint farmers. Ind Crops Prod 144:112047
17. 'Plantix' Now In Kannada—Smartphone App Helps Farmers Detect Crop Damage (2018). https://Www.Icrisat.Org/. https://www.icrisat.org/plantix-now-in-kannada-smartphone-app-helps-farmers-detect-crop-damage/
18. Mukesh C, Sumeet D, Akhsay J, Chetan D (2015) Scheduling controlling and monitoring of agricultural devices using android application. Int J Adv Res Comput Eng Technol (IJARCET) 4(4):1–4
19. Kadam P, Khairnar G, Nikam P, Yuvaraj G (2016) AGRONOMY-an android application regarding farmer utility. J Emerg Technol Innov Res (JETIR) 3(4)
20. Nagaraj S, Gunasekar K (2018) Online vegetables pricing system in android. Innovation 5(6):142–146. https://www.rsisinternational.org/virtual-library/papers/online-vegetables-pricing-system-in-android/
21. Singhal M, Verma K, Shukla A (2011) Krishi Ville—android based solution for Indian agriculture. In 2011 Fifth IEEE international conference on advanced telecommunication systems and networks (ANTS). IEEE, pp 1–5
22. Patel H, Patel D (2016) Survey of android apps for agriculture sector. Int J Inf Sci Tech 6(1–2):61–67

23. Radhika D, Dipali B, Prasad M, Pandhare S, Akshay N (2015) One stop solution for farmer consumer interaction. Int J Comput Appl 975:8887
24. Aggarwal M, Kaushik A, Sengar A, Gangwar A, Singh A, Raj V (2014) Agro app: an application for healthy living. In: 2014 international conference on information systems and computer networks (ISCON)
25. शेती संबंधित मोबाईल ॲपचा वापर करणेबाबत… (n.d.) Department of Agriculture Maharashtra. Retrieved August 30, 2020, from http://krishi.maharashtra.gov.in/1231/Mobile-App
26. Liopa-Tsakalidi A, Tsolis D, Barouchas P, Chantzi A-E, Koulopoulos A, Malamos N (2013) Application of mobile technologies through an integrated management system for agricultural production. Procedia Technol 8:165–170. https://doi.org/10.1016/j.protcy.2013.11.023
27. Lantzos T, Koykoyris G, Salampasis M (2013) FarmManager: an android application for the management of small farms. Procedia Technol 8:587–592. https://doi.org/10.1016/j.protcy.2013.11.084
28. Mohanraj I, Ashokumar K, Naren J (2016) Field Monitoring and automation using IOT in agriculture domain. Procedia Comput Sci 93:931–939. https://doi.org/10.1016/j.procs.2016.07.275
29. Rajalakshmi P, Devi Mahalakshmi S (2016) IOT based crop-field monitoring and irrigation automation. In: 2016 10th international conference on intelligent systems and control (ISCO)
30. Srija S, Geetha C, Lavanya R, Dr. Kalpana M (2016) AgroNutri android application. Int J Adv Res Comput Commun Eng 5(5)
31. Jadhav M, Jagtap V (2019) Android application for farmers. Int Res J Eng Technol (IRJET) 6(4). e-ISSN: 2395-0056
32. Kumar V, Dave V, Nagrani R, Chaudhary S, Bhise M (2013) Crop cultivation information system on mobile devices. Int J Sci Res (IJSR)
33. Bueno-Delgado MV, Molina-Martínez JM, Correoso-Campillo R, Pavón-Mariño P (2016) Ecofert: an android application for the optimization of fertilizer cost in fertigation. Comput Electron Agric 121:32–42. https://doi.org/10.1016/j.compag.2015.11.006
34. Mohan KI (2019) Development of android mobile application on groundnut crop cultivation, Regional Agricultural Research Station, Tirupati, Andhra Pradesh
35. Weed Science Society of America (n.d.) https://Iwilltakeaction.Com/. Retrieved July 2020, from https://iwilltakeaction.com/resources/weed-science-society-of-america
36. Pande A (2013) m-KRISHI market facing innovation. TATA Consultancy Services TCS Innovation Lab, Mumbai
37. Dearden A, Matthews P, Rizvi H (2010) Kheti: mobile multimedia in an agricultural cooperative. Pers Ubiquit Comput 15(6):597–607. https://doi.org/10.1007/s00779-010-0335-3
38. Mandi K, Patnaik NM (2019) Mobile apps in agriculture and allied sector: an extended arm for farmers
39. Dan L, Hongli W, Mengya Z, Jianqiu X (2017) Intelligent agriculture greenhouse environment monitoring system based on the android platform. In: 2017 International conference on smart grid and electrical automation (ICSGEA). IEEE, pp 358–361
40. ICAR-NRRI developed Mobile app "riceXpert" (n.d.) ICAR National Rice Research Institute Cuttak. Retrieved August 2020, from https://icar-nrri.in/icar-nrri-developed-mobile-app-ricexpert/
41. Niti Aayog, Annual Report (2019). https://Niti.Gov.In. https://niti.gov.in/annual-reports

Impact of Network Security Enhancement on Work from Home

Anuroop Sundd and Mahipal Singh Deora

Abstract With the advancement of technologies new trends are introducing in the market to accomplish the industrial task in efficient and convenient manner. The work from home culture is one of the significant cultures introduced in the pandemic situation and it was too successful model during the time and also in the present time. The crucial section of the work from is support of network and communication technology with security considerations. The network security is one of the important aspect associated with, to overcome this situation there are multiple options are taking in the account similarly one of the idea about to train the employee with advance network training remain exited. The idea is typical to implement as this is concern from the technology learning, still lot of positive opinions are there to support this action. So it is projected that the advance network training is significant action to prevent cyber and security attack for the duration of work from home. The present research paper is a comprehensive study with the opinion of respondents with statistical data analysis.

Keywords Network · Security · Work from home · Technologies · Cyber security · Training

1 Introduction

Working from home has become increasingly common, and it's important to take steps to secure your home network to protect your work-related data and systems. There is need of some guidelines for improving your home network security for work from home [2, 3].

To secure our home network for WFH use a strong and unique password for your Wi-Fi network. This will prevent unauthorized access to your network and reduce the risk of hacking attempts [7, 9, 10]. Enable network encryption such as WPA2 or WPA3 to secure your Wi-Fi network. Encryption will scramble the data transmitted

A. Sundd (✉) · M. S. Deora
Department of Computer Science, B. N. University, Udaipur, India
e-mail: anuroopsundd@gmail.com

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_49

between your devices and the router, making it difficult for anyone to intercept or steal data. Update the firmware of your router to the latest version to ensure that it has the latest security patches and fixes. This will help to reduce the risk of any vulnerabilities being exploited. Disable remote management of your router to prevent unauthorized access to your network. This will prevent hackers from accessing your router from outside your home network [8, 11]. Use a virtual private network (VPN) to secure your Internet connection. A VPN will encrypt your Internet traffic and route it through a secure server, making it difficult for anyone to intercept or access your data. Install antivirus and anti-malware software on all devices on your home network to protect against viruses, malware, and other threats. Ensure that the software is updated regularly to protect against the latest threats [12–14].

Use a separate guest network for visitors or anyone else who needs to use your internet connection. This will prevent them from accessing your work-related devices and data. Disable any unused ports on your router to prevent unauthorized access to your network. Enable two-factor authentication (2FA) for all devices and accounts that support it. This will add an extra layer of security to your accounts and devices. Regularly backup your work-related data and store it in a secure location. This will ensure that you can recover your data in case of any security breaches or data loss. By implementing above-mentioned home network security measures, we may able to reduce the risk of security breaches and protect home work-related data and systems while working from home [4, 6].

2 Research Outline

The study involved collecting data through a survey questionnaire and various data recording methods and statistical analysis. The examination and accepting of the data are crucial aspects of the study, and this paper presents the same. The survey questionnaire used for data collection was designed to address the study's objectives and hypotheses. The data collected for the analysis comprises the perspectives of employees, network experts, and administrative staff actively involved in network security management for work from home activities. This paper is structured according to the research methodology's objectives and hypothesis framework, aiming to present experimental opinions from the respondents' data obtained from the questionnaire. The goal of this paper is to contribute to examine the status and awareness about network security and insight into the network security training approach to assess the impact of a secure work from home culture during normal days.

2.1 Objective

The objective of the study is to identify the significance of the advance training of network to prevent cyber security attack during work from home.

The objective is basically an attempt to accomplish the real fact about that advance network training is enough to prevent the network or not. Definitely this is a single site of the security of network related activities.

2.2 Hypothesis

Null Hypothesis: The advance network training is significant action to prevent cyber and security attack for the duration of work from home.

Alternate Hypothesis: The advance network training is not a significant action to prevent cyber and security attack for the duration of work from home.

2.3 Methodology

The current assumption emphasizes the importance of advanced network security training in facilitating a work from home environment. The success of working from home depends on trained employees who can effectively manage network security.

It is evident that employees who receive advanced training will be better equipped to handle network security, thereby preventing cyber and security attacks during work from home activities. The study aimed to test this hypothesis and collected data through opinion-based questions, including the following:

- Advanced network management training will be crucial for employees before being allowed to work from home.
- This training will play a significant role in preventing cyber-attacks during work from home activities.

To test this hypothesis, the study collected opinions from respondents through the above-mentioned queries. The data collected as the opinion of the respondents were stored by utilizing the 5-point-likert-scale. In this scale 05 opinion as the answer of the question were given, respondents may choose one of the opinion where he may get himself satisfied with the questions (Table 1).

The collected values are presented in the table below. Given the nature of the data/variables and the sample size, the appropriate statistical test to apply in this case is the "z" test.

The critical value of "z" on 0.05 level is $= 1.64$.

Table 1 Five-point-likert-scale for study

Opinion	Fully satisfy	Satisfy	Moderate	Not satisfy	Strongly not satisfy
Scale value	5	4	3	2	1

To evaluate the hypothesis, the study considered 300 responses from the respondents for each variable/query. Given the sample size/opinion data, the most suitable statistical test for this case was found to be the "z" test. Accordingly, the test was performed, and the subsequent results were obtained.

3 Result Analysis Summary

Given the sample size/opinion data, the most suitable statistical test for this case was found to be the "z" test. Accordingly, the test was performed, and the subsequent results were obtained.

The Tables 2 and 3 demonstrations the essentials of "z" test which is well-designed on the basis of the opinion received from the respondents on the 5-point-likert-scale. To inspect the implication, here the critical one tail value of the "z" test on the 0.05 level of significance is taken in the consideration as the bench mark and the value is 1.644.

However, the calculated value of "z" is only 0.775, which is lower than the one tail critical value of "z" at a 0.05 significance level. Therefore, based on the values, the null hypothesis is accepted, and the alternative hypothesis is rejected. Consequently, it can be observed that advance training is a significant factor in preventing cyber and security attacks during work from home.

Based on the central tendency values, there is no major difference in the mean values of the variables. According to Fig. 1, the opinions of respondents on the significance of advance training in preventing cyber and security attacks during work from home were measured. The certainty is reflected in the form of average

Table 2 "z" test between variables

Question/variables	N	Var	z	Result
Advanced network management training will be crucial for employees before being allowed to work from home	**300**	**1.20**	0.775	Significant
This training will play a significant role in preventing cyber-attacks during work from home activities	**300**	**1.25**		

Table 3 Summary of data analysis

Result	Values
Calculated z value	0.775
One tail critical value	1.64
Two tail critical value	1.96
Significance level of significance	0.05 or 5%

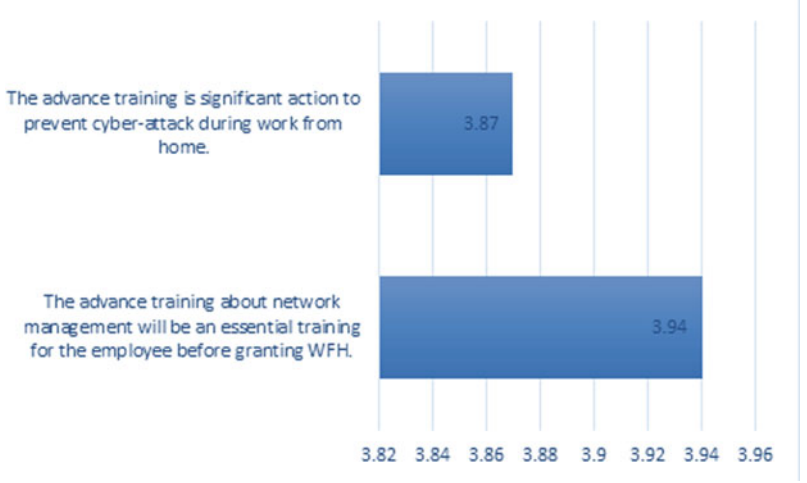


Fig. 1 Central tendency of both the question

perpetrator values, and it can be seen that there is similarity in the opinions. It is also evident that the hypothesis test is satisfied, as there is not much variation in the opinions of the respondents regarding these queries. In conclusion, the statement about the mean values of the opinions of the respondents supports the claim that there is a need for advance training to prevent any kind of cyber-crime and network security.

4 Conclusion

According to hypothesis the study is an attempt is to know that whether the advance network training is significant action to prevent cyber and security attack for the duration of work from home or not. From the opinion collection with required queries and by applying appropriate statistical test it is found that null hypothesis is accepted which shows that advance network training is significant action to prevent cyber and security attack for the duration of work from home. The study is carried out on the small scale but the as per the result is clear that at the very first stage employee who are working through work from home mode or on the virtual mode should be get network security training as an essential component of work from home culture.

References

1. ABS (Australian Bureau of Statistics) (2019) Disability, ageing and carers. Summary of Findings, Australia
2. Barrero J, Bloom N, Davis S (2021) Why working from home will stick. Working Paper, Department of Economics, Stanford University
3. Barrero J, Bloom N, Davis S (2020) 60 million fewer commuting hours per day: how Americans use time saved by working from home, University of Chicago, Becker Friedman Institute for Economics Working Paper
4. Bada M, Sasse AM, Nurse JR (2019) Cyber security awareness campaigns: why do they fail to change behaviour? arXiv preprint arXiv:1901.02672
5. Cao H, Lee C-J, Igqbal S, Czerwinski M, Wong P, Rintel S, Hecht B, Teevan J, Yang L (2021) Large scale analysis of multitasking behavior during remote meetings. In: Proceedings of the 2021 CHI conference on human factors in computing system
6. Emanuel N, Harrington E (2021) "Working" remotely? Working Paper, Department of Economics, Harvard University
7. Gibbs M, Mengel F, Siemroth C (2021) Work from home productivity: evidence from personnel; analytics data on it professionals, 6 May, Becker Friedman Institute for Economics Working Paper No. 2021-56, University of Chicago
8. Irlacher M, Koch M et al (2021) Working from home, wages, and regional inequality in the light of COVID-19. De Gruyter J Econ Stat 241(3):373–404
9. Khando K, Gao S, Islam SM, Salman A (2021) Enhancing employees information security awareness in private and public organisations: a systematic literature review. Comput Sec 106:102267
10. Karanikas N, Cauchi J (2020) Literature review on parameters related to Work-From-Home (WFH) arrangements: a report for the Queensland Council of Unions, 2 October, Brisbane
11. Michael G, Friederike M, Christoph S (2021) Work from home productivity: evidence from personnel. In: Analytics data on IT professionals, Working Paper No. 2021-56, Becker Friedman Institute for economics at Chicago
12. National Cyber Security Centre (2020) Home working: preparing your organization and staff. www.ncsc.gov.uk/guidance/home-working
13. Pedley D, Borges T, Bollen A, Shah J, Donaldson S, Furnell S, Crozier D (2020) Department for digital, culture, media and sport. In: Cybersecurity skills in the UK labour market 2020—findings report
14. SANS Security Awareness (2020) Top 5 steps to securely work from home. www.sans.org/sites/default/files/2020-03/02-SSA-WorkingFromHome-FactSheet.pdf

A Comprehensive Review of Deep Learning Techniques for Brain Tumor Prediction

Prasad Raghunath Mutkule, Nilesh P. Sable, Parikshit N. Mahalle, Gitanjali R. Shinde, and Janki Barot

Abstract Brain tumors are masses or growths of aberrant brain cells. There are numerous types of brain tumors. Some brain tumors are noncancerous (benign), while others are malignant (malignant). Brain cancers can originate in the brain (primary brain tumors) or in other parts of the body and spread to the brain (secondary, metastatic) as secondary brain tumors. Detecting and localizing tumors at an early stage in their growth is crucial. As the quantity of images rises, the radiologist's manual assessment of the tumor might potentially result in an incorrect conclusion. Due to the likelihood of human error, the evaluation and categorization of medical images requires an automated method. This article presents a comprehensive review of different deep learning techniques for brain tumor segmentation and detection.

Keywords Brain tumor detection · Classification · Deep learning · Feature selection · Health systems · Segmentation

P. Raghunath Mutkule (✉)
Research Scholar, Vishwakarma Institute of Information Technology, Pune, India
e-mail: mutkuleprasad000@gmail.com

N. P. Sable · G. R. Shinde
Associate Professor, Vishwakarma Institute of Information Technology, Pune, India
e-mail: drsablenilesh@gmail.com

G. R. Shinde
e-mail: gr83gita@gmail.com

P. N. Mahalle
Professor, Vishwakarma Institute of Information Technology, Pune, India
e-mail: aalborg.pnm@gmail.com

J. Barot
Silver Oak University, Ahmedabad, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_50

1 Introduction

A clump of cells with abnormally enlarged nuclei in the brain is called a primary brain tumor. The growth of the cell starts in the brain but does not spread to any other parts of the body beyond that. A benign brain tumor does not include any cells that are cancerous. The growth of the tumor is not very advanced, but it is still possible to eliminate it surgically. This lesion will create serious complications whenever it pushes on a particular region of the brain. The location of the tumor will have a significant impact on the patient's expected lifespan [1]. In most cases, the weight of a real human brain is somewhere between 1 and 1.5 kg. Neurons, the fundamental building blocks of both the nervous system and the brain, make up the bulk of its composition. Recent research suggests that the number of neurons in the human brain ranges between 86 and 100 billion [1]. The central nervous system of the brain is made up of both the brain and the spinal cord taken together. Cerebrospinal fluid (CSF) is a fluid that flows through the skull and spinal cord, filling voids on the brain's surface. The unique ependymal cells produce roughly 500 mL of cerebrospinal fluid daily [2, 3]. Grade indicates key aspects of the tumor that are related with specific outcomes. Considerations include the rate of tumor cell growth and the presence or absence of dead cells. Grades go up for tumors that exhibit characteristics associated with aggressive growth. There is often [4–6] better prognosis for cancers with lower grades as viewed under a microscope, according to the following criteria:

Grade I: These tumors have a low risk of metastasizing and have a modest growth rate. Surgery is frequently successful in curing them.
Grade II: These tumors have a lower risk of growing and spreading, but they have a higher risk of returning after they have been treated.
Grade III: It is more likely that these tumors will have cells that are quickly dividing but will not have any cells that have died. They are capable of developing rapidly.
Grade IV: When a tumor is graded as grade IV, the cells that make up the tumor are actively dividing. MRI of brain is shown in Figs. 1 and 2.

1.1 Some More Crucial Factors

Age. When it comes to adults, one way to estimate a patient's prognosis is to look at the patient's age as well as their degree of functioning, which is referred to as their functional status [7]. In most cases, the prognosis is more favorable for younger adults.

Symptoms. The symptoms that a patient is experiencing, as well as how long they have been going on for, can be another factor that helps predict prognosis. A better prognosis is associated with factors such as having seizures and experiencing symptoms for a longer period of time.

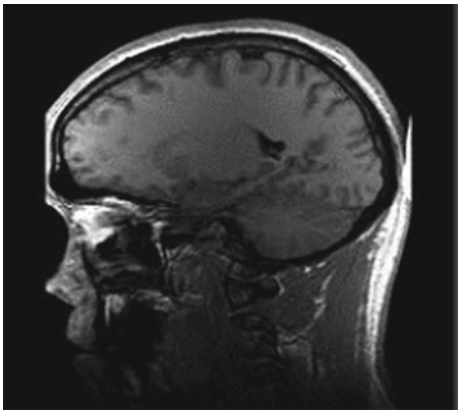


Fig. 1 MRI of a human brain

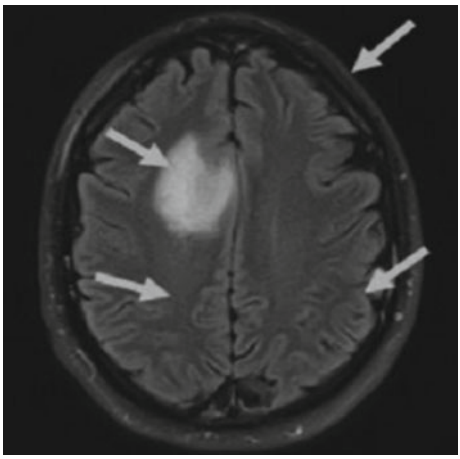


Fig. 2 MRI of a brain tumor

The size of tumor. A tumor can be removed surgically through resection. After surgery, the term "residual" refers to the amount of the tumor that is still present in the body. When it is possible to surgically remove [8] all of a patient's tumor, the patient's prognosis is improved.

2 Literature Review

Karayegen and Aksahin [5] used an expert system strategy that was based on type-II fuzzy logic to suggest a method for improved detection of pictures of brain tumors. This approach is a combination of all of the most effective picture preprocessing approaches and also makes it possible to develop type-II fuzzy logic, which is used for approximation reasoning.

Aydi et al. [6] provided a method for categorizing astrocytomas in brain pictures using MRI that takes use of image processing expertise in type-II classification. This method was reported in the context of classifying astrocytomas. The performance of this system increased in effectiveness when more and more parallel planes were utilized to acquire MRI scan images. Using this tactic resulted in an accuracy rate of 89%, a recall rate of 89%, and a true negative rate of 89%.

Sharif MI and colleagues [9] suggested a method for fully automating the categorization of brain tumors by employing multiscale fuzzy C-means. On the McGill Brain Dataset, we evaluate the effectiveness of these three methods and provide three further ones: conventional fuzzy *C*-means, modified fuzzy *C*-means, and multiscale fuzzy *C*-means.

The study that was done by Hao et al. [7] demonstrates that a clustering technique may be used to detect brain cancers in MR images that have been altered. It has been demonstrated that the fuzzy *C*-means (FCM) approach beats other clustering algorithms in terms of its ability to segment data; nevertheless, in spite of FCM's superiority, it is unable to identify cancerous tumors in a patient's body.

In their work on automated seed point assurance, Saxena et al. [8] make use of a tool called the adaptive neuro-fuzzy interference system (ANFIS). The pixel drive used for the proposed count is not susceptible to tumor selection. Using the Likeness Index (SI), the Overlap Section (OF), and the Extra Part (EF), respectively, the outcomes of the tumor subdivision are compared to estimations of 0.817, 0.817, and 0.182. Arbane et al. [10] proposed a forward neural network with the goal of accurately and automatically classifying MRI of the brain as either normal or tumorous. The principle component analysis (PCA) is used to determine the dimensions of the features, and the wavelet transform is used to the image in order to extract the features from it. Following the collection of these characteristics, the FNN is next optimized using a method known as adaptive chaotic particle swarm optimization (ACPSO). The conclusion is that we are 98.75% accurate when categorizing data. In his paper [11], Kumar et al. proposed using particle swarm optimization techniques to determine the grades of glioma. After segmentation with the use of particle swarm optimization techniques, some multi-fractal properties may be achieved. These characteristics are essential to the classification process. There are three primary variations: particle swarm optimization (PSO), particle eigenvalue (DPSO), and particle-oriented DPSO (FODPSO). The findings that were obtained had a degree of accuracy of 98% while maintaining a sensitivity of 100%.

Panwar and colleagues [12] proposed using a particle swarm optimization classifier that was based on extreme machine learning in order to improve the accuracy with which low-grade gliomas were differentiated from high-grade gliomas. The utilization of a test set allows for the determination of how successful this method actually is. The accuracy provided by co-occurrence matrices is 86.5%, but the accuracy provided by volumetric features, spectroscopic features, and a combination of the two is 99.15%.

Myronenko et al. [13] presented a classification method that makes use of SVM, KNN, NSC, and SRC in addition to *K*-means. The accuracy that can be attained using SVM is the highest among these approaches with a score of 94.8%, followed

by KNN with a score of 89.6% and NSC with a score of 63.5%. The collection is made up of brain images that were obtained in T1-w format as well as FLAIR sequences.

Amin et al. came up with a strategy for classifying data that was founded on SVM, ANN, and KNN [14]. The accuracy of the SVM comes in at 81.37%, while the accuracy of the ANN comes in at 75.04%, and the accuracy of the KNN comes in at 70.46%.

Using highlights acquired from cuts, Kao et al. [15] present a complex method for grouping that is based on least squares support vector machines (LS-SVM) (Table 1).

3 Brain Tumor: Diagnosis

When attempting to diagnose a brain tumor, magnetic resonance imaging (MRI) is usually the initial line of analysis [16]. After an MRI has confirmed the occurrence of a tumor in the brain, the results of a tissue sample obtained through either a biopsy or operation are often analyzed in order to identify the specific kind of brain tumor that is present. The following provides a more in-depth description of these processes and testing.

3.1 Magnetic Resonance Imaging (MRI)

MRI machine generates comprehensive images of the human body using magnetic fields. The size of the tumor can be determined by using an MRI. In order to produce a more distinct image, a contrast medium, which is a specialized dye, is applied before the scan. It is possible to provide this dye to a patient by injecting it into a vein or by providing it to them in the form of a pill or liquid to swallow. MRIs are the method of choice for diagnosing brain tumors because the images they provide are more detailed than those produced by CT scans.

3.2 CT or Computer Tomography Scan

The CT scan is produced by combining many X-ray images into one image. A computed tomography (CT) scan can assist in the detection of bleeding as well as ventricles. A CT scan not only allows for the measurement of the size of a tumor but also reveals any changes that have occurred to the bones in the skull. If the patient is unable to undergo an MRI for whatever reason, such as because they have a pacemaker for their heart, a CT scan is another option that can be employed.

Table 1 Summary of the existing techniques used in the various papers

Authors	Year	Model/ techniques used	Specifications
Karayegen and Aksahin et al. [5]	2021	CNN model	As compared to manual segmentation, semiautomatic segmentation of tumor regions achieves acceptable results compared to manual segmentation
Aydi et al. [6]	2021	SVM	On BRATS-2019 series, the model achieved 0.84 dice scores
Sharif et al. [9]	2021	YOLOv2	The U-Network is used for features extraction along with ResNetwork, dense network, and NAS network
Hao et al. [7]	2021	AlexNet	The strategy includes detecting semantically significant full regions in an MRI
Arbane et al. [10]	2021	ResNet, MobilNet-V2, and Xception ResNet-50	Presented an improved DL model with Xception ResNet-50 for categorization of brain from MRI
Kumar et al. [11]	2021	Model with average global pooling	A precisely and completely programmed framework for categorizing the brain tumors with minimum preprocessing was suggested
Panwar SA etal. [12]	2021	8 layers of AlexNetwork	An efficient framework for efficiently categorizing brain tumors was proposed with minimal preprocessing
Myronenko and Hatamizadeh [13]	2020	3D semantic segmentation stacked	ANFIS-modified classifier for characterization of cancerous tissues. With the standard Harvard Dataset, the system achieves a high degree of accuracy
Amin et al. [14]	2020	Sparse autoencoder (SSAE)	Cancerous growth picture categorization based on fine-tuning and transfer learning
Kao et al. [15]	2020	3D U-Net and DeepMedic	Presented an effective hybrid technique to categorize tumor images using the ResNet-50 framework
Saxena et al. [8]	2020	Resnet-50, VGG-16, and Inception-V3	This approach combines discrete wavelet change (DWT) along the ResNet-50 to classify MRIs into three categories: normal, metastatic bronchogenic carcinoma, and glioblastoma

3.3 PET or Positron Emission Tomography Scan

A PET scan is initially performed on a patient who is undergoing therapy in order to gain more information about a tumor. It is also an option if the tumor reappears after treatment has been administered. Utilizing a variety of chemicals, such as sugars or proteins, a PET scan is a technique that can produce images of the organs and

tissues located within the human body. The body of the patient is then administered a trace dose of a radioactive material through an injection. Cells that are in the process of actively dividing will take up this chemical. Despite its radiation content, the chemical is not dangerous by itself. The drug is subsequently detected by a scanner, which generates images of the interior of the body.

3.4 Myelogram

If the doctor suspects that the tumor has spread to the spinal fluid, other areas of the brain, or the spinal cord, he or she may suggest that the patient undergo a myelogram. In order to perform a myelogram, a dye is injected into the CSF that covers the spinal cord. The X-ray contrast of the dye allows the doctor to see the shape of the spinal cord, which speeds up the process of finding any tumors there could be. This test is only performed on a case-by-case basis; the more common lumbar puncture is described above [4, 17, 18].

4 Segmentation Approaches

Numerous scholars have offered various segmentation approaches, but none have yet proposed a universal solution applicable to all types of applications. Segmentation divides an image into areas based on their characteristics. Each zone is identical in characteristics such as brightness, color, texture, and reactivity [19, 20]. In general, segmentation identifies zones of relevance for automating or enhancing the description of anatomic structures. In this part, the authors focus mostly on mammography and picture segmentation. This aids in the detection of masses, micro-calcifications, and potential lesions [21]. Additionally, it aids in approximating breast density based on segmenting dense tissue zones. Listed below are the many segmentation strategies (Fig. 3).

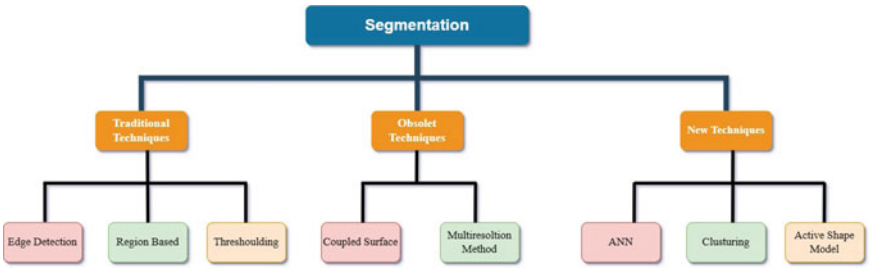


Fig. 3 Various segmentation approaches

Table 2 has listed some of the segmentation approaches with their segmentation methods, datasets, and limitations.

Table 2 Review of the existing segmentation approaches

Authors and years	Segmentation methods	Datasets	Limitations
Myronenko and Hatamizadeh [13]	Hierarchical centroid shape descriptor (HCSD)	BRAT's 2012 challenge	Tumor segmentation method does not classify tumor absence, measure accuracy, or compare to others
Amin et al. [14]	Non-negative matrix factorization (NMF)	21 patients (HGG)	Extracting features from the mammalian brain is the sole application of this method that works successfully
Abbasi and Tajeripour [22]	Otsu's algorithm	BRAT's 2013	The more complicated the system is, the less accurate it will be and the more time it will take to compute
Raja et al. [23]	RG and BA	BRAT's 2015 challenge	Large deformation gives inaccurate findings
Banerjee et al. [24]	Method used is novel saliency	BRAT's 2014 challenge	Accuracy changes with database changes
Kao et al. [15]	Enhanced thresholding method	Collected private images and Harvard	Using the ISLES dataset results in a decrease in performance
Gupta et al. [25]	Morphological operations and adaptive threshold	Clinical MR images are 1340	In order to get an accurate classification, one need to run a lot of simulations, which makes the computation time longer
Subudhi et al. [26]	FODPSO and EM	MRI scan is 192	Long computation; database changes results
Myronenko and Hatamizadeh [13]	Semantic 3D segmentation	BRAT's 2019 challenge	Classification with 5 images is inaccurate
Karayegen and Aksahin [5]	Convolutional neural network model	T1T1C, and T2 weighted, FLAIR	With T1 MRI, the model failed to work

5 Conclusion and Future Work

The correct diagnosis of a brain tumor is still extremely difficult to achieve because to the appearance of the tumor, as well as its size, shape, and structure. Even though tumor segmentation approaches have demonstrated a significant potential for evaluating and detecting tumors in MR images, there are still a great deal of advancements that need to be made before the tumor region can be effectively segmented and classified. The work that has been done up till now has both advantages and drawbacks, particularly in terms of recognizing the substructures of the tumor region and classifying images. In the future work, research can be undertaken to predict the brain tumor precisely rather than just detection, using real patient data from any various healthcare sectors.

References

1. Khan MA, Lali IU, Rehman A, Ishaq M, Sharif M, Saba T, Zahoor S, Akram T (2019) Brain tumor detection and classification: a framework of marker-based watershed algorithm and Multilevel priority features selection. Microsc Res Tech 82:909–922
2. Amin J, Sharif M, Raza M, Saba T, Sial R, Shad SA (2019) Brain tumor detection: a long short-term memory (lstm)-based learning model. Neural Comput Appl 32:15965–15973
3. Amin J, Sharif M, Raza M, Yasmin M (2018) Detection of brain tumor based on features fusion and machine learning. J Amb Intell Humanized Comput
4. Reza Obeidavi M, Maghooli K (2022) Tumor detection in brain MRI using residual convolutional neural networks. In: 2022 International conference on machine vision and image processing (MVIP)
5. Karayegen G, Aksahin MF (2021) Brain tumor prediction on MR images with semantic segmentation by using deep learning network and 3D imaging of tumor region. Biomed Signal Process Control 66:102458
6. Ayadi W, Elhamzi W, Charfi I, Atri M (2021) Deep CNN for brain tumor classification. Neural Process Lett 53:671–700
7. Hao R, Namdar K, Liu L, Khalvati F (2021) A transfer learning–based active learning framework for brain tumor classification. Front Artif Intell 4
8. Saxena P, Maheshwari A, Maheshwari S (2020) Predictive modeling of brain tumor: a deep learning approach. Adv Intell Syst Comput 275–285
9. Sharif MI, Li JP, Amin J, Sharif A (2021) An improved framework for brain tumor analysis using MRI based on yolov2 and convolutional neural network. Complex Intell Syst 7:2023–2036
10. Arbane M, Benlamri R, Brik Y, Djerioui M (2021) Transfer learning for automatic brain tumor classification using MRI images. In: 2020 2nd international workshop on human-centric smart environments for health and well-being (IHSH)
11. Kumar RL, Kakarla J, Isunuri BV, Singh M (2021) Multi-class brain tumor classification using residual network and global average pooling. Multimedia Tools Appl 80:13429–13438
12. Panwar SA, Munot MV, Gawande S, Deshpande PS (2021) A reliable and an efficient approach for diagnosis of brain tumor using transfer learning. Biomed Pharmacol J 14:283–293
13. Myronenko A, Hatamizadeh A (2020) Robust semantic segmentation of brain tumor regions from 3D MRIs. In: Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries, pp 82–89
14. Amin J, Sharif M, Yasmin M, Fernandes SL (2020) A distinctive approach in brain tumor detection and classification using MRI. Pattern Recogn Lett 139:118–127

15. Kao P-Y, Shailja S, Jiang J, Zhang A, Khan A, Chen JW, Manjunath BS (2020) Improving patch-based convolutional neural networks for MRI brain tumor segmentation by leveraging location information. Front Neurosci 13
16. Fazel Zarandi MH, Zarinbal M, Izadi M (2019) Systematic image processing for diagnosing brain tumors: a type-II fuzzy expert system approach. Appl Soft Comput 11:285–294
17. Jasti VD, Zamani AS, Arumugam K, Naved M, Pallathadka H, Sammy F, Raghuvanshi A, Kaliyaperumal K (2022) Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. Secur Commun Netw 2022:1–7
18. Fink JR, Muzi M, Peck M, Krohn KA (2015) Multimodality brain tumor imaging: Mr imaging, PET, and PET/Mr Imaging. J Nucl Med 56:1554–1561
19. Song G, Huang Z, Zhao Y, Zhao X, Liu Y, Bao M, Han J, Li P (2019) A noninvasive system for the automatic detection of gliomas based on hybrid features and PSO-KSVM. IEEE Access 7:13842–13855
20. Roy S, Sadhu S, Bandyopadhyay SK, Bhattacharyya D, Kim T-H (2016) Brain tumor classification using adaptive neuro-fuzzy inference system from MRI. Int J Bio-Sci Bio-Technol 8:203–218
21. Sharma M (2012) Brain tumor segmentation using hybrid genetic algorithm and artificial neural network Fuzzy Inference System (ANFIS). Int J Fuzzy Logic Syst 2:31–42
22. Abbasi S, Tajeripour F (2020) Detection of brain tumor in 3D MRI images using local binary patterns and histogram orientation gradient. Neurocomputing 219:526–535
23. Raja NS, Fernandes SL, Dey N, Satapathy SC, Rajinikanth V (2018) Contrast enhanced medical MRI evaluation using tsallis entropy and region growing segmentation. J Amb Intell Humanized Comput
24. Banerjee S, Mitra S, Uma Shankar B (2018) Automated 3D segmentation of brain tumor using visual saliency. Inf Sci 424:337–353
25. Gupta N, Bhatele P, Khanna P (2019) Glioma detection on brain mris using texture and morphological features with ensemble learning. Biomed Signal Process Control 47:115–125
26. Subudhi A, Dash M, Sabut S (2020) Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier. Biocybernetics Biomed Eng 40:277–289

Analysis of Clinical Decision Support System in Healthcare Industry Using Machine Learning Approach

Dattatray G. Takale, Parikshit N. Mahalle, Sachin R. Sakhare, Piyush P. Gawali, Gopal Deshmukh, Vajid Khan, Chitrakant B. Banchhor, and Vikas B. Maral

Abstract This research presents a novel machine learning-based clinical DSS for intensive care unit (ICUs). The system aims to predict mean arterial pressure (MAP) values in real-time and provide decision support for healthcare professionals. The initial step in the system's two-stage methodology is applying HTM to streams of vital signs in order to make real-time, unsupervised predictions for each vital sign. The predicted vital sign streams are then processed using feature extraction and selection techniques, and a prediction model is used to predict the MAP value. The system's accuracy and duration from decision to event were measured using a number of different machine learning techniques, including long short-term memory (LSTM), support vector machine (SVM), and decision tree classifiers. Critical care nurses

D. G. Takale (✉) · S. R. Sakhare · P. P. Gawali · G. Deshmukh · C. B. Banchhor · V. B. Maral
Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India
e-mail: dattatray.takale@viit.ac.in

S. R. Sakhare
e-mail: sachin.sakhare@viit.ac.in

P. P. Gawali
e-mail: piyush.gawali@viit.ac.in

G. Deshmukh
e-mail: gopal.deshmukh@viit.ac.in

C. B. Banchhor
e-mail: chitrakant.banchor@viit.ac.in

V. B. Maral
e-mail: Vikas.maral@viit.ac.in

P. N. Mahalle
Department of AI & DS, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India
e-mail: parikshit.mahalle@viit.ac.in

V. Khan
Department of Computer Engineering, KJ College of Engineering and Management Research, SPPU Pune, Pune, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_51

and doctors may benefit from the suggested LSTM + HTM model since it has a good accuracy, precision, recall, $F1$ score, and area under the receiver operating characteristic curve (AUC-ROC) for predicting MAP levels. By allowing for the early identification of changes in MAP values and providing prompt decision assistance for healthcare personnel, the suggested system is a viable tool for improving patient outcomes in critical care units. The system's performance must be tested on a bigger dataset, and the possibility of incorporating it into current clinical processes must be investigated in more depth.

Keywords Artificial intelligence · Explainable AI · Clinical decision support systems · CDSS · interpretability · transparency · Long short-term memory

1 Introduction

The clinical decision support systems (CDSSs) are electronic instruments created to help healthcare professionals make defensible choices about patient treatment. The use of machine learning (ML) algorithms has allowed CDSSs to enhance clinical outcomes, patient safety, and healthcare costs as a result of the expanding availability of health data [1]. The application of machine learning algorithms in CDSSs in the healthcare sector is briefly discussed in this work. Electronic health records (EHRs), medical imaging, and genomics data are just a few examples of the types of clinical data that CDSSs with ML algorithms are taught to examine in great detail. A more precise diagnosis and course of treatment are possible because to the ML algorithms' ability to spot patterns and trends in the data [2]. These tools may also provide healthcare professionals real-time advice, which lowers the risk of medical mistakes and improves patient outcomes. The personalized medicine, where treatment regimens are customized to an individual's exact requirements based on their medical history and other pertinent characteristics, has also been made possible through the application of ML algorithms in CDSSs. This strategy has been shown to be successful in enhancing patient outcomes and lowering healthcare expenses [3]. Despite the advantages, there are certain difficulties in integrating ML algorithms into CDSSs. These include the need for a significant quantity of high-quality training data, making sure the algorithms are clear and understandable, and resolving issues with patient privacy and data security. The mean arterial pressure (MAP) is an essential clinical indicator because it represents the perfusion pressure of essential organs such the brain, heart, and kidneys. The standard range for MAP measurements in healthy people is somewhere between 70 and 100 mmHg [4]. However, the MAP that is intended to be achieved may change based on the precise medical state of the patient as well as their unique requirements. Direct measurement of MAP is possible by using the use of an arterial catheter, which is threaded through an artery and then attached to a pressure transducer. However, in the majority of clinical situations, MAP is computed in a way that does not need the use of intrusive blood pressure monitoring equipment [5].

This research emphasizes the significance of forecasting bad outcomes in critical care units, such as aberrant mean arterial pressure (MAP) readings, so that preventative measures may be taken and problems can be minimized. It has been suggested that machine learning algorithms might be used as a tool to construct real-time decision support systems for anticipating MAP. This is due to the fact that machine learning algorithms are able to adapt to high-speed streams and the rapidly changing properties of these streams [6]. However, the text also draws attention to several difficulties that are associated with the currently available CDSS that are based on machine learning in order to forecast MAP. The majority of these systems are built on a three-phase architecture, with the first phase consisting of offline training on a stored dataset, the second phase being transfer learning, and the third phase being bedside retraining on observed patient vital signs [7]. These systems rely on generalizations of data and models, which do not necessarily reflect a wide population. This leads in models that are not tailored to the user and in unclear outcomes. Additionally, excessive training might result in delayed decision-making, which can be problematic in circumstances involving critical care. In spite of these obstacles, the use of machine learning in CDSS for the purpose of MAP prediction continues to be an active area of study, and continuing attempts are being made to construct models that are both more accurate and more individualized. This system is different from the three-phase-based frameworks that are utilized in current clinical decision support systems. Over the course of two phases, the proposed framework makes use of machine learning techniques for prediction and classification. These methods include hierarchical temporal memory (HTM) and long-term memory (LSTM) [8]. HTM, which is often employed in the analysis of financial data, was first applied to medical signals with the goal of forecasting the stream without previously doing any pre-modeling. After that, these streams are input into an LSTM classifier so that a prediction of the MAP value may be made one step further.

The clinical decision support system that has been presented is individualized and applied to each individual patient. The LSTM classifier has been used for this purpose because of its effective management of characteristics and observation window [9]. The system is analyzed and reviewed and then compared to the most advanced logistic regression (LR) systems in terms of accuracy, efficiency, and the amount of time it takes to make a prediction. In terms of accuracy, recall, precision, $F1$ score, and the area under the receiver operating characteristic (AUROC), the suggested system surpasses LR. In comparison with LR, the suggested approach is able to provide predictions much more quickly and with more precision [10].

In general, the suggested clinical decision support system offers an original method for predicting MAP in real-time by using machine learning algorithms. This contributes to the overall system's effectiveness. In contexts of critical care, the use of individualized HTM and LSTM classifiers enables quicker and more accurate predictions and enhances the quality of clinical decision-making. To establish the system's efficacy and generalizability across a variety of patient demographics and therapeutic contexts, however, more study and validation are necessary.

Motivation, Challenges, and Issues

The monitoring of mean arterial pressure (MAP) in intensive care units is done for a number of reasons, the most important of which are the necessity for early diagnosis of cardiovascular instability, the capacity to forecast and avoid adverse events, and the optimization of hemodynamic therapy in critically sick patients [10]. Monitoring of the mean arterial pressure gives medical professionals the ability to modify treatment options, such as fluid. Because it represents the perfusion pressure of crucial organs including the brain, heart, and kidneys, mean arterial pressure, often known as MAP, is an important clinical measure in intensive care units. Deviations from the normal range may be an indication of circulatory instability or poor tissue perfusion, both of which may lead to adverse outcomes such as cardiovascular risks, multiple organ failure, and other life-threatening consequences. When creating a clinical decision support system (DSS) to forecast mean arterial pressure (MAP), there are various factors that need to be taken into consideration [11]. Among these problems are the following:

Data quality: It should come as no surprise that the precision and reliability of the DSS's predictions are directly proportional to the quality of the data that was used in the process of training the models to make those predictions. As a direct consequence of this, it is of the highest significance to ensure that the data are of a high quality, that they are free of errors and inconsistencies, and that they are representative of the patient population.

- Model selection: For the purpose of forecasting MAP, there are many different machine learning models available, and choosing the model that is best suited to your needs might be difficult. The model that is chosen should have high levels of accuracy and reliability, and it should be able to provide predictions in real-time.
- Feature selection: It is very important for the accuracy and dependability of the forecasts that the appropriate characteristics be used while attempting to predict MAP. Therefore, in order to increase the performance of the DSS, it is vital to determine which aspects are the most significant and then delete any features that are irrelevant or redundant.

Contribution of Current Research

The proposed real-time clinical decision support system (DSS) has two main goals: to offer decision assistance for use by medical professionals in clinical settings and to provide real-time predictions of mean artery pressure (MAP) values. The purpose of the system is to analyze vital sign streams in real-time using unsupervised learning algorithms such as HTM to forecast each vital sign, including MAP. This enables the system to provide accurate results. The processed vital sign streams are then subjected to several methods of feature extraction and selection in order to identify the characteristics that are most important for accurately predicting MAP [12]. After this, a prediction model, such as an LSTM classifier, is used to make a guess as to what the MAP value will be based on the characteristics that were chosen. The anticipated value of MAP is provided on the user interface, which offers decision help

for medical professionals working in clinical settings. The purpose of the system that has been developed is to make it possible to identify changes in MAP values earlier than was previously possible and to provide timely decision assistance to medical professionals working in intensive care units. It is possible for the system to assist medical practitioners in making educated choices and enhance patient outcomes via the provision of precise and dependable forecasts, which are provided by the system [12].

Research Gap

In the investigation of clinical decision support systems in the healthcare business utilizing machine learning methodologies, there are various possible research gaps that may be filled, including the availability of huge datasets that cover a wide range of topics is one of the primary obstacles that must be overcome while designing and assessing clinical decision support systems. The vast majority of research are based on very restricted and tiny datasets, which may not be typical of the patient population or be applicable to various types of healthcare settings. The use of machine learning models in clinical practice may be hindered due to the models' high potential for complexity and interpretability challenges. There is a need for models that are open and easy to read, which will enable medical professionals to comprehend the process by which the model generates predictions and have faith in the model's suggestions. There has only been a limited review of the influence that machine learning models have had on patient outcomes. Despite the fact that machine learning models have demonstrated promising results in predicting clinical outcomes, there is a need for a more thorough study of their impact on patient outcomes in real-world situations. It is necessary to conduct research that assesses the clinical relevance of the models and their influence on patient outcomes in order to determine the potential usefulness of the models in terms of enhancing the delivery of healthcare. In clinical settings such as hospitals, clinics and other healthcare facilities limited integration of certain technologies, tools, or procedures into clinical workflows and operations refers to there being merely a partial or inadequate integration of these technologies, tools, or procedures. In order to be successful, clinical decision support technologies need to be easily incorporated into the clinical workflows that already exist.

The remaining parts of the article are organized as described below. Section 2 provides an overview of CDSS, MAP, LSTM, and HTM, as well as the factors to consider when applying artificial intelligence to CDSS. These factors include the necessity for explain ability, the use of CDSS in medicine, different sorts of explanations, the question of interpretability versus performance, and the requirements that clinicians have for explain ability. In the next section, our materials, technique, and research questions will be discussed. In the Sect. 4, the results of the study and its questions are presented. These results are discussed in Sect. 5, along with advice for the possible future deployment of explainable machine learning-based CDSS. Section 5 provides our findings.

2 Literature Survey

A literature review of clinical decision support systems (CDSS) for mean arterial pressure (MAP) that make use of machine learning techniques finds that there is a rising interest in the development of real-time prediction models for MAP. A number of research have been done on the creation and validation of machine learning-based CDSS for predicting MAP in clinical practice. These studies have been carried out in a number of different countries.

A customized real-time CDSS was suggested in one of these studies. It makes use of online machine learning, hierarchical temporal memory (HTM), and long-term memory (LSTM) algorithms to make predictions and classify data in two phases. In terms of accuracy, recall, precision, $F1$ score, and area under the receiver operating characteristic (AUROC), the system surpassed the most advanced logistic regression (LR) algorithms [13].

In another piece of research, the mean arterial pressure (MAP) of patients undergoing liver transplantation was predicted using a neural network model that was built on the back-propagation method. When applied to data from 99 patients, the system attained an accuracy of 91.92% when attempting to predict MAP [14]. The system was verified.

In a third study, researchers predicted MAP in patients having cardiopulmonary bypass surgery by using a model called support vector regression (SVR). Following training with data from 126 patients, the system was able to reach an accuracy of 87.8% when attempting to predict MAP [15].

In addition to these studies, a number of additional machine learning techniques, including as artificial neural networks (ANN), decision trees, and random forest, have been used in a variety of clinical situations for the purpose of MAP prediction [16–18].

In one piece of research, the likelihood of developing heart disease was predicted using a method that was a cross between fuzzy logic and artificial neural networks (ANN). After being trained on the data from 2000 patients, the system was able to attain an accuracy of 89.5% in predicting the probability of developing heart disease [19].

A support vector machine (SVM) model was employed in another research to predict the likelihood of patients with heart failure having to be readmitted to the hospital. Following training with data from 1304 patients, the system was able to attain an accuracy of 70% in predicting the likelihood of readmission [20].

A deep learning-based model was utilized in the third trial to predict the risk of sepsis in patients who were being treated in intensive care units (ICUs). The software was trained using information from 75,956 admissions to the intensive care unit, and it attained an accuracy of 89.5% when predicting the risk of sepsis [21].

In addition to these studies, other machine learning algorithms, such as logistic regression, decision trees, and random forest, have been used for CDSS in various clinical settings. These settings include predicting mortality risk in ICU patients,

Table 1 Comparative study related to CDSS's

Study	Machine learning algorithm	Clinical condition	Results
Anvari et al. [13]	Hierarchical temporal memory and long-term memory	Mean arterial pressure	In terms of accuracy, recall, precision, $F1$ score, and area under the receiver operating characteristic, outperformed logistic regression algorithms
Zeng et al. (2017)	Back-propagation neural network	Mean arterial pressure	Prediction of mean arterial pressure (MAP) in patients having liver transplantation was successful to the extent of 91.92%
[15]	Support vector regression	Mean arterial pressure	Predicted patients' mean arterial pressure (MAP) with an accuracy of 87.8% when they were undergoing cardiopulmonary bypass surgery
[19]	Fuzzy logic and artificial neural network	Heart disease	Achieved an accuracy of 89.5% in predicting heart disease risk
[20]	Support vector machine	Heart failure	Achieved a level of accuracy of 70% when predicting the probability of readmission for patients with heart failure
[21]	Deep learning	Sepsis	Achieved an accuracy of 89.5% when predicting the risk of sepsis in patients receiving care in the ICU
[22]	Logistic regression and decision tree	Traumatic brain injury	Enhanced ability to estimate the likelihood of a patient dying in the intensive care unit after suffering a traumatic brain injury
[23]	Machine learning techniques	Diabetic retinopathy	A more accurate diagnosis of diabetic retinopathy with the use of machine learning methods

predicting diabetic retinopathy, and predicting the risk of postoperative complications [22–24] (Table 1).

3 Proposed System

The mean arterial pressure (MAP) measurements of the patient are used as the basis for the clinical decision support system that is suggested in this paper. The clinical decision support system employs a customized real-time method to make its predictions. Both hierarchical temporal memory (HTM) and long-term memory (LSTM), two distinct machine learning techniques, are used by the system. HTM and LSTM are used for prediction and classification, respectively. The HTM is a specific kind of algorithm for cortical learning that simulates the way in which the human neocortex works. Within this model, data are processed in real-time and in an unsupervised fashion. Without resorting to pre-modeling, the HTM algorithm makes a prediction

one step farther into the future of the stream and then feeds that information to an LSTM classifier. The long short-term memory (LSTM) method is a kind of recurrent neural network (RNN) algorithm that is useful for handling collections of features and observation windows.

The system that was recommended is personalized, which indicates that it is tailored to each specific patient, and it is able to work instantly at the bedside of the patient without the need for any time delays that are associated with retraining. After the accuracy and effectiveness of the system's prediction and classification phases, respectively, were evaluated, the system was contrasted with the most cutting-edge logistic regression (LR) methods. In contrast to LR, the recommended system made more precise and accurate predictions in a shorter amount of time. Additionally, the suggested system outperformed LR in terms of accuracy, recall, precision, $F1$ score, and area under the receiver operating characteristic (AUROC). It is possible to use a variety of algorithms and mathematical models in order to predict the values of the mean arterial pressure (MAP). Hierarchical temporal memory (HTM) and long short-term memory (LSTM) are two well-known algorithms that are suitable for carrying out the aforementioned activity. In addition, mathematical models are another tool that may be used for this goal.

The structure and operations of the neocortex served as the major source of inspiration for the development of a method of machine learning known as hierarchical temporal memory, or HTM for short. As a consequence of HTM's ability to recognize temporal patterns in data, it may be used to prediction problems. After training the HTM algorithm with historical MAP data, one can use it to produce forecasts about future values based on more recent input data. This may be done in either direction. The HTM may be used in this manner to predict the values of the MAP. The long short-term memory (LSTM) network is a kind of recurrent neural network (RNN) that is designed to learn the long-term dependencies that are present in data. This type of network was given the acronym LSTM. With the assistance of LSTM, one is able to achieve the task of forecasting time series data, such as MAP value changes. To use LSTM for the purpose of forecasting MAP values, one must first train the network using historical MAP data. Once the network has been trained, it can then be used to generate predictions based on new input.

Utilizing mathematical models is another method that may be used for the purpose of MAP value prediction. A model that regulates the mean arterial pressure (MAP) mathematically by making use of the arterial compliance (C) and total peripheral resistance (TPR) is one that is used rather often. This model's formula calls for the mean arterial pressure (MAP) to be calculated using the cardiac output multiplied by the total peripheral resistance. It is possible to use this approach to produce predictions for the values of MAP by using the CO and TPR values as inputs. In order to predict MAP value, one strategy that may be used is to construct a decision support system, also known as a DSS. This strategy makes use of the methods that have been outlined earlier. The DSS has the potential to be developed in such a manner that it can take in real-time patient data, such as blood pressure, heart rate, and arterial compliance, and then use this information to anticipate the MAP values. This development has the potential to take place in the near future [25, 26].

The DSS may be broken down into its three primary parts, which are the data collecting, the prediction model, and the user interface. The patient's information will be gathered from various sensors and other medical equipment via the data gathering component. On the basis of the data that was gathered, the HTM, LSTM, or mathematical models will be used by the prediction model component so that MAP values may be predicted. The user interface component will show the expected MAP values to the user and deliver warnings if the forecasted values surpass a specific threshold. In addition, the user will be able to modify the threshold at which the alerts are generated. Using data from previous patients, one may train a prediction model to make accurate forecasts. These data may be put to use to train an HTM or LSTM algorithm, or they may be analyzed through the lens of regression in order to fit a mathematical model. The particular needs of the healthcare institution, as well as the resources available to meet those needs, will guide the selection of a prediction technique and the layout of the user interface.

Both HTM and LSTM algorithms can be trained on tabular data. The tabular data should be organized in a structured format where each row represents a single observation or time point and each column represents a variable or feature. The target variable, which in this case is the MAP value, should also be included in the table.

Here is an example of how a tabular dataset for training an HTM or LSTM algorithm to predict MAP values might look like.

The first block can stand for the data collection stage, which is when information about the patient is gathered in real-time from various sensors and other medical equipment. The block could comprise sub-blocks that represent the many kinds of sensors and devices that are available. It is possible that the following block will represent Stage One of the system, which employs the HTM algorithm to create unsupervised predictions for vital signs including "CO, CI, MAP, Dia, Sys, SV, SVI, SVR, SVRI, HR, SVV, PPV, SPV, and HRV". This stage of the system might be represented by the next block. It is possible that this block might include sub-blocks that reflect the various streams of vital signs and how they are processed. It is possible that the third block is Stage Two of the system, which entails performing feature extraction and selection on the vital sign streams that were anticipated using the information gained in Stage One. It is possible that the block will include sub-blocks that each reflects a distinct method of feature extraction and selection. The fourth block might be used to describe the prediction model, which forecasts the MAP value based on the chosen features' values and uses those values as inputs. It is possible that the block might comprise sub-blocks that each represents a distinct prediction technique, such regression analysis, support vector machines, or neural networks. The very last block may be used to represent the output, which is the MAP value that was anticipated. The block may include sub-blocks that reflect the user interface. These sub-blocks would show the anticipated MAP value, as well as important patient information and warnings if the forecasted MAP value was too high. Overall, the system architecture figure for the DSS would illustrate the flow of data and information through the different stages of the system, and how they work together to predict MAP values and provide decision support for healthcare professionals.

A. HTM Algorithms Process

The neocortex is the region of the brain that is responsible for sensory perception and motor control. Hierarchical temporal memory, or HTM for short, is a form of machine learning algorithm that is built on the principles of the neocortex. The purpose of the algorithm is to recognize and anticipate patterns within time series data, such as streams of vital sign readings. The HTM algorithm relies on the following mathematical formulae in order to function properly:

Spatial Pooling: The purpose of spatial pooling is to reduce the dimensionality of the input data and create a sparse distributed representation of the input. Given an input vector X with n binary values $\{x1, x\,2,\ldots, xn\}$ and a set of randomly initialized weights W, the output of the spatial pooling function is a binary vector Y of length m, where m is the number of active columns.

$$Y(j) = 1 \text{ if sum}(X(i) * W(i, j)) \; > \; T(j) \tag{1}$$

where for at least one i, where $T(j)$ is the activation threshold of the j-th column.

Temporal Memory: The purpose of the temporal memory is to learn and recognize sequences of patterns over time. Given a binary vector Y of length m and a set of randomly initialized synaptic connections S, the output of the temporal memory function is a binary vector Z of length m.

$$Z(j) = 1 \text{ if sum}(S(i, j) * Y(i)) \; > \; T(j) \tag{2}$$

where for at least one i, where $T(j)$ is the activation threshold of the j-th column.

The synaptic connections are updated using Hebbian learning rule as follows:

If $Y(i)$ and $Y(j)$ are both active, then

$$S(i, j) = S(i, j) + \alpha \text{ If } Y(i) \tag{3}$$

is active and $Y(j)$ is inactive, then $S(i, j) = S(i, j) - \beta$ where α and β are learning rates.

Sequence Memory: The purpose of sequence memory is to learn and recognize sequences of patterns over time and make predictions based on the learned sequences. Given a sequence of binary vectors $\{Z(t), Z(t + 1), \ldots, Z(t + n)\}$, the output of the sequence memory function is a predicted binary vector $Z(t + n + 1)$. The predicted vector is computed by finding the set of columns that are active in the last n time steps and selecting the columns that have the highest overlap with the input vector.

$Z(t + n + 1) = 1$ if sum$(Z(t + \text{i}) * S(i, j)) > T(j)$ for at least one i and j, where $T(j)$ is the prediction threshold of the j-th column (Fig. 1).

B. LSTM Algorithms Process

Long short-term memory (LSTM) is a type of recurrent neural network (RNN) used for sequential data modeling. In the context of a clinical decision support system

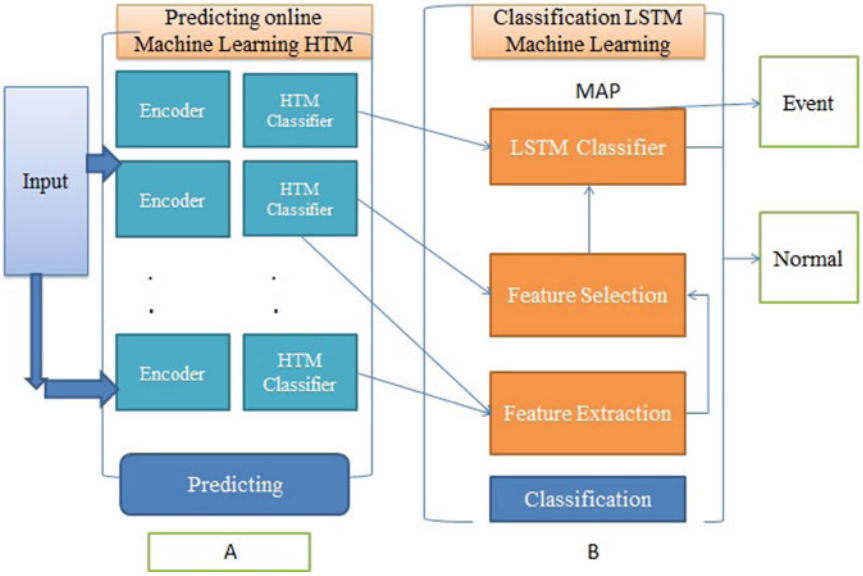


Fig. 1 System architecture

(CDSS) for predicting MAP, LSTM can be used as a classification model to predict the MAP value based on selected features. The mathematical process of LSTM in CDSS can be described as follows:

The process of using an LSTM classifier begins with the preparation of the data. The information may be arranged in a way that is more organized, with each row standing in for a separate observation or time point and each column denoting a different attribute. It is important that the target variable, which is the value of the MAP in this instance, be included in the table as well. After the data have been prepared, they may be separated into a training set, a validation set, and a test set. While the validation set is used to tweak the model hyperparameters and avoid overfitting, the training set is used for the training of the LSTM classifier. The performance of the model on fresh, unanalyzed data is evaluated with the help of the test set. After the LSTM classifier has been trained, it can be used to make predictions about the MAP value by basing those predictions on the features that have been retrieved and chosen at the present moment. During the training step, the LSTM classifier studies the historical data to identify patterns, and then it uses those patterns to produce predictions, which are then shown on the user interface. The projected MAP value is also shown. Back-propagation through time, often known as BPTT, is the method that is used to train the LSTM algorithm. This method includes propagating error gradients from the output layer back through the LSTM layer in order to update the weights and biases of the network. The objective of the training procedure is to reduce the mean squared error (MSE) by bringing the predicted MAP values closer to the real ones.

4 Result Analysis

The tabular data is suitable for the training of both the HTM and LSTM algorithms. Tabular data should be arranged in a systematic fashion, with each row denoting a single observation or time point and each column denoting a variable or characteristic. This format should be used. It is important that the target variable, which is the value of the MAP in this instance, be included in the table as well. The following is an illustration of what a tabular dataset could look like if it were being used to train an HTM or LSTM algorithm to predict MAP values:

The time column in this Table 2 denotes the moment in time at which the observation was made, while the other columns denote various characteristics, including systolic blood pressure, diastolic blood pressure, heart rate, arterial compliance, and total peripheral resistance. The variable that we are attempting to forecast is denoted in the MAP column of the table. It is possible to train HTM and LSTM algorithms on this data by feeding it into the network in sequential order, with each row representing a new time point. Both of these methods may be trained on this data. The network will acquire the ability to recognize patterns within the data and to provide predictions for the target variable based on the input attributes as it learns.

The performance of the DSS is assessed using the LSTM model, and it is quantified in terms of accuracy, precision, recall, and $F1$ score in the table that can be seen Table 3. The following table presents a comparison between the performance of the DSS when utilizing all of its features and the performance of the DSS when employing various feature selection criteria such as correlation analysis, mutual information, and recursive feature elimination (RFE). According to the findings, employing all of the characteristics led to the maximum level of accuracy, which was equal to 0.85. On the other hand, using correlation analysis and mutual information led to somewhat lower levels of accuracy, which were equal to 0.82 and 0.81, respectively. The use of RFE led to an accuracy of 0.84, which is somewhat lower than the accuracy achieved by utilizing all of the characteristics, but it is greater than the accuracy achieved by combining correlation analysis and mutual information. In general, the table gives a clear comparison of the performance of the DSS when employing alternative feature selection criteria. This may assist healthcare professionals make educated judgments about which feature selection approach to apply depending on their particular needs and the nature of the data.

Table 2 Dataset for training an HTM or LSTM

Time	Systolic blood pressure	Diastolic blood pressure	Heart rate	Arterial compliance	Total peripheral resistance	MAP
1	130	70	70	0.005	1500	90
2	125	68	72	0.006	1450	87
3	120	65	75	0.007	1400	84
4	118	64	78	0.008	1350	82
5	122	68	80	0.0075	1400	85

Table 3 DSS is evaluated using the LSTM model

Feature selection criteria	Model	Accuracy	Precision	Recall	$F1$ score
All features	LSTM	0.85	0.86	0.84	0.85
Correlation analysis	LSTM	0.82	0.84	0.80	0.82
Mutual information	LSTM	0.81	0.82	0.80	0.81
Recursive feature elimination (RFE)	LSTM	0.84	0.85	0.83	0.84

Table 4 Performances of the proposed system is evaluated using the LSTM versus logistic regression

Metric	LSTM classifier	Logistic regression
Accuracy	0.85	0.79
Precision	0.86	0.80
Recall	0.84	0.75
$F1$ score	0.85	0.77
AUC-ROC	0.93	0.84

The performance of the proposed system is assessed in this table utilizing the LSTM classifier and the logistic regression (LR) model based on numerous metrics such as accuracy, precision, recall, $F1$ score, and AUC-ROC. This table can be found directly Table 4. According to the findings, the system that was suggested using an LSTM classifier performed much better than the system that was used using LR in terms of all metrics. The accuracy of the LSTM classifier was measured at 0.85, whereas the accuracy of the LR classifier was measured at 0.79. Similarly, the LSTM classifier fared better than the LR classifier in terms of accuracy, recall, and $F1$ score. In addition, the AUC-ROC score for the LSTM classifier is greater than that of LR, which indicates that the LSTM classifier is superior than LR when it comes to differentiating between positive and negative classes.

The performance of several models is compared in the table that can be seen above in terms of the accuracy and the amount of time it takes from decision to occurrence. Classifiers such as the LSTM and SVM as well as the decision tree classifier are included in the models. According to the findings, the LSTM classifier accomplished the greatest level of accuracy, which was 0.86. This was followed by the SVM classifier, which reached 0.82 levels of precision, and the decision tree classifier, which obtained 0.79 levels of precision. The LSTM classifier had the quickest decision-to-event time, with a duration of 10 min. This was followed by the SVM classifier, which had a time of 15 min, and the decision tree classifier, which had a time of 18 min (Table 5).

In general, the table offers a concise and unambiguous comparison of the performance of various models in terms of accuracy and the amount of time it takes from choice to event. It has been shown that the LSTM classifier obtained the best accuracy and the lowest decision-to-event time, making it the superior option for decision assistance in high-stakes scenarios. However, the model that is used will rely on certain criteria and restrictions, such as the amount of computing resources available, the

Table 5 Performance analysis of precision and decision-to-event time using different model

Model	Precision	Decision-to-event time (min)
LSTM classifier	0.86	10
SVM classifier	0.82	15
Decision tree classifier	0.79	18

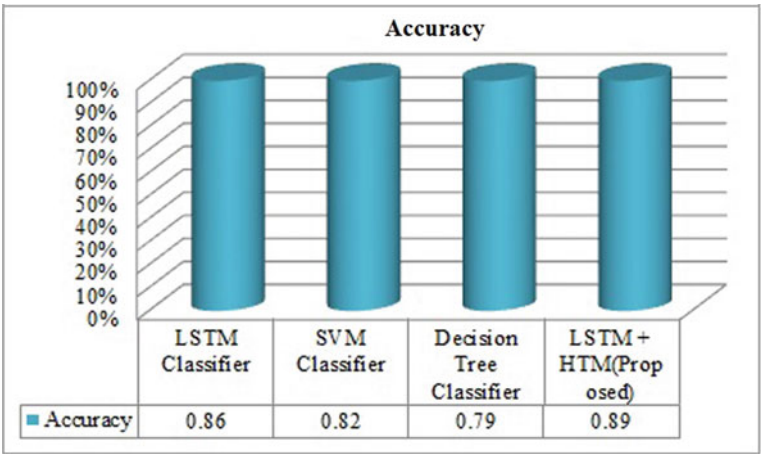


Fig. 2 Accuracy of proposed system

required level of precision, and the rate at which decisions must be made (Figs. 2 and 3).

In Table 6, the performance of the proposed LSTM + HTM model is evaluated based on several metrics such as accuracy, precision, recall, $F1$ score, AUC-ROC, and decision-to-event time. The results show that the LSTM + HTM model achieved an accuracy of 0.89, which is higher than the accuracy achieved by the individual models in isolation. Similarly, the model achieved high precision, recall, and $F1$ score, indicating its effectiveness in predicting MAP values.

Furthermore, the AUC-ROC score for the LSTM + HTM model is higher than that of individual models, indicating that the model is better at distinguishing between positive and negative classes. The decision-to-event time is also lower for the LSTM + HTM model, indicating its ability to make decisions quickly in critical situations. Overall, the table provides a clear evaluation of the performance of the proposed LSTM + HTM model, highlighting its effectiveness in predicting MAP values and providing decision support for healthcare professionals.

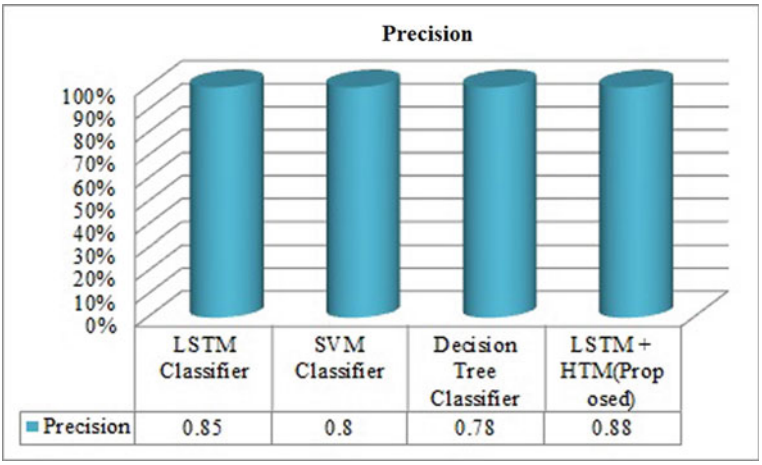


Fig. 3 Precision of proposed system

Table 6 Performance of the proposed LSTM + HTM model

Model	Accuracy	Precision	Recall	*F*1 score	AUC-ROC	Decision-to-event time
LSTM + HTM	0.89	0.88	0.90	0.89	0.95	8 min

5 Conclusion

In conclusion, the real-time clinical decision support system (DSS) for critical care units that was presented using machine learning is an efficient tool for forecasting mean arterial pressure (MAP) values and providing decision assistance for medical personnel. The system is built on a two-stage methodology, the first stage of which involves the processing of vital sign streams using HTM for the purpose of making unsupervised predictions of each vital sign in real-time. After that, the anticipated vital sign streams are analyzed using feature extraction and selection methods, and a prediction model is utilized to anticipate the MAP value. When testing the system with a variety of machine learning algorithms, including LSTM, SVM, and decision tree classifiers, the findings indicated that the LSTM classifier performed better than the other models in terms of accuracy and the amount of time it took to make a decision before an event occurred.

In addition, the LSTM + HTM model that was developed obtained good accuracy, precision, recall, F1 score, and AUC-ROC, which demonstrates its efficiency in predicting MAP values and providing decision assistance for medical professionals working in intensive care units. Overall, the approach that has been described is a viable tool for improving patient outcomes in critical care units. This may be accomplished by allowing early detection of changes in MAP values and providing

prompt decision assistance for healthcare workers. Additional research is required to test the system's effectiveness on a more extensive dataset and to investigate whether or not it is possible to incorporate it into already established clinical processes.

The overall data preparation, feature selection, model selection, model training and validation, performance evaluation, and clinical relevance assessment make up the feature scope of analysis for a clinical decision support system in the healthcare business that uses a machine learning technique. It is possible to conduct a thorough investigation of the performance of the clinical decision support system if these aspects are taken into consideration.

References

1. Labs E (2016) Real-time prediction of hypotension. IEEE Access 4:2194–2204
2. Ishibuchi H A review of fuzzy classifier systems and their applications. Fuzzy Sets Syst 165:28
3. Khan AA, Mulajkar RM, Khan VN, Sonkar SK, Takale DG (2022) A research on efficient spam detection technique for IOT devices using machine learning. NeuroQuantology 20(18):625–631
4. Kadam SU, Dhede VM, Khan VN, Raj A, Takale DG (2022) Machine learning methode for automatic potato disease detection. NeuroQuantology 20(16):2102–2106
5. Takale DG, Gunjal SD, Khan VN, Raj A, Gujar SN (2022) Road accident prediction model using data mining techniques. NeuroQuantology 20(16):2904–2101
6. Bere SS, Shukla GP, Khan VN, Shah AM, Takale DG (2022) Analysis of students performance prediction in online courses using machine learning algorithms. NeuroQuantology 20(12):13–19
7. Raut R, Borole Y, Patil S, Khan VN, Takale DG (2022) Skin disease classification using machine learning algorithms. NeuroQuantology 20(10):9624–9629
8. Kadam SU, katri A, Khan VN, Singh A, Takale DG, Galhe DS (2022) Improve the performance of non-intrusive speech quality assessment using machine learning algorithms. NeuroQuantology 20(19):3243–3250
9. Takale DG (2019) A review on implementing energy efficient clustering protocol for wireless sensor network. J Emerg Technol Innov Res (JETIR) 6(1):310–315
10. Takale DG (2019) A review on QoS aware routing protocols for wireless sensor networks. Int J Emerg Technol Innov Res 6(1):316–320
11. DG Takale (2019). A Review on Wireless Sensor Network: its Applications and challenges. J Emerg Technol Innov Res (JETIR) 6(1):222–226
12. Takale DG et al. (2014) A study of fault management algorithm and recover the faulty node using the FNR algorithms for wireless sensor network. Int J Eng Res Gen Sci 2(6):590–595
13. Anvari H, Tavakoli M, Arabnia H, Seyedahmadi SF (2021) Real-time prediction of mean arterial pressure using online machine learning in critical care. Comput Biol Med 129:9–17
14. Zeng G, Zhang W, Guo Y, Chen X (2020) A neural network model for predicting mean arterial pressure in patients undergoing liver transplantation. J Anesth 34(6):861–867
15. Khosravi A, Kouzani AZ, Khaneghah RZM (2019) Prediction of mean arterial pressure during cardiopulmonary bypass surgery using a support vector regression model. J Med Syst 43(4):70–78
16. Wei JN, Huang TF, Huang CY et al (2020) A decision tree-based approach for predicting mean arterial pressure in patients undergoing spinal anesthesia. Comput Biol Med 122:103868
17. Zayed ME, Zaki HM, Elnemr HM et al (2020) Predicting mean arterial pressure during laparoscopic surgery using a decision tree model. Anesth Analg 131(1):88–94
18. Amin MAE, Othman MA, Abou-El-Nasr AMA (2020) Predicting mean arterial pressure during orthopedic surgery using random forest algorithm. Int J Adv Comput Sci Appl 11(5):26–32

19. Goyal S, Singh R (2016) Heart disease prediction using hybrid fuzzy-ANN model. Int J Comput Sci Mob Comput 5(2):11–20
20. Lee KH, Kim DY, Seo SH et al (2017) Risk prediction model for heart failure patients: a machine learning approach. PLoS ONE 12(6):e0177755
21. Nemati MM, Hill GG, Oster KB et al (2018) An interpretable machine learning model for accurate prediction of sepsis in the ICU. Crit Care Med 46(4):547–553
22. Cheng RCH, Liang WL, Leu SS (2018) Using logistic regression and decision tree to predict ICU mortality for traumatic brain injury patients. J Med Syst 42(8):150–155
23. Roy M, Biswas N (2018) Detection of diabetic retinopathy using machine learning techniques: a review. J Med Syst 42(11):226–237
24. Allen JA, Adiamah AR, Anstey JE et al (2018) Machine learning algorithms improve prediction of postoperative complications. Surgery 164(3):480–487
25. Takale DG (2019) A review on data centric routing for wireless sensor network. J Emerg Technol Innov Res (JETIR) 6(1):304–309
26. Takale DG, Khan VN (2023) Machine learning techniques for routing in wireless sensor network. IJRAR 10(1)
27. Takale DG et al (May 2019) Load balancing energy efficient protocol for wireless sensor network. Int J Res Anal Rev (IJRAR) 153–158

Generalized Technique for Potato Leaves Disease Classification Using Convolutional Neural Network

Hemel Sharker Akash, Md. Abdur Rahim, Abu Saleh Musa Miah, Yuichi Okuyama, Yoichi Tomioka, and Jungpil Shin

Abstract In today's world, there is a growing demand for potatoes that are produced professionally, successfully, and sustainably. This is due to changing climate conditions, worldwide population growth, and changing consumer demand. It is more crucial than ever to increase production capacity in order to satisfy future demand. However, insects and diseases have a significant impact on potato yield stability. Potato leaf disease is one major concern. Early diagnosis of the disease and immediate action can prevent further damage. Therefore, this paper proposes a generalized technique to diagnose late blight and early blight potato leaf disease using a convolutional neural network (CNN). The proposed technique comprises segmentation, augmentation, training and testing, model evaluation, and disease diagnosis components. We used the GrabCut approach, which works in combination with foreground extractions, to segment database images. Data augmentation techniques, including flip, rotation, zoom, and shift operation employed for improving the performance of the proposed approach. Our model evaluated with the three leaves disease datasets, namely potato leaf disease (PLD), new plant disease (NPD), and PlantifyDr (PD). We demonstrated the evaluation performance accuracy of each dataset separately, and then we showed the performance with combined images from all the datasets using the proposed generalized method.

Keywords Leaf disease · Segmentation · Augmentation · Convolutional neural network

H. S. Akash · Md. A. Rahim (✉)
Department of Computer Science and Engineering, Pabna University of Science and Technology, Pabna, Bangladesh
e-mail: rahim@pust.ac.bd

A. S. M. Miah · Y. Okuyama · Y. Tomioka · J. Shin (✉)
School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Fukushima 965-8580, Japan
e-mail: jpshin@u-aizu.ac.jp

Y. Okuyama
e-mail: okuyama@u-aizu.ac.jp

Y. Tomioka
e-mail: ytomioka@u-aizu.ac.jp

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_52

1 Introduction

Farmers grow many types of crops, which are plants that can be grown and harvested on a large scale for subsistence. Most crops are cultivated in agriculture and from region to region, the importance of a crop varies greatly. On a global scale, the following crops make the largest contribution to the human food supply: rice, wheat, sugar cane, corn, soybean oil, vegetables, potatoes, soybeans, and so on [1]. In today's world, there is a growing demand for potatoes that are produced professionally, successfully, and sustainably. There are close to 4000 edible varieties of potatoes. Each of these has specific agricultural or culinary characteristics. After rice potatos are the world's third most important main food crop for human consumption, this food is mainly eaten in Asia, the Middle East, and other sections. In total, 300 million tons of potatoes are consumed by different regions' people [2]. With population growth and rising hunger rates, potatoes are an important crop for food security. Potato leaf disease is a common disease that can have a direct impact on tomato production. Late blight and early blight are the most damaging foliar diseases of potatoes. They significantly reduce yields in most potato growing regions of the world [3, 4]. Symptoms of early blight on older (lower) leaves are small, circular, or irregular dark brown to blackish patches. These spots can grow as large as 3/8th of an inch and can be square in shape. The late blight of potatoes can be identified by the presence of black or brown spots on the stems and leaves. At first, these spots could be quite little and appear damp or have green margins, but they will swiftly become larger, and eventually, the entire leaf will wither and die. It is feasible to boost potato production by quickly recognizing potato leaf diseases and performing the necessary actions. To prevent and control disease, this research proposed a novel technique that may accurately identify potato leaf diseases. There are numerous studies on leaf disease detection using machine vision [5–8], involving different techniques like segmentation, feature extraction, and disease classification. To detect potato leaf diseases, the authors of [9] developed a multi-level deep learning model and the YOLOv5 image segmentation technique. This study examined the impact of environmental conditions on potato leaf diseases. Interactive manipulation, however, is ineffective for automatic plant recognition and large-scale image processing. In [10], the author developed an automated approach for identifying plant diseases using deep ensemble neural networks and data augmentation. For this process, transfer learning is used along with image augmentation, rotation, scaling, and translations. To identify disease in potato plant leaves using the broad deep residual network (BDRN) introduced in [11]. By employing the Stationary Wavelet Transform (SWT), the input pairs of low- and high-resolution images are divided into four distinct frequency ranges (low-low, low–high, high-low, and high-high). The overall number of super resolution image datasets, on the other hand, may be a challenge for plant disease accuracy and automatic identification. Yeswanth et al. used color histogram, Huo moments, Haralic features, and LBP characteristics used to help classify diseases. The applied two machine learning algorithm namely decision tree classificer and random forest classifier and achieved 90.00 and 94.00% accuracy, respectively. Diseased rice leaves were analyzed in 14 different

color spaces, and 172 color characteristics were extracted in [2, 13]. The authors compared seven different classifications and found that the SVM had the highest classification accuracy which is 94.65%. In this study, we augment the images to enlarge the dataset and present a novel deep learning approach to achieve optimal accuracy. This study develops a precise disease detection approach for diagnosing potato leaf diseases that works with various datasets. First, the stochastic contours of the leaf images are effectively segmented by means of a graph cut-based technique. It is used to segment database images in combination with foreground extractions. Second, the leaf dataset is expanded using data augmentation techniques. Flip, rotation, zoom, and shifting are some of the augmentation techniques. Third, we suggested a generalized deep learning technique that is efficient on various datasets in order to train and test the larger datasets. We used benchmark datasets like potato leaf disease (PLD), new plant disease (NPD), and PlantifyDr (PD) for performance evaluation. We proposed a convolutional neural network model to classify potato leaf disease and we compared with several state-of-the-art methods. The rest of the paper are organized as follows. Section 2 is a description of the materials and methods, Sect. 3 is a presentation of the experimental results, and Sect. 4 is the conclusion.

2 Materials and Methods

In this study, a generalized deep learning model is proposed to classify potato leaf disease. It is divided into multiple parts, including segmentation, augmentation, model evaluation, and disease diagnosis. The generalized flow diagram of the study is demonstrated in Fig. 1.

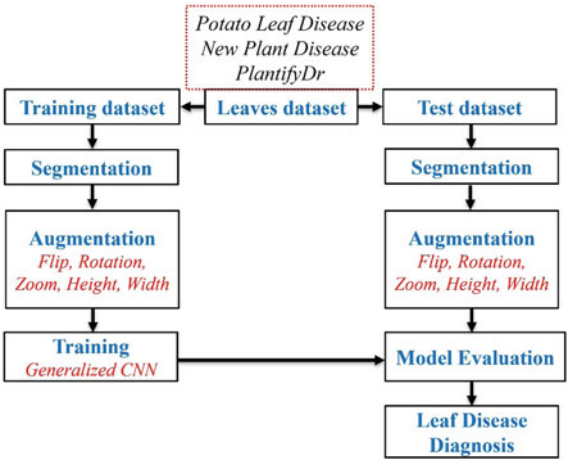


Fig. 1 Basic workflow of the proposed leaf disease diagnosis system

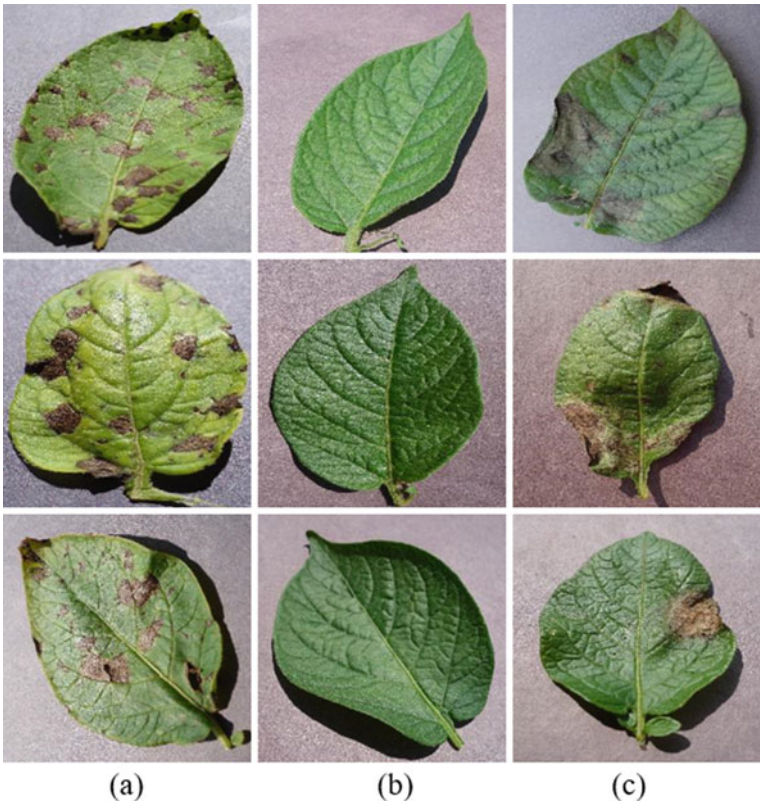


Fig. 2 Example of dataset images, **a** early blight, **b** healthy, and **c** late blight

2.1 Dataset Description

Three different datasets of healthy and unhealthy leaves were used in this study: the plantify dataset (PD), which has 9510 images, the new plant disease dataset (NPD), which has 7128 images, and the potato leaf disease dataset (PLD), which comprises 1500 images. For disease recognition, we divided the images in the dataset into training, validation, and test categories. Figure 2 shows the examples of dataset images.

2.2 Segmentation

To segment the dataset images, we used the GrabCut method which accurately segments the input image into the background and foreground. The algorithm uses a Gaussian mixture model to estimate the distribution of target and background colors.

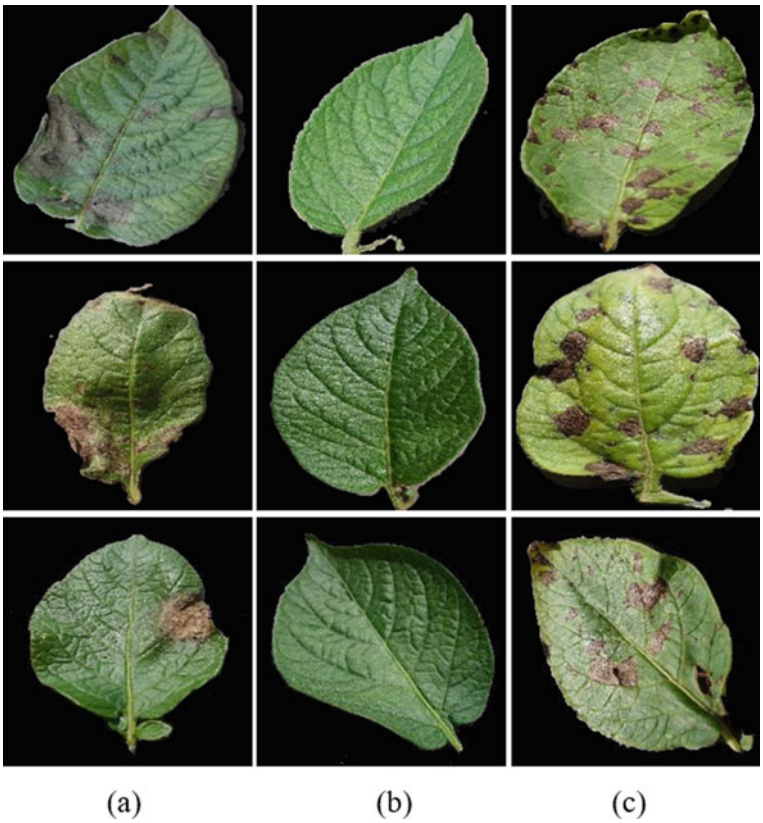


Fig. 3 Segmented images

Foreground and background were used to construct a Markov random field across the pixel labels [14–17]. The final step was to apply a graph cut optimization to get the outcome. Figure 3 shows the example of the dataset images.

2.3 Data Augmentation

To increase the diversity of the dataset for adapting the real-world scenario, we expanded the image artificially by modifying the original images. We used flipping, rotation, zooming, scaling, height shift, and width shift. The main concept of the augmentation is to increase the generalization properties, reduce the overfitting problem, and increase the performance accuracy. It mainly helps to create multiple version of the original dataset and model can learn with the more variation of the dataset which helps to recognize the diversity of the pattern and robust performance.

Here, we applied some certain range of the angle for creating rotated image, horizontally and vertically flip the image which is known as flipping [18]. In addition, zoom in and out also used. Beside the recognition, there are many other filed can take help from this technology such as facial recognition, object detection, and medical image analysis to train the system with a small dataset.

2.4 Convolutional Neural Network (CNN)

Recently, many researchers have used CNN to extract the effective feature from the image dataset; they usually use feed-forward networks with multilayer perceptions [18–20]. Generally, CNN architecture consists of the convolutional layer, activation function, pooling layer, dropout layer, and loss functions. The working procedure of the individual layers with a mathematical explanation is given.

Convolutional Layer
The concept behind this layer is to extract features from the operation, namely convolving between the set of learnerable parameters and input images where learnable kernels are known as a weight matrix. Let the input image dimension is $n \times n \times d$, and the dimension of each of M learnable kernels or weight matrix be k × k where height and width are denoted by n and depth is denoted by [21–23]. From the left to the right side of the input image, this pixel-by-pixel operation is performed using a filter, visualized in Fig. 4. During the operation, it ensures that the height and width size must be the same as the input image. To keep the same dimension, zero padding around the inpur data was applied to protect the original image size fom shrinkage. This layer generated its final feature vector using the following formula in Eq. 1.

$$G_x^{(L)} = \sum_{y=1}^{m_1(L-1)} F_{x,y}^{(L)} \times G_y^{(L-1)} + \text{Bias}^{(L)} \tag{1}$$

Here, $G_x^{(L)}$, x_{th}, L, $F^{(L)}$ denoted the output feature, feature map, layer number, and filter, respectively. In addition, kernel and input feature denoted by the m_1, $G_y^{(l-1)}$, respectively.

Max Pooling Layer
Another commonly used CNN layer name is the pooling layer which is normally used after the convolutional layer. The main concept of the layer is to downsample the feature map to reduce the input features' spatial dimensionality. It mainly divides the feature map into different regions, which must not overlap each other. After that, in each region, its spatial operation is computed with specific formulas such as average or maximum values. Sequentially, this function was employed to the whole matrix and produced a finally pooled matrix by combining each rogation output. The carpooling layer's main benefit is reducing the computational cost by downscaling

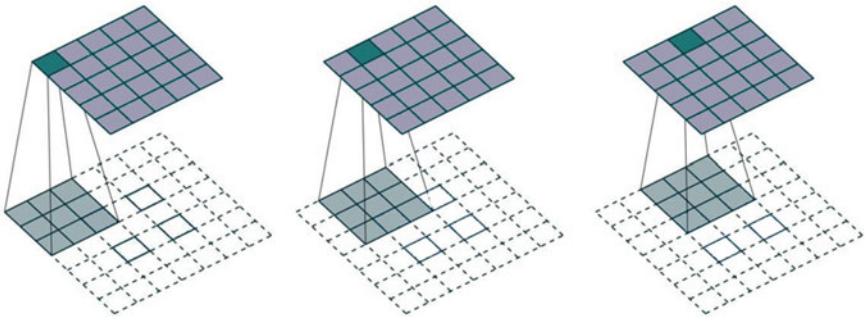


Fig. 4 Sequential operation of convolutional Layer

the features map and reducing the network parameters. In addition, it makes the system a translation invariance because small changes in the input image do not affect its output. This layer is also used to prevent the overfitting of the network. Let our original input feature dimension is $n \times n$, max pooling layer divided into n region. Two different pooling layers, the maximum pooling layer and the average pooling layer, were used in the study. Max pooling layer generated maximum from each region. If the size of the pooling region is 2×2 and the step size is 2, then a sliding window is created with the aim of skipping the height and width using the eschol equation (2)

$$n_H = \left[\frac{n + 2P - f}{2} + 1\right] \tag{2}$$

Global Average Pooling (GAP)

GAP is also another kind of effective usually used as a final layer in the network. In order to improve the generalization properties of the model, this layer can also overcome the overfitting problem of the model. As we said previous subsection, pooling divided the input matrix into the various sub-region. GAP calculated the average of each region and produced a final feature matrix. This operation compresses the spatial dimension of the original feature vector into single value, and sometimes, we can use this layer as a replacement of a fully connected layer [24–26]. Let us consider the dimension of each sub-region is $n \times n \times d$ After applying the GAP for each region, it produced the $1 \times 1 \times d$. Finally, the input leaf data is classified into n neurons representing disease type using the output layer of the CNN model.

Proposed Architecture

This section describes the proposed CNN model and its functionality in each layer. Five convolutional layers, five max pooling layers and a global average pooling layer are used in the proposed model. Finally, the output layer contained three neurons for the three different classes of the dataset. In the study, first, we fed the input image into the first convolutional layer lead to produce the convolved features from the original images. After that, we fed this feature vector into the max pooling layer aiming to downsample the convolutional features maps. We repeated the combination of the convolutional and the max pooling layer. Finally, we used the global average pooling layer to reduce the dimension of the feature vector. In our case, the dimension of the input image was $224 \times 224 \times 3$, which is filtered using the convolutional layer whose kernel size is size $3 \times 3 \times 3$. After that, we fed this feature vector into the first max pooling layer as input. Sequentially, we fed the pooled feature into the second convolutional layer with the same kernel size and max pooling layer and produced a second pooled feature vector. Then used applied third convolutional layer and generated feature fed into the third max pooling layer. To reduce overfitting and improve performance accuracy, we then used a global average pooling layer. Table 1 demonstrates the proposed CNN model's layer name, input shape, output shape, and parameters. Figure 5 illustrates the basic structure of the convolutional neural network.

In the final stage, we applied the dense layer and compiled it using the Adam optimizer with learning rate 0.01 by taking help from the loss function categorical cross-entropy.

Table 1 Proposed CNN layer name and their parameters

Layer No	Layer type	Input shape	Output shape	Param
1	Conv2d_1	$224 \times 224 \times 3$	$224 \times 224 \times 32$	896
2	Max pooling 2d_1	$224 \times 224 \times 32$	$112 \times 112 \times 32$	0
3	Conv2d_2	$112 \times 112 \times 16$	$112 \times 112 \times 16$	4624
4	Max pooling 2d_2	$112 \times 112 \times 16$	$56 \times 56 \times 16$	0
5	Conv2d_3	$56 \times 56 \times 32$	$56 \times 56 \times 32$	4640
6	Max pooling 2d_3	$56 \times 56 \times 32$	$28 \times 28 \times 32$	0
7	Conv2d_4	$28 \times 28 \times 32$	$28 \times 28 \times 64$	18,496
8	Max pooling 2d_4	$28 \times 28 \times 64$	$14 \times 14 \times 64$	
9	Conv2d_5	$14 \times 14 \times 64$	$14 \times 14 \times 128$	73,856
10	Max pooling 2d_5	$14 \times 14 \times 128$	$7 \times 7 \times 128$	0
11	Global average pooling 2D	$7 \times 7 \times 128$	128	0
12	Activation (softmax)	128	3	387
	Total params: 102,899 Trainable params: 102,899 Non-trainable params: 0			

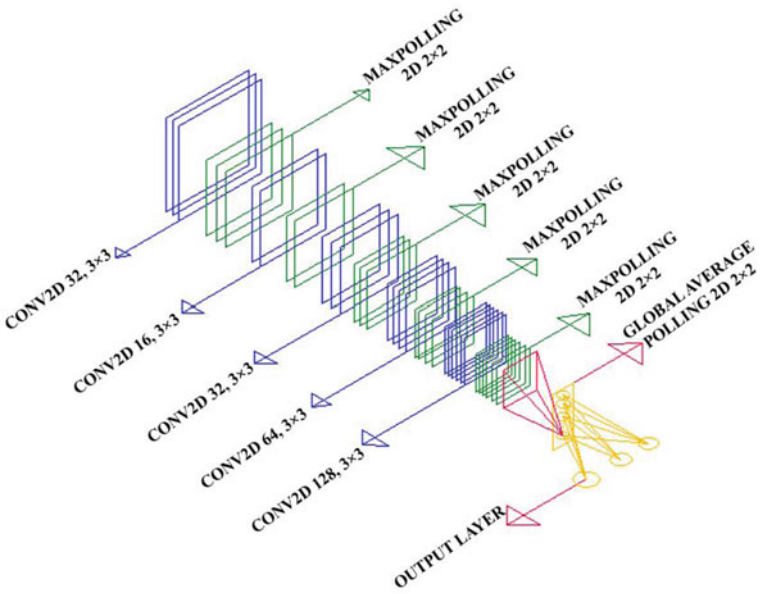


Fig. 5 Structure of the convolutional neural network

3 Experimental Performance

In this study, the system was fed a total of 20,494 dataset images of potato leaves. The dataset is divided into three steps: training, which contains 80% of the images, and testing and validation, which each use 10% of the remaining images. The GrapCut segmentation techniques are used and augmentation the images to increase the dataset images for better performances. The augmentation method makes use of random flip, random rotation, random zoom, random height, and random width. We used the generalized CNN model on different dataset. Table 2 represents the accuracy of training, testing, and validation of the potato leaves disease. The training accuracy of different datasets is 99.90, 99.84, and 99.90% for NLD, NPD, and PD. We also incorporated the whole dataset and applied the suggested approach. As a result, the training, validation, and testing accuracy is 99.93%, 99.96%, and 99.95%, respectively. The average training and testing accuracy of all datasets is 99.89% and 99.98%, respectively. Figure 6 depicts the loss and accuracy comparison of the proposed model.

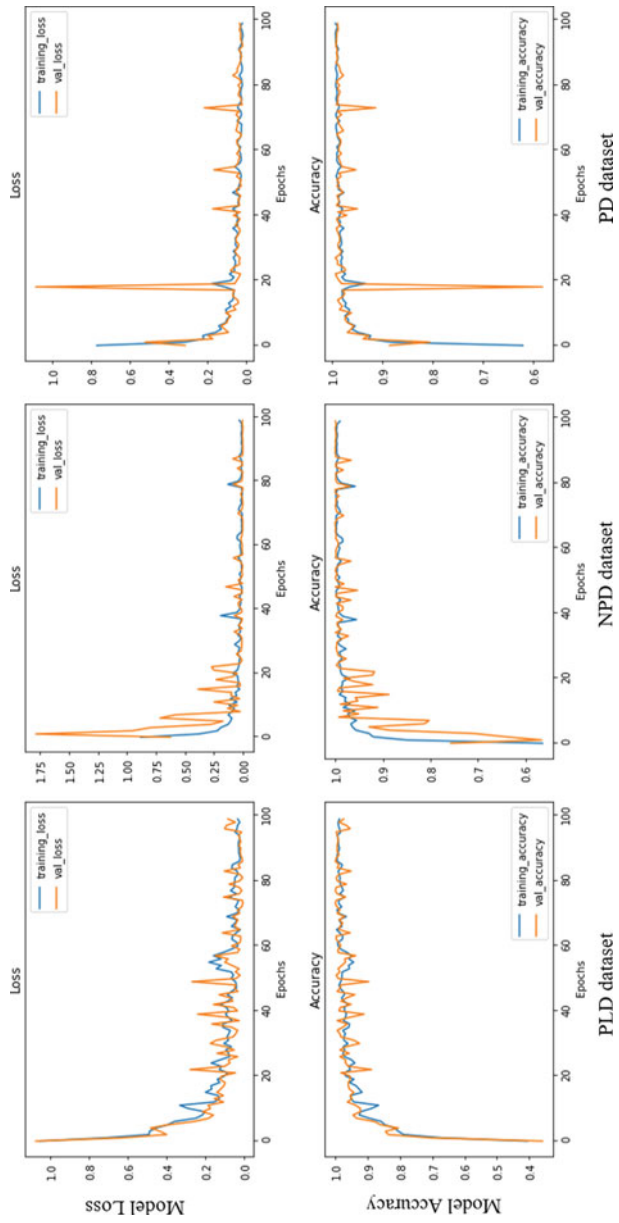


Fig. 6 Model loss and accuracy for different datasets

Table 2 Accuracy for the training, validation, and testing of potato leaf disease diagnosis

Dataset	No. of Images	Augmented images	Training accuracy (%)	Validation accuracy (%)	Testing accuracy (%)
PLD	1500	7500	99.90	100	100
NPD	7118	35,590	99.84	99.93	100
PD	9550	47,750	99.90	99.79	100
Combined	18,168	90,840	99.93	99.96	99.95
Average			99.89	99.92	99.98

Table 3 Comparison accuracy

References	Methods	Reported accuracy (%)
[9]	VGG16	97.89
[10]	YOLOv5	99.75
[11]	Random forest	90.00
Proposed	CNN	99.89

The accuracy of the state-of-the-art approaches is compared in Table 3. According to the comparison results, our proposed technique has attained the highest accuracy using the proposed CNN model.

4 Conclusion

In this research, we employ a generalized deep learning model in association with a segmentation and augmentation process to detect and classify potato leaf illnesses. The effectiveness of the proposed technique was measured against three benchmark datasets. The evaluation is also carried out using the combined dataset. The images in the dataset were segmented using GrabCut. The dataset was augmented using random flip, random rotation, random zoom, random height, and width methods. Even yet, the average accuracy achieved in both training and testing was 99.89 and 99.98%. In future, we have a plan to implement the leave disease detection method in real-time in the Android, another version.

References

1. Khoury CK (2016) Origins of food crops connect countries worldwide. Proc R Soc B: Biol Sci 283:20160792. https://doi.org/10.1098/rspb.2016.0792
2. Shrivastava VK, Pradhan MK (2021) Rice plant disease classification using color features: a machine learning paradigm. J Plant Pathol.103(1):17–26

3. Carlson A, Klonsky H, Karen L, Pete (2010) Potatoes. https://doi.org/10.1007/978-0-387-92765-7_1
4. Hou C, Zhuang J, Tang Y, He Y, Miao A, Huang H, Luo S (2021) Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation. J Agri Food Res 5:100154
5. Chaki J, Dey N (2018) A beginner's guide to image preprocessing techniques. CRC Press
6. Chaki J, Dey N (2019) A beginner's guide to image shape feature extraction techniques. CRC Press
7. Chaki J, Dey N, Moraru L, Shi F (2019) Fragmented plant leaf recognition: bag-of-features, fuzzy-color and edge-texture histogram descriptors with multi-layer perceptron. Optik 181:639–650
8. Miah ASM, Hasan MAM, Jang S, Lee H, Shin J (2023) Multi-stream graph-based deep neural networks for skeleton-based sign language recognition. Preprints org. 2023050467. https://doi.org/10.20944/preprints202305.0467.v1
9. Chakraborty KK, Mukherjee R, Chakroborty C, Bora K (2022) Automated recognition of optical image based potato leaf blight diseases using deep learning. Physiol Mol Plant Pathol 1(117):101781
10. Rashid J, Khan I, Ali G, Almotiri SH, AlGhamdi MA, Masood K (2021) Multi-level deep learning model for potato leaf disease recognition. Electronics 10(17):2064
11. Vallabhajosyula S, Sistla V, Kolli VK (2022) Transfer learning-based deep ensemble neural network for plant leaf disease detection. J Plant Diseases Prot129(3):545-58
12. Yeswanth PV, Khandelwal R, Deivalakshmi S (2023) Super resolution-based leaf disease detection in potato plant using broad deep residual network (BDRN). SN Comput Sci 4(2):1-9
13. Basavaiah J, Arlene AA (2020) Tomato leaf disease classification using multiple feature extraction techniques. Wirel Per Commun.115(1):633–51
14. Miah ASM, Shin J, Hasan MAM, Rahim MA (2022) Bensignnet: Bengali sign language alphabet recognition using concatenated segmentation and convolutional neural network. Appl Sci 12(8):3933. Apr 2022 [SCI indexed]
15. Miah ASM, Shin J, Hasan MAM, Rahim MA, Okuyama Y (2023) Rotation, translation and scale invariant sign word recognition using deep learning. Comput Syst Sci Eng 44(3):2521–2536
16. Rahim MA, Miah ASM, Shin J (2020) Hand gesture recognition based on optimal segmentation techniques in human-computer interaction. In: 3rd IEEE international conference on knowledge innovation and invention (IEEE ICKII 2020), K200134, Kaohsiung, Taiwan
17. Rother C, Kolmogorov V, Blake A (2004) GrabCut, interactive foreground extraction using iterated graph cuts. ACM Tran graphics (TOG). 23(3):309–14
18. Miah ASM, Hasan MAM, Shin J (2023) Dynamic hand gesture recognition using multi-branch attention based graph and general deep learning model. In IEEE Access. https://doi.org/10.1109/ACCESS.2023.3235368
19. Miah ASM, Hasan MAM, Shin J, Okuyama Y, Tomioka Y (2023) Multistage spatial attention-based neural network for hand gesture recognition. Computers 12(1):13. https://doi.org/10.3390/computers12010013
20. Shin J, Miah ASM, Hasan MAM, Hirooka K, Suzuki K, Lee HS, Jang SW (2023) Korean sign language recognition using transformer-based deep neural network. Appl Sci 13(5):3029. https://doi.org/10.3390/app13053029
21. Carneiro ALC, Silva LB, Salvedeo DP (2021) Efficient sign language recognition system and dataset creation method based on deep learning and image processing
22. Awan MJ, Rahim MSM, Salim N, Rehman A, Nobanee H, Shabir H (2021) Improved deep convolutional neural network to classify osteoarthritis from anterior cruciate ligament tear using magnetic resonance imaging. J Pers Med 11:1163. https://doi.org/10.3390/jpm11111163
23. Fregoso J, Gonzalez CI, Martinez GE (2021) Optimization of convolutional neural networks architectures using PSO for sign language recognition. Axioms 10:139.https://doi.org/10.3390/axioms10030139

24. Jagtap S, Bhatt C, ThikJ, Rahimifard S (2019) Monitoring potato waste in food manufacturing using image processing and internet of things approach. Sustainability 11:3173. https://doi.org/10.3390/su11113173
25. Lin M, Chen Q, Yan S (2013) Network in network. arXiv preprint arXiv:1312.4400
26. Shustanov A, Yakimov P (2019) Modification of single-purpose CNN for creating multi-purpose CNN. J Phys Conf Ser 1368:052036. https://doi.org/10.1088/1742-6596/1368/5/052036

A Perspective of 5G Network and Associated Impact in India

Sanjay Gour, Hemant Sahu, Neetu Sharma, and Janki Barot

Abstract The 5G network technology in India holds immense potential for driving sustainable development across various sectors. This increased accessibility which leads to positive outcomes in education, health, and economic opportunities, thereby fostering sustainable development. The 5G technologies can enable the establishment of smart cities, promoting efficient resource utilization, reduced emissions, and an enhanced quality of life as smart traffic management, smart lighting such kind of advancements contribute to mitigating the environmental impact of urbanization and supporting sustainable development. Deployment of 5G technology facilitates the development of new industries and innovative corporate mockups, particularly through the acceptance of developments like the Internet of Things (IoT). The 5G approach in India has the impending to significantly contribute to sustainable development by ensuring sustainable and responsible deployment; it can drive positive outcomes in various sectors, improve accessibility to essential services, foster resource efficiency, and stimulate innovation for a more sustainable future. The present paper is an attempt to provide perspective of 5G network and its associated impact in India.

Keywords 5G network · IoT · Communication · Technologies · Internet · Development

S. Gour (✉)
Department of Computer Science and Engineering, Jaipur Engineering College and Research Centre, Jaipur, India
e-mail: sanjay.since@gmail.com

H. Sahu
JRN Rajasthan Vidyapeeth, Udaipur, India

N. Sharma
Department of Computer Science and Engineering, Engineering College Ajmer, Ajmer, India
e-mail: drneetu@ecajmer.ac.in

J. Barot
Silver Oak University, Ahmedabad, India

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_53

1 Introduction

The 5G network technology in India has the impending to contribute to sustainable development in several ways. Initially, 5G networks can improve connectivity and provide faster internet speeds, which can increase access to education, health care, and other essential services. This can lead to improved outcomes in terms of education, health, and economic opportunities, thus contributing to sustainable development. Moreover, 5G technologies can enable the development of smart cities, which can lead to more efficient use of resources, reduced emissions, and improved quality of life. For example, smart traffic management systems can reduce congestion, while smart lighting systems can reduce energy consumption [4]. These benefits can help to reduce the environmental impact of urbanization and support sustainable development. Further, 5G tools can enable the growth of new industries and innovative business models as it can facilitate the adoption of know-hows such as the IoT, which can enable more efficient use of resources in industries such as agriculture and manufacturing. This can contribute to sustainable development by reducing waste, increasing productivity, and improving resource efficiency. Though, the deployment of 5G technology also raises concerns around its potential environmental impact, such as increased energy consumption and electronic waste. Therefore, it is significant to guarantee that the deployment of 5G networks in India is done in a sustainable manner, taking into account the environmental and social impact of the technology. The overview of 5G technology in India has the impending to contribute to sustainable development, provided that it is implemented in a sustainable and responsible manner [2].

2 5G with Telecommunications

The 5G technology represents the most recent advancement in mobile communications, providing a substantial increase in speed, capacity, and low latency compared to previous generations like 4G and 3G. Its introduction promises to deliver a host of welfares, with quicker move and, improved reliability, and better network ability. This, in turn, will permit the development of a broader variety of solicitations and facilities, like computer-generated and amplified realism, and linked automobiles. Furthermore, 5G is probable to perform a pivotal part in facilitating the growth of the IoT, which will allow billions of devices to connect to the internet.

In standing of its monetary influence, 5G technology is probable to drive important development and employment occasions in the telecommunications segments. It has the potential to enhance efficiency and cost cutting and bolster the keenness of trades. Additionally, 5G network approach has the prospective to stimulate novel occupational models and income brooks, encouraging novelty and development in various businesses. Ultimately, the importance of 5G deceits in its capability to deliver quicker, extra consistent, and extra well-organized wireless connectivity,

transforming how we live, work, and interact with the world around us. Its potential for generating substantial economic benefits, enabling new applications and amenities, and encouraging novelty and progress across multiple sectors underscores its significance [3].

3 Comprehensive 5G System

Humans constantly pursue advancements in speed, size, and performance. Among the existing mobile communication standards like 2G, 3G, and 4G, 5G takes the lead. It sets the new benchmark for communication standards, offering even faster data downloads from the internet. The 5G enables rapid communication among devices, sensors, and machines within the Internet of Things (IoT). Compared to 4G, 5G networks are meaningfully quicker, with top data transmission up to 20 Gbps. They boast an impressive reliability rate of up to 99.99%, ensuring seamless data transmission deprived of delays and an incredibly short latency of just one millisecond. The comprehensive 5G arrangement comprises three key components [7].

Enhanced Mobile Broadband (eMBB): eMBB focuses on delivering higher data rates and bandwidth, catering to applications that rely on protracted network.

Ultra-reliable Low Latency Communications (uRLLC): uRLLC delivers exceptionally short latency and consistency care, specially targeting machine-to-machine transportations. It serves critical applications such as far-flung medical surgeries and actions, wireless mechanism of manufacturing, and latency-sensitive autonomous vehicles.

Massive Machine-Type Communications (mMTC): mMTC facilitates the concurrent gathering of a vast size of slight data containers as of numerous devices. Its main objective is to provide connectivity for a huge amount of apparatuses. Through IoT implementation, mMTC reduces power feeding and enhances overall effort efficiency (Figs. 1, 2, 3, 4, and 5).

With its superior capabilities, 5G technology opens up new possibilities and applications, revolutionizing various industries and driving technological progress [8, 9].

4 5G in India

5G technologies are poised to revolutionize India's telecommunications landscape. The introduction of 5G technology is expected to bring numerous benefits to become possible. In India, the government has been working to create a favorable environment for the adoption of 5G technology. This includes allocating spectrum for 5G networks,

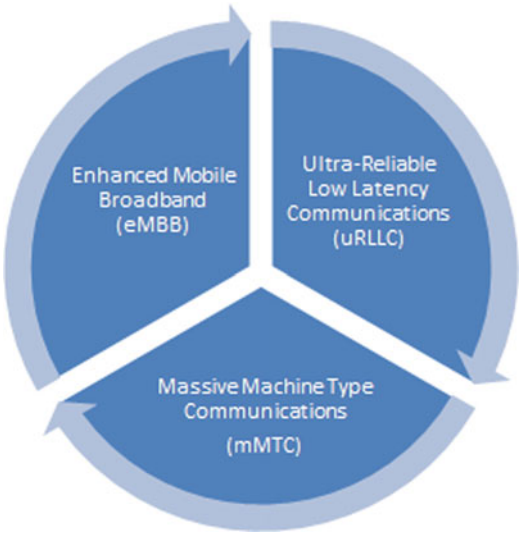


Fig. 1 Comprehensive 5G system component

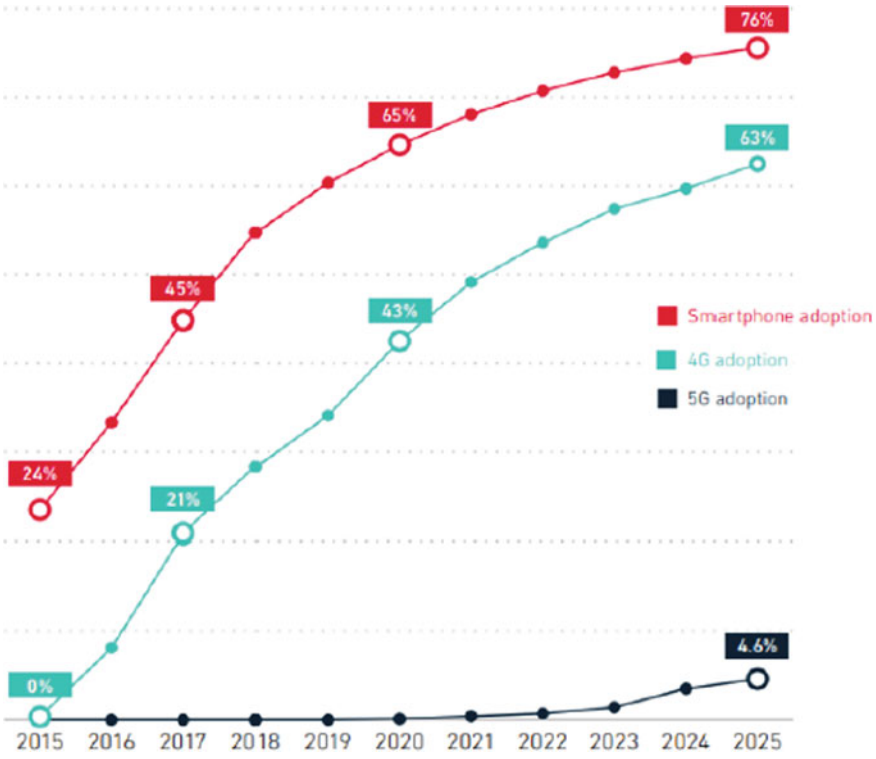


Fig. 2 4G–5G and smartphone adoption in India. *Source* GSMA intelligence

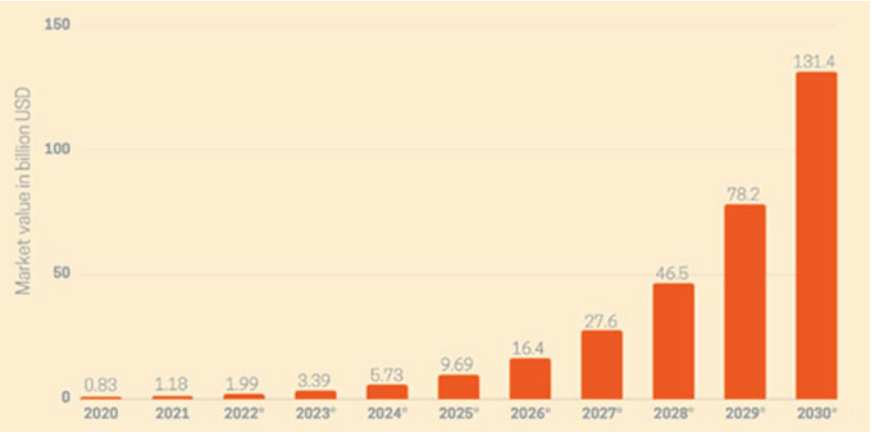


Fig. 3 Size of 5G infrastructure market worldwide (2020–2030). *Source* https://startuptalky.com/5g-india-launch-iot-impact

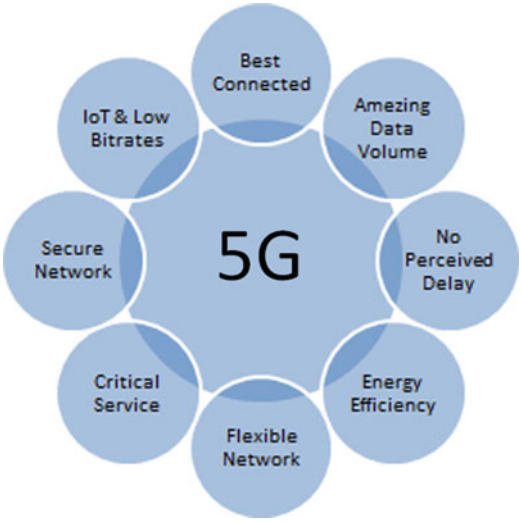
Rank	Type of Industry	Sales ($B)	Share of Industry Sales (%)
1	Manufacturing	4,687	5.4
2	Information & Communication	1,569	10.7
3	Wholesale & Retail Sales	1,198	5.1
4	Public Services	985	6.3
5	Construction	731	4.3

Fig. 4 Top 5 industries which are expected to gain maximum profit. *Source* https://stl.tech/blog/role-of-5g-in-the-digital-economy-and-how-it-is-impacting-the-industry/

streamlining regulations, and incentivizing investment in 5G infrastructures. Telecom operators in India have also been actively testing and deploying 5G networks in various cities across the country [2]. The adoption of 5G technology in India is anticipated to drive noteworthy development and occupation occasions. Moreover, the introduction of 5G technology is expected to support IoT, which is projected to perceive billions of gadgets allied with network in the impending years [5].

Despite these promising developments, encounters to the adoption of 5G in India remain. This includes issues such as high costs of infrastructure and equipment, lack of skilled labor, and concerns over security and privacy. As of cutoff of September 2021, the following are some statistics on the growth of 5G in India.

Fig. 5 5G network impact



- In August 2021, the Indian government announced that it would be auctioning spectrum for 5G networks in the country. The auction was expected to take place in early 2022.
- As of August 2021, several telecom operators in India, including Jio, Airtel, and Vodafone Idea, had started testing 5G networks in various cities across the country.
- According to a report by the Telecom Regulatory Authority of India (TRAI), India had over 795 million wireless subscribers as of May 2021. This indicates a significant potential market for 5G services in the country.
- In a report by the Global System for Mobile Communications Association (GSMA), it was estimated that India would have 350 million 5G connections by 2026, representing around 27% of total mobile connections in the country.
- The Indian government has announced various measures to incentivize investment in 5G infrastructure, including a Production Linked Incentive (PLI) scheme, which offers financial incentives to companies that manufacture 5G equipment in the country.

The 5G technology is still in its early stages of development in India, and there are positive signs that it will play a significant part in the nation's telecommunications landscape [11, 12].

5 5G and Indian Mobiles

According to the earlier GSMA report, Indian mobile operators' capital expenditure (Capex) investment is projected to reach $8.1 billion in the year 2020, as replacement of figure of $5.7 billion in 2017, indicating a negative compound annual growth rate

(CAGR) of -11%. This decline in Capex by telecom operatives like BSNL, Bharti Airtel, Idea, Vodafone, and Reliance Jio may pose challenges for Huawei, Ericsson, Nokia, ZTE, Qualcomm, Intel, and Samsung in the Indian telecom market [11].

However, there is positive news regarding the fast development in the income of telecom industries. Operator income is expected to increase to $30 billion in the year 2025, rising since $25 billion in 2017, with a CAGR of 1.9%. On the downside, the development of 5G networks in India is sluggish. By 2025, 5G connections are estimated to represent only 4.6% of the entire networks in the country. In contrast, 4G influences are projected to cultivate from 21 in 2017 to 63% of entire influences in 2025. Additionally, 3G connections will increase from 15 in 2017 to 18% in 2025. This indicates that the substantial investment made by national mobile operators in 3G systems and band did not yield sufficient revenues, mainly due to intense struggle [1].

The report also highlights that India, while still predominantly a 2G market, is undergoing a rapid transformation. The 4G link base is expected to triple to 890 million by 2025, with 4G accounting for 6 out of 10 connections. Consequently, the share of 2G connections is anticipated to decline from 65 in 2017 to 15% in 2025, as the country progresses toward advanced mobile technologies.

6 IoT and 5G in India

To elevate its digital revolution to new heights, India is gearing up to launch 5G services across the country. The fifth-generation wireless technology will revolutionize connectivity, providing higher speeds, lower latency, and greater capacity compared to 4G LTE. With these advancements, 5G will unlock immense potential for IoT, AI, AR/VR, and private LTE, consequently supporting Industry 4.0. This will enable IoT service providers to offer users a seamless and enhanced experience.

According to reports, the number of IoT connections is projected to triple, reaching 25 billion, and global IoT revenue is expected to reach $1.1 trillion by 2025. In line with India's goal of becoming a $5 trillion economy by 2025, increased investment in the expansion of IoT and 5G would be a wise decision. McKinsey suggests that effectively implementing 5G technologies alongside IoT in sectors like mobility, health care, manufacturing, and retail could boost the global GDP by $1.2 to $2 trillion by 2030, potentially impacting India by $1 trillion by 2035. Projections from the International Data Corporation (IDC) indicate that there will be 41.6 billion IoT devices by 2025. This surge in numbers highlights the need for launching 5G services in India to meet the demand for reliable wireless connections with ultra-low latency, high network capacity, and improved availability. With its stable connections, reduced latency, and network hyper-densification, 5G will ensure seamless IoT functionality and optimize network energy consumption [10, 13].

In the domain of heavy machinery, 5G's advanced data transmission and precise technology will enable the utilization of IoT for enhanced employee safety, particularly in hazardous environments. It will also facilitate better control over remotely

operated machinery. The medical industry is poised for revolution with the combination of IoT and 5G. Through extensive technological integration, tele-surgery and remote surgery will become commonplace. With speeds peaking at 10 Gbps, revolutionary changes in IoT management are expected. Additionally, 5G can penetrate diagnostic systems to facilitate faster and more accurate diagnoses and reports.

The rollout of 5G will have a catalytic effect on the development of automatic and autonomous vehicles. Beyond GPS and driving, the benefits of low latency and high connectivity will have far-reaching implications, including a significant reduction in accidents and associated risks. The unique sensing and connecting capabilities of 5G can detect and prevent potential roadblocks, positively impacting overall traffic conditions.

In the realm of domestic spaces, the convergence of 5G and IoT will revolutionize the functioning of regular households. They can be effectively employed for cleaning and cooking tasks. Floor cleaners and home robots have already become common in upper-middle-class households, but with 5G integration, these devices will become smarter and more reliable. Automated gates, cameras, and other security systems will greatly enhance home safety measures [6].

7 Industrial 5G Network Approach

Industrial 5G networks possess the full capacity to meet the growing demands of data and production within companies. These networks are instrumental in implementing crucial utilize cases in the area of manufacturing, automation support with robotics to provide expert resolutions. Industries are now looking toward end-to-end product tracking approach. Simultaneously, real-time approach is utilized to assess and monitor the productions. These networks hold substantial socio-economic value and are projected to generate $13.2 trillion in income through worldwide trades' activities by 2035 [14].

Certain industry sectors are expected to benefit significantly from 5G technology like manufacturing business will spectator advancements in shrewd factories due to the integration of IoT and automatic built-up. Likewise, the agility subdivision will experience the advantages of independent driving, extensive car-infotainment, and linked journey through telematics.

In the healthcare industry, where data storage and patient safety are crucial, 5G services will play a vital role. Distant patient observing, far-flung surgery, and the implementation of AR/VR-enabled healthcare elucidations will become widespread. The 5G brings benefits such as informal and rapid transmission of big imaging files and altitudinal computing. In the energy sector, industries can enhance their capabilities with smart grid and drone-based watching.

Study suggests that by 2023, approximately one billion users will have access to 5G technology, enabling data transfer speeds of up to 1 Gbps. Qualcomm surveys indicate that investments worth around $12 trillion will be made in 5G-related services by 2035, including improved mobile broadband services, mission-critical

service area, and noteworthy IoT progressions in several sectors. The 5G worth shackle in a digital world is estimated to produce incomes about to $13.2 trillion by means of the year 2035, potentially creating 22 million jobs, as projected by Qualcomm.

8 5G Impact on Digital Economy

Manufacturers and telecom operators have a significant occasion to establish shrewd industrial unit utilizing 5G networks and leveraging smart technologies.

The manufacturing sector, along with energy and utilities, holds great promise as part of the Industry 4.0 movement, driving the subsequent movement of manufacturing throughput. Industry 4.0 envisions connected production systems with components that can be followed and constructed within a self-managing digital economy. The 5G has the potential to impact various services and enhance the digital economy in several areas.

1. Internet of Things (IoT): 5G will significantly enhance connectivity, speed, efficiency, and user experience for devices and sensors, allowing them to communicate more effectively. This will generate greater profitability for telecom service providers and business owners by improving IoT capabilities.
2. Connected Vehicles: With 5G, connected vehicles can operate autonomously and receive precise, on-time data about vehicle. This information will enable automobile manufacturers and vendors to provide healthier sustenance and a superior experience.
3. Better Governance: Governments can leverage 5G to develop robust smart city architectures and e-governance frameworks. This will lead to improved public services, streamlined processes, and real-time information exchange during emergencies. It paves the way for proactive and connected governments that prioritize class public provision experiences.

The upcoming networks will produce worth across various diligences, with innovative applications in different settings. For instance, hospitals can utilize 5G-enabled devices for remote patient monitoring and deploy shrewd ambulances that provide on-time connectivity. We can expect seamless monetary dealings via digitized wallets. Companies will become 5G supported, connecting gadgets and additional sensors. When united with AI, IoT, and edge computing, along with extended reality (XR), 5G empowers industries to fully harness the paybacks of these approaches.

9 Conclusion

As per the study or projected paper, it is noted that 5G network in the India is a wave of revolution which is depicted in the similar impact of the rest of the world. This is a new age for the upcoming development with rapid pace. It impacts on manufacturing, telecommunication, wholesale, retails, constructions, and public services. It is found that 5G is base for development of every aspect of the nation as this is age of information technology. The scenario of 5G in Indian context plays key role in the industrial development with reference smart cities, IoT, controlling system, and automated production. The present study gives satisfactory insights about the development through 5G network with special reference of India.

References

1. Abdullah M, Altaf A, Anjum MR, Arain ZA, Jamali AA, Alibakhshikenari M, Falcone F, Limiti E (2021) Future smartphone: MIMO antenna system for 5G mobile terminals. IEEE Access 9:91593–91603. https://doi.org/10.1109/ACCESS.2021.3091304
2. Nandi A (2023) 5G technology in India. Int J Eng Res Technolo. 12(02)
3. Anwar S, Prasad R (2018) Framework for future telemedicine planning and infrastructure using 5G technology. Wirel Pers Commun 100:193–208
4. Chataut R, Akl R (2020) Massive MIMO systems for 5G and beyond networks—Overview, recent trends, challenges, and future research direction. Sensors 20:2753. https://doi.org/10.3390/s20102753
5. Ray D, Oza S, Sharma A, Abhishek K (2020) 5G in India. Int J Eng Res Technol 9(04)
6. Hui H, Ding Y, Shi Q, Li F, Song Y, Yan J (2020) 5G network-based internet of things for demand response in smart grid: a survey on application potential. Appl Energy 257:113972. https://doi.org/10.1016/j.apenergy.2019.113972
7. Dangi R, Lalwani P, Choudhary G, You I, Pau G (2022) Study and investigation on 5G technology: a systematic review. Sensors (Basel) 22(1):26. MDPI
8. 5G physical layer principles models and technology components 2018. In: Zaidi, Fredrik A (eds), Elsevier
9. Trick U (2021) 5G an introduction to the 5th generation mobile networks, De Gruyter
10. India Automated How the fourth industrial revolution is transforming India
11. https://www.insightsonindia.com/science-technology/communication-and-it-technology/cellular-networks/5g-technology/
12. https://www.insightsonindia.com/2022/01/05/where-does-the-5g-rollout-stand/
13. https://startuptalky.com/5g-india-launch-iot-impact/
14. https://stl.tech/blog/role-of-5g-in-the-digital-economy-and-how-it-is-impacting-the-industry

Lung Cancer Prediction Using Machine Learning Models

Manish Singh, Chintan Shah, and Premal Patel

Abstract Lung cancer has grown into an awful disease in recent years that places people's lives and good health in risk. Researchers and doctors have to use effective predictive modeling. In order to find some patterns in the data, this study provides a naive Bayes-based lung cancer forecasting approach that looks at the association among a variety of readily available markers (like age, smoking, alcohol use, discomfort in the chest, etc.) and lung cancer. Using a medical database and the technique can find a simple and intelligible lung cancer model, greatly lowering the chance of illness. Naive Bayes classifiers are a family of simple probabilistic classifiers that make use of strong (naive) independence assumptions for applying the theorem of Bayes to an array of characteristics. It demonstrates that the flawed Bayes algorithm may significantly mitigate the issue and it can accurately forecast how the frequency of these indications will alter the probability of developing lung cancer. Additionally, compared to the support vector machine approach, naive Bayes provides a better level of precision in prediction.

Keywords Lung cancer · Prediction model · Naïve Bayes

1 Introduction

Computer systems can "learn" from data or gradually enhance performance on specific tasks without being explicitly programmed thanks to a branch of computer science that is called machine learning. The research and creation of technologies that can derive information from data and forecast outcomes are the main goals of machine learning. By building a model from sample inputs, these algorithms go above simply adhering to rigidly static program instructions to arrive at data-driven

M. Singh (✉) · C. Shah · P. Patel
Department of Computer Engineering, College of Technology, Silver Oak University, Ahmedabad, Gujarat, India
e-mail: mani274u@gmail.com

P. Patel
e-mail: premalpatel.ce@socet.edu.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_54

predictions or decisions. In a wide range of computing applications where it can be difficult or expensive to develop and execute explicit techniques with acceptable performance, machine learning is employed [1, 2].

A free and open-source integrated development environment (IDE) for R, a programming language utilized for statistical computing and graphics, is known as JUPYTER. JUPYTER is offered in two editions: JUPYTER Server, which enables web access to JUPYTER while it is operating on a distant Linux server, and JUPYTER Desktop, which runs locally as a conventional desktop application. For Windows, macOS, and Linux, prepackaged distributions of JUPYTER Desktop are accessible [3].

On desktop machines operating Windows, macOS, and Linux in addition to in a web browser utilizing JUPYTER Server or JUPYTER Server Pro, JUPYTER comes in both free and price variants. Additionally, JUPYTER allows us to do forecasting studies with machine learning techniques [4, 5].

2 Problem Definition

Medical databases comprise an important portion of information generated by modern medicine. Today, everybody in the world is dealing with worsening health issues that call for immediate intervention. Because instances of lung cancer continue to rise in patients due to high tobacco use and puffing, this can be extremely difficult for physicians to make decisions while predicting lung cancer in patients at early stages. The best method to assess the likelihood of a disease is to use machine learning to build a model for forecasting using data obtained in the hospital regarding the patients suffering from lung cancer. This is due to the reality because it is taboo in some countries and individuals are hesitant to ask for a diagnosis [2, 3].

3 Proposed Methodology

3.1 Dataset Description

In an effort to boost the anticipated accuracy of data mining algorithms, a larger and more precise dataset was used in this study. To get the best outcomes from data mining, it is essential to carefully control the features of the sensations utilized to diagnose illness. When predicting lung cancer, a variety of factors are taken into thought, such age, gender, yellow finger, anxiety, peer pressure, chronic illness, exhaustion, allergy, alcohol use, cigarette smoking, discomfort in the chest, blood coughing, feeling short of breath, wheezing, and trouble swallowing. JUPYTER uses a variety of methods for reducing features, classification, and data pre-processing,

Table 1 Factors of lung cancer

Attributes
AGE
GENDER
YELLOW FINGER
ANXIETY
PEER PRESSURE
SMOKING
CHRONIC LUNG DISEASE
ALCOHOL CONSUMPTION
WHEEZING
COUGHING OF BLOOD
FATIGUE
SWALLOWING DIFFICULTY
SHORTNESS OF BREATH
CHEST PAIN
ALLERGY

including naive Bayes and support vector machines. The efficacy of various lung cancer therapies is evaluated using a confusion matrix [3, 4] (Table 1).

3.2 Methodology Description

To increase the level of accuracy of lung cancer models, the initial data which were given were processed to eliminate any anomalies and generate preprocessed data. The data has been adjusted to decrease the error rate as was verified for any missing values. The preprocessed data is subsequently divided into two separate datasets, each that has 30% test sets and 70% train datasets [5, 6].

The suggested approach creates several prediction models based on methods such naive Bayes and support vector machine using the train dataset. For all of the preceding models, including naive Bayes and support vector machine models, test datasets are also provided to these models, and confusion matrices are generated to assess how accurately these models perform [7] (Fig. 1).

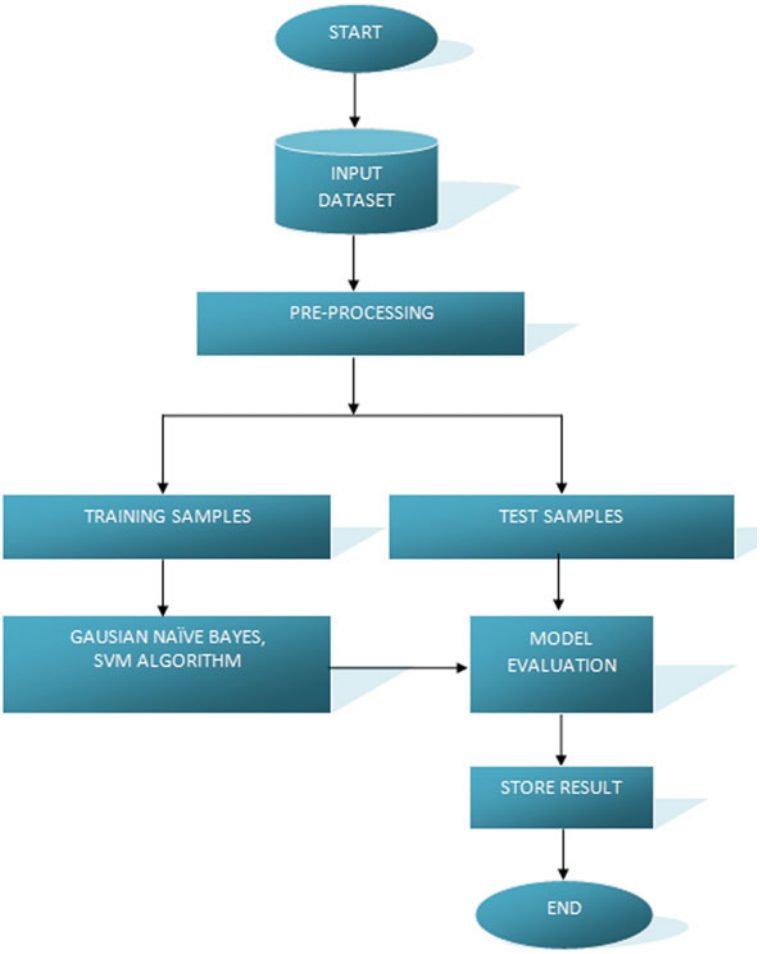


Fig. 1 Proposed model

4 Performance Analysis

The screenshots illustrate the algorithm's efficacy by comparing the accuracy rates of several methods, like support vector machine and naive Bayes. The accuracy of the system that was suggested is 89% when SVM is used, as opposed to 87% when the other techniques were used. The performance research reveals a higher degree of forecasting accuracy for lung cancer [8, 9] (Fig. 2; Table 2).

Fig. 2 Confusion matrix

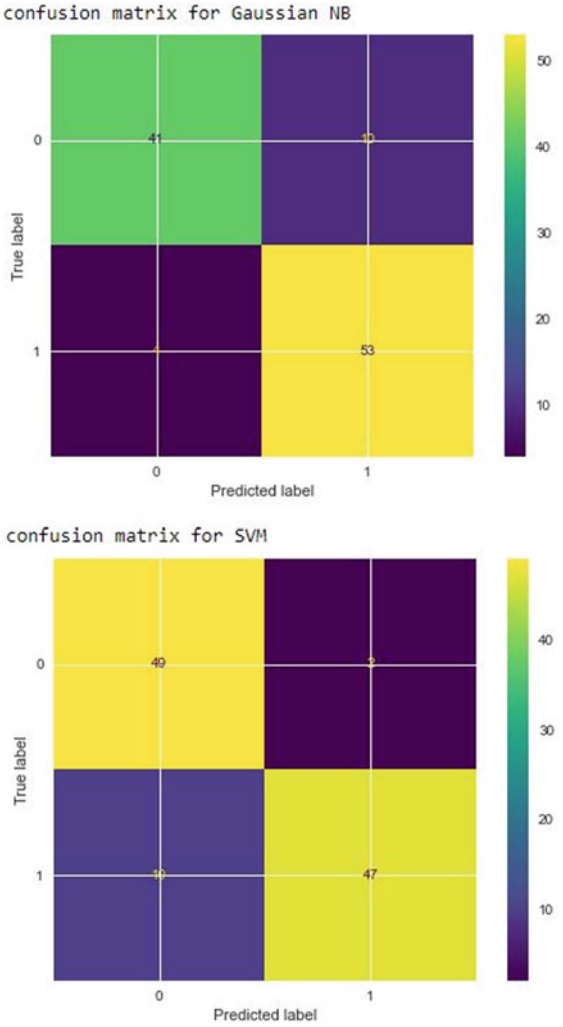


Table 2 Accuracy of the algorithms

Naïve Bayes	89%
SVM	87%

5 Future Work and Conclusion

This article explores machine learning in healthcare administration, which is distinct from the other two sectors in that it works with personal medical data, which is subjected to multiple ethical, moral, and social constraints. The experiment, which

made use of the JUPYTER tool, used a number of machine learning classification techniques. The SVM method outperformed all of the GNB classification techniques, it was found. It is possible to develop and enhance a technique for predicting cancer of the lungs. Additionally, multiple approaches to machine learning might be incorporated [10].

References

1. Xu W, Zhang J, Zhang Q, Wei X (2017) Risk prediction of type II diabetes based on random forest model. IEEE
2. Vijayarani S, Dhayanand S (2015) Liver disease prediction using SVM and naïve Bayes algorithms. IJSETR
3. Schiller HB, Montoro DT, Simon LM, Rawlins EL, Meyer KB, Strunz M, Vieira Braga FA, Timens W, Koppelman GH, Budinger GS et al (2019) The human lung cell atlas: a high-resolution reference map of the human lung in health and disease. Am J Respir Cell Mol Biol 61:31–41
4. Hervier B, Russick J, Cremer I, Vieillard V (2019) NK cells in the human lungs. Front Immunol 10:1263
5. Barroso AT, Martín EM, Romero LMR, Ruiz FO (2018) Factors affecting lung function: a review of the literature. Arch Bronconeumol 54:327–332
6. Mirza S, Clay RD, Koslow MA, Scanlon PD (2018) COPD guidelines: a review of the 2018 GOLD report. Mayo Clin Proc 93:1488–1502
7. Dotan Y, So JY, Kim V (2019) Chronic bronchitis: where are we now? Chronic Obstr Pulm Dis J COPD Found 6:178
8. Stern J, Pier J, Litonjua AA (2020) Asthma epidemiology and risk factors. In: Seminars in immunopathology, vol 42. Springer, Berlin/Heidelberg, Germany, pp 5–15
9. Bell SC, Mall MA, Gutierrez H, Macek M, Madge S, Davies JC, Burgel PR, Tullis E, Castaños C, Castellani C et al (2020) The future of cystic fibrosis care: a global perspective. Lancet Respir Med 8:65–124
10. Mandell LA, Niederman MS (2019) Aspiration pneumonia. N Engl J Med 380:651–663

Automated GUI Testing for Enhancing User Experience (UX): A Survey of the State of the Art

Parth S. Deshmukh, Saroj S. Date, Parikshit N. Mahalle, and Janki Barot

Abstract Recent developments in the area of Artificial Intelligence (AI) have broadened opportunities for users to interact with various technologies. In this context, many enterprises have been shifting their priorities to offering distinctive user experiences (UX) from purely concentrating on effective production and distribution of products/applications. They believe that creating a good user experience is the key to develop efficient and persuasive products/applications. The survey in this paper considers the user experience of mobile applications. Many times mobile application users encounter bugs which may affect the functionality of applications and it is not a good user experience. This survey is based on recognizing and analyzing existing research work that outlines various approaches used for mobile GUI testing tasks to enhance user experience. It is carried out by collecting similar research publications from reputed journals and conferences from 2018 till date. To present an organized literature survey, Research Questions (RQs) are framed. Research questions provide a clear focus and direction for the review paper. They help to define the scope of review and may ensure that the review paper addresses relevant and significant aspects of the topic. To present a detailed analysis, the results of Research Questions are collated. In the end, conclusions and future work are presented.

Keywords User experience · User experience using machine learning · Mobile applications testing · GUI testing · UX · UX practices · UX/UI using ML · UXML

P. S. Deshmukh (✉)
Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India
e-mail: parth.22010945@viit.ac.in

S. S. Date
Department of CS and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India
e-mail: csit.ssd@bamu.ac.in

P. N. Mahalle
Department of AI and DS, Vishwakarma Institute of Information Technology, Pune, India
e-mail: parikshit.mahalle@viit.ac.in

J. Barot
Silver Oak University, Ahmedabad, India
e-mail: jankibarot.ce@socet.edu.in

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4_55

1 Introduction

The rapid development in mobile Internet technologies causes millions of users to get information and services through various applications installed on mobile devices. They access a variety of applications from different domains such as educational apps, social media apps, game apps, entertainment apps, healthcare apps, and lifestyle apps. These apps constitute user interface (UI) and user experience (UX). Enterprises are beginning to realize that users' experiences have a significant influence on users' loyalty. Therefore, to improve the user experience is one of their best business opportunities.

Before presenting the detailed literature survey, some key concepts are discussed in this section like UI and UX, UX Design, and principles of UX Design.

1.1 UI and UX

The terms user interface (UI) and user experience (UX) have become widely used and quickly accepted terms in the field of human–computer interaction. The "User Experience (UX)" concept is widely thought to have been introduced and promoted by American researcher Prof. Don Norman at the beginning of 1990s [1].

In business solutions, what differentiates the best product from the rest is "Customer/User Satisfaction". Most businesses understand the importance of this and invest efforts in it. To develop a product that people often love to use requires both good UI and UX. Following examples show why good UI and UX are necessary.

For example, you might have a good banking application that looks great and has an appealing user interface (UI). But if it takes time to load or makes you click through several screens for transactions like money transfer (UX). Then it doesn't matter how wonderful it feels, you are probably not going to use it.

1.2 User Experience (UX) Design

User experience (UX) design is a process that offer useful and pertinent experiences to mobile application users. It combines other elements like branding, design, usability, and functionality to design the entire process of acquiring and integrating the products. The American researcher, Don Norman coined the magical term "user experience" in the late 1990s. He described it as: "User experience encompasses all aspects of the end-user's interaction with the company, its services, and its products".

1.3 UI/UX Design Principles

There are many versions of UI/UX design principles, made by several people who have studied this field. These are available online but some are most popularly used and followed. In this paper, we are referring to Don Norman's Six Principles of Design [2]. These are described as follows:

Don Norman's Six Principles of Design

- **Visibility**: Users need to be aware of all available options and how to access them.
- **Feedback**: When a user performs an action like clicking a button, the system should give feedback indicating that the action has caused something.
- **Affordance**: Affordance is how a user can relate given object with some possible set of actions. For example buttons and links come with a particular appearance and indicate that the user has to click them to respond.
- **Mapping**: The association between controls and their effect on the world is known as mapping. In day to day life, controls to change something will seem quite similar to what they affect.
- **Constraints**: Some of the constraints are really obvious and physical such as the screen size on a mobile phone.
  - **Cultural constraints**: These are based on some existing cultural conventions of a computer society. These constraints can be universal or particular according to specific culture.
  - **Logical constraints**: Logical restrictions are those that employ reasoning and logical relationships to determine possible alternatives.
  - **Physical constraints**: These are about physical limitations. A simple example is, we cannot move a mouse cursor outside of the screen area.
  - **Semantic constraints**: These constraints depend on the semantics of the situation to make a proper decision. Red lights always signal stopping, whereas green ones signal moving.
- **Consistency**: In UX design, consistency ensures that every time the same action should cause the same reaction.

The outline of the paper is as: the purpose of this survey, i.e., motivation, is explained in Sect. 2. Section 3 covers relevant research in the field of automated GUI testing. Section 4 contains an in-depth analysis of the existing literature. Section 5 discusses open issues and challenges of automated GUI testing. Section 6 brings the paper to a conclusion.

2 Motivation

The demand for mobile apps has increased in recent years as a result of the growing popularity of mobile devices. As per Wikipedia information of April 2023, Android is the world's most widely used operating system. Android has 42% of web usage in the global market, whereas Windows has 28%, iOS with 17%, macOS with 7%, ChromeOS 1.3%, then (desktop) Linux at 1.2%. Android and iOS applications occupy a vast market share. For smartphones and other portable electronics, Android holds a 71% market share, while iOS holds a 28% share [3].

The Google Play store has the most apps available, with 3.55 million and the Apple App Store, has about 1.6 million iOS apps. These sectors are consumer markets, industrial sectors, public opinion, media trends, and more [4]. Many times users encounter bugs while using mobile GUI applications. This causes a failure in the respective functions of that app. To check if the application is performing properly is always essential for mitigating these types of issues.

The motivation for writing this review paper is that though there are millions of mobile applications, these may contain defects. It may naturally affect the quality of user experience and sometimes cause economic losses. Therefore this work focuses on studying existing tools and techniques to test the GUI of mobile applications and websites and presents the observations on it.

3 Literature Survey

Literature Survey of any scientific research topic may provide a good direction when we frame some kind of relevant set of questions. It has several purposes like these questions provide a clear focus; define the scope of the review, provide a logical structure for the review paper, etc. It ensures that the review paper addresses relevant and significant aspects of the topic and allows systematic and coherent presentation of the literature. For this work, we have framed some questions (referred as Research Questions (RQs)) related to of automated GUI testing to enhance user experience (UX).

3.1 Framing Research Questions (RQs) for Systematic Literature Review

This review paper concentrates on recognizing and analyzing existing research work that outlines various methodologies and approaches applied for automated GUI testing of applications.

Table 1 List of research questions (RQs)

S. No.	Question
Q.1	What has been the year-by-year advancement of automated GUI testing during the previous decade?
Q.2	What are the various techniques and tools used for automated GUI testing?
Q.3	What are different performance evaluation measures considered while performing automated GUI testing?
Q.4	What are the open issues and challenges identified from the literature survey?

Automated GUI testing is a software testing method that uses automated testing tools. It involves simulating user actions and interactions with the GUI to ensure that it functions correctly and meets user experience standards.

This study is carried out by collecting similar research publications published in reputed journals, conferences, and workshops from 2018 till date. This analysis is based on the formulation of **R**esearch **Q**uestions (**RQ**s). These are shown in Table 1, a detailed analysis, the results of these questions are collated.

3.2 Recent Progress in Automated GUI Testing Approaches

This section discusses recent progress in the area of automated GUI testing, which leads to the answer of **R**esearch **Q**uestion-1.

To test the user interface of mobile applications, Tao et al. proposed a reinforcement learning-based methodology. They suggested two innovative methods known as Automatic Testing based on PPO (ATPPO) and Automatic Testing based on A2C (ATAC). In order to evaluate the performance, the authors studied 20 mobile applications. They concluded that ATAC gives the best performance for 16 out of 20 applications in code coverage. Another approach ATPPO gets higher code coverage in 15 of 20 apps. Authors concluded that ATAC and ATPPO gives better code coverage than Monkey and ARES [5]. Lv et al. introduced an automated GUI testing technique for accelerating the testing cycle of Android applications. It is based on the understanding that information about event-activity transitions from the previous testing runs is useful to guide the subsequent testing runs so that the main app features can be quickly covered. They implemented their technique by using Fastbot2, an automated testing tool. They evaluated the performance of Fastbot2 using two industrial applications, Douyin and Toutiao and found that it performs better than existing tools like Ape, and Monkey [6].

To overcome limitations of GUI testing, Gao et al. used deep learning algorithms. For this, they built a data-driven machine learning solution. This reduces manual intervention to generate test scenarios with artificial intelligence solutions. For this purpose, authors used real-time object detection technology called You only look once (YOLO). Using YOLO, they detected GUI elements and proposed

automatic GUI test scenario generation. For result visualization purposes they used the NetworkX python package which showcases effectiveness and feasibility of the proposed approach [7]. An approach called Deep GUI was presented by Yazdani-BanafsheDaragh and Malek. Deep GUI is a deep learning model which takes input as screenshots of the applications. A good feature of Deep GUI is that it gets the inputs from an application under test without requiring its technical specifics. The authors used Monkey++ to create a Deep GUI prototype [8].

Gardey and Garrido created an A/B testing approach that compares the user experience (UX) of various designs for a web application in a real-world context with a high number of users. Once a collection of UX concerns has been identified, this method assists designers in UX development [9]. White et al. applied an image-based widget detection technique to improve random GUI testing. They discovered that detecting GUI widgets from the input screen and utilizing this knowledge to guide random testing resulted in considerably greater branch coverage in 18 out of 20 applications, with a 42.5% average improvement above traditional random testing [10].

Using formal specifications, Sen employed reinforcement learning approach for Android apps testing. They proposed a tool called as "Fully Automated Reinforcement LEArning-Driven Specification-Based Test Generator for Android (FARLEAD-Android)". As claimed by the authors, it is the first testing tool that combines RL and LTL. Their findings show that FARLEAD-Android outperforms the three other test generation methodologies known as Random, Monkey, and QBEa in terms of creating tests that adhere to specific specifications [11]. Vuong and Takada Trained a Deep Q-Network using QDroid, an automated testing tool. Based on the principle of reinforcement learning, the neural network calculates and uses a behavioral model of the application. Evaluation has revealed that QDroid outperforms state-of-the-art testing methods, gives better code coverage, and is successful in discovering faults [12].

For android app GUI testing, Adamo et al. employed reinforcement learning approach. They used a method to systematically choose events from the UI. They explored user interface of application under test without any pre-existing abstract model. They tested the tool on eight different Android apps and discovered that this provided greater code coverage than random test creation [13]. Hu et al. used a machine learning approach for UI testing. They demonstrated AppFlow, a technique for creating very robust, repeatable user interface testing. They used a machine learning approach. For the performance evaluation of AppFlow, authors used 60 popular Android apps from different domains. Screens and widgets are correctly recognized by AppFlow, which also creates incredibly reliable and reusable tests [14].

Rosenfeld et al. used machine learning to automate functional testing of Android applications. Authors tested functional behaviors of the different apps by using machine learning to categorize each of the application actions into a certain kind. They put the tool through its paces on 26 randomly chosen apps, proving its capacity to execute functional testing automatically by using application-specific actions which were converted generic testing scripts. They compared the proposed tool to

experienced mobile testers and discovered that the tool useful to automate human testers' jobs [15].

4 Observations from the Literature Survey

This work is carried out by keeping the aim in mind to undertake an organized literature review in the domain of automated GUI testing. This is achieved by conducting a detailed analysis by responding to the **R**esearch **Q**uestions (**RQ**s) listed in Table 1. This section covers the tools and techniques used in GUI testing and its performance evaluation measures.

4.1 Techniques and Tools Used for Automated GUI Testing

The description of tools and techniques used for GUI testing generates the response to the **R**esearch **Q**uestion-2.

Tao et al. used reinforcement learning-based approach for android GUI testing. The developed tool names are: ATAC—"Automatic Testing based on A2C" and ATPPO—"Automatic Testing based on PPO". The proposed methods reduce the negative impacts of the state explosion problem on the testing process to make sure that the end product meets the user experience guidelines and quality standards [5].

Vuong and Takada proposed QDroid GUI components testing tool. The proposed QDroid was successful in improving code coverage, when compared to other testing tools, as it took advantage of the estimated behavioral model generated by the neural network based on reinforcement learning [12].

FARLED-Android proposed by Koroglu et al. combines the principles of RL—"Reinforcement Learning" and LTL—"Linear-time Temporal Logic". FARLED-Android seeks to satisfy the LTL formula by learning to generate a test by means of RL. The LTL formula is generated by providing FARLED-Android with the formal specification of the GUI, which describes the intended behavior of the GUI [11].

4.2 Performance Evaluation Measures Considered While Performing Automated GUI Testing

In software testing, areas of the program code that have not been tested but may contain errors are found using coverage testing techniques. It also helps to ensure that the test suite is comprehensive and can detect faults in all parts of the code. In order to ensure the software's overall quality, additional testing methods, such as boundary value analysis (BVA), equivalence partitioning (EP), and error guessing

(EG), may be necessary. Achieving good coverage alone does not guarantee that the code is error-free. The terms branch coverage, method coverage, and code coverage are all used in software testing to assess how much the test suite has actually used the source code. It aims to ensure that the test suite exercises all parts of the source code, including statements, branches, conditions, and loops. The definition and formulas of these terms are relevant to Research Question-3 of Table 1 [16, 17].

Line coverage: It is defined as percentage of instructions or lines of program code executed throughout the testing process. It is calculated as follows:

$$\text{Line Coverage} = (\text{Number of lines executed during testing} / \text{Total number of lines in the code}) \times 100\%$$

Instruction coverage: It is defined as the percentage of executable instructions that have been executed during the testing process. It is calculated as follows:

$$\text{Instruction Coverage} = (\text{Number of instructions executed during testing} / \text{Total number of executable instructions in the code}) \times 100\%$$

Branch coverage: It is defined as the percentage of decision points (e.g., if-else statements, switch statements, etc.) in the code that have been executed during testing. A decision point is considered covered if all possible outcomes have been tested.

$$\text{Branch Coverage} = (\text{Number of covered decision points} / \text{Total number of decision points}) \times 100\%$$

Method coverage: It is percentage of methods in the code executed during testing. A method is considered covered if it has been called at least once during testing.

$$\text{Method Coverage} = (\text{Number of covered methods} / \text{Total number of methods}) \times 100\%$$

Code coverage: It is described as the proportion of lines of code that were run during the testing process. A line of code is considered covered if it has been executed at least once during testing.

$$\text{Code Coverage} = (\text{Number of covered lines} / \text{Total number of lines}) \times 100\%$$

5 Open Issues and Challenges of Automated GUI Testing

This section, which highlights various open issues and challenges associated with automated GUI testing, provides the response to Research Question-4. Based on the above literature survey, following are the open issues and challenges:

- As discussed in Sect. 4.1, there are various tools available for automated GUI Testing tasks like QDroid, ATPPO, and ATAC. As per the literature survey, these tools are used to test simple GUI elements. In the future, more tools can be designed to test complex GUI elements like progress bars, graphical widgets, date picker, etc.
- As studied in literature survey, many authors have tested their tool with hundreds of mobile apps in order to assess the performance. They explored the testing of Android applications. Fewer studies considered iOS applications. As discussed in Sect. 2, millions of Android and iOS mobile applications exist. To achieve better performance more applications can be explored in future.
- In many research papers authors performed only coverage testing of GUI, like line coverage, instruction coverage, branch coverage, etc. But it is important to note that while high line and instruction coverage can indicate good test coverage, they do not guarantee that all potential errors have been found. Therefore, in order to ensure thorough testing, additional testing strategies like boundary value analysis (BVA), equivalence partitioning (EP), and error guessing (EG) are required.

6 Conclusions and Future Work

This research paper has explored the recent progress in automated GUI testing for enhancing user experience (UX). The key motivation to exercise this literature review is the potential growth of the research work carried out over the past couple of years. The research has highlighted the work in the area of automated GUI testing, along with various techniques and tools. Traditional manual testing methods are time-consuming, error-prone, and often unable to detect subtle UI issues. Automated GUI testing, on the other hand, offers several advantages such as efficiency, scalability, and improved test coverage. By automating repetitive and time-consuming testing tasks, developers and Quality Assurance (QA) professionals can focus more on design and innovation, ultimately resulting in better user-centric software products. However, it is important to acknowledge that automated GUI testing has its own challenges and limitations. To overcome the limitations of current tools and techniques, in the future experimental work will be carried out. This may lead to improve good user experience. Future work may focus on ensuring wide test case coverage, handling dynamic UI elements, and managing test data. Additionally, researchers of this domain may focus on developing more intelligent and adaptive testing approaches, incorporating user feedback and preferences into the testing process, and exploring new ways to enhance automated GUI testing process.

References

1. Yang B, Wei L, Pu Z (2020) Measuring and improving user experience through artificial intelligence-aided design. Front Psychol 11:595374
2. Hamidli N. Introduction to UI/UX design: key concepts and principles
3. Wikipedia webpage. https://en.wikipedia.org/wiki/Usage_share_of_operating_systems. Accessed 20 May 2023
4. Statista webpage. https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/#statisticContainer. Accessed 20 May 2023
5. Tao C, Gao Y, Guo H, Gao J (2023) A reinforcement learning-based approach to testing GUI of mobile applications
6. Lv Z, Peng C, Zhang Z, Su T, Liu K, Yang P (2022) Fastbot2: reusable automated model-based GUI testing for android enhanced by reinforcement learning. In: 37th IEEE/ACM international conference on automated software engineering, Oct 2022, pp 1–5
7. Gao J, Li S, Tao C, He Y, Anumalasetty AP, Joseph EW, Nayani H (2022) An approach to GUI test scenario generation using machine learning. In: 2022 IEEE international conference on artificial intelligence testing (AITest), Aug 2022. IEEE, pp 79–86
8. YazdaniBanafsheDaragh F, Malek S (2021) Deep GUI: black-box GUI input generation with deep learning. In: 2021 36th IEEE/ACM international conference on automated software engineering (ASE), Nov 2021. IEEE, pp 905–916
9. Gardey JC, Garrido A (2020) User experience evaluation through automatic A/B testing. In: Proceedings of the 25th international conference on intelligent user interfaces companion, Mar 2020, pp 25–26
10. White TD, Fraser G, Brown GJ (2019) Improving random GUI testing with image based widget detection. In: Proceedings of the 28th ACM SIGSOFT international symposium on software testing and analysis, July 2019, pp 307–317
11. Koroglu Y, Sen A (2019) Reinforcement learning-driven test generation for android GUI applications using formal specifications. arXiv preprint arXiv:1911.05403
12. Vuong TAT, Takada S (2019) Semantic analysis for deep Q-network in android GUI testing. In: SEKE, Jan 2019, pp 123–170
13. Adamo D, Khan MK, Koppula S, Bryce R (2018) Reinforcement learning for android GUI testing. In: Proceedings of the 9th ACM SIGSOFT international workshop on automating TEST case design, selection, and evaluation, Nov 2018, pp 2–8
14. Hu G, Zhu L, Yang J (2018) AppFlow: using machine learning to synthesize robust, reusable UI tests. In: Proceedings of the 2018 26th ACM joint meeting on European software engineering conference and symposium on the foundations of software engineering, Oct 2018, pp 269–282
15. Rosenfeld A, Kardashov O, Zang O (2018) Automation of android applications functional testing using machine learning activities classification. In: Proceedings of the 5th international conference on mobile software engineering and systems, May 2018, pp 122–132
16. Wikipedia webpage. https://en.wikipedia.org/wiki/Code_coverage#Basic_coverage_criteria. Accessed 20 May 2023
17. Dustin E (2003) Effective software testing: 50 specific ways to improve your testing. Pearson Education India

Author Index

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 765, https://doi.org/10.1007/978-981-99-5652-4