Lecture Notes in Networks and Systems 754

Milan Tuba
Shyam Akashe
Amit Joshi *Editors*

ICT Infrastructure and Computing

Proceedings of ICT4SD 2023, Volume 3

Springer

Lecture Notes in Networks and Systems

Volume 754

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 Infrastructure and Computing

Proceedings of ICT4SD 2023, Volume 3

Springer

Editors
Milan Tuba
Singidunum University
Belgrade, Serbia

Shyam Akashe
ITM University
Gwalior, Madhya Pradesh, India

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, Gujarat, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-99-4931-1 ISBN 978-981-99-4932-8 (eBook)
https://doi.org/10.1007/978-981-99-4932-8

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

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: GR 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, Ghana, United Arab Emirates, 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 organized Physical at Goa and Virtual on Zoom including 1 inaugural and 1 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 socio-economic 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-day 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 was born in Lucani in 1952 and completed primary and secondary education in Osijek. He received B.S. in Mathematics, M.S. in Mathematics, M.S. in Computer Science, M.Ph. in Computer Science, and Ph.D. in Computer Science from University of Belgrade and New York University. From 1983 to 1994, he was in the USA first as a graduate student as well as a teaching and research assistant at Vanderbilt University in Nashville and Courant Institute of Mathematical Sciences, New York University, and later as an assistant professor of Electrical Engineering at Cooper Union Graduate School of Engineering, New York. During that time, he was the founder and the director of Microprocessor Lab and VLSI Lab, the leader of scientific projects, and a supervisor of many theses. His research interests include mathematical, queuing theory, and heuristic optimizations applied to computer networks, image processing, and combinatorial problems.

Shyam Akashe is a 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. He has participated in numerous national and international conferences and presented over 100 papers.

Amit Joshi is currently the director of Global Knowledge Research Foundation and also an entrepreneur researcher who has completed his master's and research in the areas of cloud computing and cryptography in medical imaging. He has an experience of around 10 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 an 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 organized more than 50 national and international conferences and programs in association with ACM, Springer, IEEE, to name a few across different countries including India, UK, Europe, USA, Canada, Thailand, Egypt, and many more.

Contributors

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

Shlok Agarwal Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

N. Ajay S J C Institute of Technology/Computer Science and Technology, Chickballapur, India

J. K. Ajin KCG College of Technology, Karapakkam, Chennai, India

Jangam Akhila School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, India

M. Anu Mohan New Horizon College of Engineering, Bangalore, India

Pisati Anusha School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, India

Richard Apau Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana

I. Aravindhan KCG College of Technology, Karapakkam, Chennai, India

Michael Asante Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana

Siddhi Bafna Medi-Caps University, Indore, India

T. Jagan Balaji Department of EEE, Rajalakshmi College of Engineering, Chennai, Tamil Nadu, India

Janki Barot Silver Oak University, Ahmedabad, India

Devang Bathwar Department of Electronics and Communication, Institute of Technology, Nirma University, Ahmedabad, India

Sneha Belapurkar Medi-Caps University, Indore, India

Shubham Bhange Symbiosis Institute of Geoinformatics, Symbiosis International (Deemed University), Pune, India

M. S. Bharath Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Jitendra Bhatia Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, India

Divyesh Bhatnagar MBM University, Jodhpur, India

Yash Bhise VES Institute of Technology, Mumbai, India

Anna Bon Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

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

Ruchi Chakkarwar Shah and Anchor Kutchhi Engineering College, Mumbai, India

Dipanwita Chakrabarty Haldia Institute of Technology, Haldia, West Bengal, India

Ruby Chanda Symbiosis Institute of Management Studies, Symbiosis International (Deemed) University, Pune, India

Saumya Chaturvedi School of Engineering and Technology, Sharda University, Greater Noida, India

Archana Chaudhari Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

Vaibhav Chaudhari JSPM's Rajarshi Shahu College of Engineering, Pune, India

Anamika Chauhan Department of Information Technology, Delhi Technological University, New Delhi, India

Damayanti Chavan Department of Computer Engineering, International Institute of Information Technology (IIT), Pune, India

Trushita Chaware School of ECE, Dr. Vishwanath Karad MIT-WPU, Pune, India

Shakti Raj Chopra School of Electronics and Electrical Engineering, Lovely Professional University Phagwara, Phagwara, India

Sandeep Chourasia School of Management Studies, Sangam University, Bhilwara, Rajasthan, India

Dhruvendra Kumar Chourishi Department Of Computer Science and Engineering, Government Women's Polytechnic College, Bhopal, Madhya Pradesh, India

Pranay Dadheech Department of CSE, Geetanjali Institute of Technical Studies, Udaipur, India

Rohit Dadhel JSPM's Rajarshi Shahu College of Engineering, Pune, India

Ramraj Dangi Department of CSE, SCSE, Manipal University Jaipur, Jaipur, Rajasthan, India

Rahul Dhaigude Symbiosis Institute of Management Studies, Symbiosis International (Deemed) University, Pune, India

Soumya Kanti Dhara Haldia Institute of Technology, Haldia, West Bengal, India

Saurabh Dhavale JSPM's Rajarshi Shahu College of Engineering, Pune, India

Saksham Dube Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Tanya Dubey VES Institute of Technology, Mumbai, India

Shailesh Gahane Faculty of Sciences and Technology, School of Allied Sciences, Datta Meghe Institute of Higher Education and Research, Wardha, India

Lokesh P. Gaikwad Vishwakarma Institute of Information Technology, SPPU, Pune, India

Arunangshu Giri Haldia Institute of Technology, Haldia, West Bengal, India

Gaurav Gogisetty Department of CSE, SCSE, Manipal University Jaipur, Jaipur, Rajasthan, India

Vinaya Gohokar School of ECE, Dr. Vishwanath Karad MIT-WPU, Pune, India

Janhavi Gosavi Department of Computer Engineering, Shah and Anchor Kutchhi College, Mumbai, India

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

Sanjeev Gour Department of Computer Science, Career College, Bhopal, Madhya Pradesh, India

Nithin Vajuvalli Narayana Gowda Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

Anil Gupta MBM University, Jodhpur, India

Gunjan Gupta A2G Enterprises Pvt. Ltd., Noida, UP, India

Jalal J. Hamad Ameen Electrical Engineering, College of Engineering, University of Salahaddin—Hawler, Erbil, Iraq

Ari Happonen LUT University, Lappeenranta, Finland

James Ben Hayfron-Acquah Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana

Ritesh Hiware Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Yashwant S. Ingle Vishwakarma Institute of Information Technology, Pune, India; SKNCOE, Pune, India

K. Iniyan KCG College of Technology, Karapakkam, Chennai, India

Dhanraj Jadhav JSPM's Rajarshi Shahu College of Engineering, Pune, India

Vinayak P. Jakite Vishwakarma Institute of Information Technology, SPPU, Pune, India

Ashwini Jarali Department of Computer Engineering, International Institute of Information Technology (IIT), Pune, India

J. Jayapriya Christ University, Bangalore, Karnataka, India

Aditya U. Joshi Vishwakarma Institute of Information Technology, SPPU, Pune, India

Manasi Joshi Department of Computer Engineering, International Institute of Information Technology (IIT), Pune, India

Sudarshan Joshi School of ECE, KLE Technological University, Hubballi, India

Neerajaa Kadam Department of Computer Engineering, Shah and Anchor Kutchhi College, Mumbai, India

Hridesh Kamal VES Institute of Technology, Mumbai, India

Pratika Kamath Shah and Anchor Kutchhi Engineering College, Mumbai, India

Meghana Kanneganti Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

M. Karuppasamy Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India

Kumar Keshamoni Department of ECE, Koneru Lakshmaiah Education Foundation, Hyderabad, India

Darshi Khalasi Department of Electronics and Communication, Institute of Technology, Nirma University, Ahmedabad, India

Mohammed Sinan Khan Department of Artificial Intelligence and Machine Learning, BMS Institute of Technology and Management, Bengaluru, India

S. Kirivarshan KCG College of Technology, Karapakkam, Chennai, India

L. Koteswara Rao Department of ECE, Koneru Lakshmaiah Education Foundation, Hyderabad, India

Krutthika Hirebasur Krishnappa Department of Computer Science, Southern University and A&M College, Baton Rouge, LA, USA

Apurva Kulkarni Department of Computer Engineering, International Institute of Information Technology (IIT), Pune, India

Isha Kulkarni Shah and Anchor Kutchhi Engineering College, Mumbai, India

Kunal Kulkarni School of ECE, Dr. Vishwanath Karad MIT-WPU, Pune, India

Pradnya Kulkarni School of ECE, Dr. Vishwanath Karad MIT-WPU, Pune, India

Rajesh Kumar Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, Rajasthan, India

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

Vimal Kumar School of Computer Science, University College Dublin, Dublin, Ireland

Malaram Kumhar Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, India

R. Madhura Visvesvaraya Technological University Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

Parikshit Mahalle Vishwakarma Institute of Information Technology, Pune, India

A. Bharathi Malakreddy Department of Artificial Intelligence and Machine Learning, BMS Institute of Technology and Management, Bengaluru, India

K. Malathi KCG College of Technology, Karapakkam, Chennai, India

R. Manasa Visvesvaraya Technological University Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

T. N. Manjunath BMSIT&M/Computer Science and Technology, Bengaluru, India

Jainam S. Mardia Department of EEE, Rajalakshmi College of Engineering, Chennai, Tamil Nadu, India

Aayush Mehta Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Shivangi Mehta Devang Patel Institute of Advance Technology and Research, Changa, Anand, Gujarat, India

Trapti Mishra Medi-Caps University, Indore, India

Sakshi Modi Medi-Caps University, Indore, India

Priyanka D. More Vishwakarma Institute of Information Technology, SPPU, Pune, India

Kapil Mundada Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, India

Sachin Naik Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Pune, India

Dnyaeshwari Nakhate Faculty of Sciences and Technology, School of Allied Sciences, Datta Meghe Institute of Higher Education and Research, Wardha, India

Pranshu C. B. S. Negi Indian Institute of Technology (Banaras Hindu University), Varanasi, UP, India

Abhisheik Nimpure Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Shrey Ojha Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Vanishree Pabalkar Symbiosis International (Deemed University), Pune, India

Srinivas Padmanabhuni AiEnsured, Bengaluru, India

Antria Panayiotou Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

S. S. Pandey Institute of Medical Sciences, Banaras Hindu University, Varanasi, UP, India

And Prashant Panse Medi-Caps University, Indore, India

Aryaman Parihar Department of CSE, SCSE, Manipal University Jaipur, Jaipur, Rajasthan, India

Pritee Parwekar SRM Institute of Science and Technology, NCR Campus, Ghaziabad, UP, India;
School of Life Sciences, JNU, Delhi, India

Mayank Patel Department of CSE, Geetanjali Institute of Technical Studies, Udaipur, India

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

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

Aniket Y. Patil Vishwakarma Institute of Information Technology, Pune, India

Sumedh Patti Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, India

Reshma Pohankar Faculty of Sciences and Technology, School of Allied Sciences, Datta Meghe Institute of Higher Education and Research, Wardha, India

Nithi Poojari Shah and Anchor Kutchhi Engineering College, Mumbai, India

K. Poorani Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India

Mekala Pranav School of Electronics and Electrical Engineering, Lovely Professional University Phagwara, Phagwara, India

Sagar Pujari VES Institute of Technology, Mumbai, India

Leena Ragha Ramrao Adik Institute of Technology, D. Y. Patil Deemed to Be University, Navi Mumbai, Maharashtra, India

Ramya Rajamanickam National Informatics Centre, New Delhi, India

Tejas Rajguru Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, India

Anil Rajput Higher Education Department, OSD, Government of Madhya Pradesh, Bhopal, Madhya Pradesh, India

G. Ramesh New Horizon College of Engineering, Bangalore, India

Abhishek Rathi Department of Information Technology, Delhi Technological University, New Delhi, India

Tarun Kumar Rathi Department of Information Technology, Delhi Technological University, New Delhi, India

Harish Rathod Gujarat Technological University, Ahmedabad, Gujarat, India

Parth Rathod Devang Patel Institute of Advance Technology and Research, Changa, Anand, Gujarat, India

Ravi Rathod Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

Narendra Singh Rathore Geetanjali Institute of Technical Studies, Udaipur, India

Riya Raut Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

Chandan Singh Rawat VES Institute of Technology, Mumbai, India

Faiz raza Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

H. Sai Vardhan Reddy KLE Technological University, Hubballi, India

T. Charandeep Reddy Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Francis Saa-Dittoh University for Development Studies, Tamale, Ghana

Annam Rohan Venkata Sai School of Electronics and Electrical Engineering, Lovely Professional University Phagwara, Phagwara, India

Annam Rohith Naga Sai School of Electronics and Electrical Engineering, Lovely Professional University Phagwara, Phagwara, India

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

Riya Sakhiya Devang Patel Institute of Advance Technology and Research, Changa, Anand, Gujarat, India

Adrinil Santra Haldia Institute of Technology, Haldia, West Bengal, India

Puskraj Savji Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, India

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

Riya Shah Devang Patel Institute of Advance Technology and Research, Changa, Anand, Gujarat, India

Nuzhat Faiz Shaikh SKNCOE, Pune, India

Sayali Shambharkar Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, India

G. Shanmmukha Srinivasa Reddy New Horizon College of Engineering, Bangalore, India

Ajay Kumar Sharma Department of CSE, Geetanjali Institute of Technical Studies, Udaipur, India

Neeraj Sharma Indian Institute of Technology (Banaras Hindu University), Varanasi, UP, India

Shiru Sharma Indian Institute of Technology (Banaras Hindu University), Varanasi, UP, India

Ameya Shetty Department of Computer Engineering, Shah and Anchor Kutchhi College, Mumbai, India

M. R. Shrihari S J C Institute of Technology/Computer Science and Technology, Chickballapur, India

D. A. Shubhashree S J C Institute of Technology/Computer Science and Technology, Chickballapur, India

A. R. Shylaja S J C Institute of Technology/Computer Science and Technology, Chickballapur, India

Saroja V. Siddamal KLE Technological University, Hubballi, India

C. Sindhu Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Harmandeep Singh Department of Information Technology, Delhi Technological University, New Delhi, India

Ritu Singh Sanskriti University, Mathura, India

Vijander Singh Department of ICT and Science, Faculty of Information Technology and Electrical Engineering, NTNU–Norwegian University of Science and Technology, Ålesund, Norway

Vivek Kumar Singh Department of Data Science and Engineering, Indian Institute of Science Education and Research, Bhopal, India;
Center of Environmental Sensing and Modeling, Singapore-MIT Alliance for Research and Technology, Singapore, Singapore

Deepak Sinwar Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, Rajasthan, India

Vipin Kumar Solanki Sanskriti University, Mathura, India

G. Jai Sreedhar Department of EEE, Rajalakshmi College of Engineering, Chennai, Tamil Nadu, India

D. Subba Rao Department of ECE, Siddhartha Institute of Engineering and Technology, Hyderabad, India

L. V. Subbareddy New Horizon College of Engineering, Bangalore, India

S. Sudharsan Department of ECE, Rajalakshmi College of Engineering, Chennai, Tamil Nadu, India

T. V. Sumithra Ramrao Adik Institute of Technology, D. Y. Patil Deemed to Be University, Navi Mumbai, Maharashtra, India

M. Sundharam New Horizon College of Engineering, Bangalore, India

Yalla Ebenezer Sunith School of Electronics and Electrical Engineering, Lovely Professional University Phagwara, Phagwara, India

Boyapally Swathi School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, India

T. Kavya taneesha Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Puneshkumar U. Tembhare Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Divyesh Thakur Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

Apoorva Thorat Department of Computer Engineering, International Institute of Information Technology (IIT), Pune, India

Vinodray Thumar Department of Computer Engineering, Vishwakarma Government Engineering College, GTU, Ahmedabad, India

Prerna Tiwari School of Management Studies, Sangam University, Bhilwara, Rajasthan, India

Frimpong Twum Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana

Komal Ugemuge Faculty of Sciences and Technology, School of Allied Sciences, Datta Meghe Institute of Higher Education and Research, Wardha, India

Sanket Ujgare JSPM's Rajarshi Shahu College of Engineering, Pune, India

Usman Ahmad Usmani Universiti Teknologi Petronas, Perak, Malaysia

Roopchand Reddy Vanga Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Prashant Vats Department of CSE, SCSE, Manipal University Jaipur, Jaipur, Rajasthan, India

Sadanand Venkataraman Department of Artificial Intelligence and Machine Learning, BMS Institute of Technology and Management, Bengaluru, India

Anant Verekar Department of Computer Engineering, Shah and Anchor Kutchhi College, Mumbai, India

Rajkumar Verma Imagineer Private Limited, Bangalore, India

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

M. Vinay Christ University, Bangalore, Karnataka, India

Pinki Vishwakarma Department of Computer Engineering, Shah and Anchor Kutchhi College, Mumbai, India

Anom S. Warbhuvan Vishwakarma Institute of Information Technology, SPPU, Pune, India

Junzo Watada Waseda University, Tokyo, Japan

K. P. Yashaswini Visvesvaraya Technological University Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

Traffic Classification of Software-Defined Networks Using Machine Learning

Sudarshan Joshi and Suneeta V. Budihal

Abstract The classification of network traffic has been a primary topic in the growing field of networking engineering. It is a foremost task for the Internet Service Providers (ISPs) to have knowledge about the types of network applications that are flowing into a network. The network operators or the ISPs need to manage the overall performance of a network. To accomplish these requirements, a learning-based traffic classification is proposed in this paper. Two machine learning (ML) algorithms Logistic Regression (supervised) and K-means clustering (unsupervised) were used to classify DNS, Telnet, Ping and voice traffic flows with Distributed Internet Traffic Generator (D-ITG). The accuracy is 78.89% indicating that the Logistic Regression models perform better when compared to K-means clustering.

Keywords Network traffic · Classification · Logistic Regression · K-means clustering

1 Introduction

Network classification is getting popular over these years since it provides the ability to automatically detect the application which has generated a given stream of packets from direct view and individual view or packet distribution which flows through the network. This capability can be useful in many tasks which have great interest to carriers, ISPs and network administrators. For any traffic management activities from pricing, treatment and security functions, traffic segregation is one important and basic need. There are different ways and are also prioritised for the network traffic classification which are port-based approach, payload-based approach and statistics-based. The payload-based approach focuses on DPI review process which uses application marks as part of the payload to separate conferences. Software-

http://www.kletech.ac.in.

S. Joshi · S. V. Budihal (✉)
School of ECE, KLE Technological University, Hubballi 580031, India
e-mail: suneeta_vb@kletech.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_1

Defined Networking (SDN) is an emerging, dynamic and cost saving structure which makes it ideal for the scope of bandwidth. OpenFlow has defined network protocol which help to manage and direct traffic between routers and switch.

Traditional method includes identifying based on port number and in this case payload inspection do not work because of dynamic and encrypted nature of current traffic. This project tries to attempt to use monitored and targeted ML algorithms to differentiate different traffic flows based on required bandwidth, QoS and their system based on different flow rate details. As network traffic classification using ML is growing rapidly. Its main strategies is in cyber security. ML classification and flow integration can also help identifies network vulnerability and potential issues to network infrastructure. Lastly, identifying applications or web-based agreements is important in predicting future trends and ensuring that the network can meet demand. Therefore, the segregation of traffic for SDN-based networks using a ML process.

2 Literature Survey

The various types of methods that can be used for traffic classification [1] are port-based, payload-based and statistical-based. During applications creation of port-based classification cannot be employed, dynamic ports are used. Identification of a pattern for the application is done in case of payload-based (DPI). For the purpose of recognising patterns, regular expressions are used. Pattern-based categorization is severely constrained in its ability to update the generalised pattern for many applications due to the applications' rapid expansion. The DPI technique is unable to categorise encrypted communications. The description about some of the papers referred for this project is mentioned below. Authors in [1] focused on different types of network classification approaches. Scrambled organise activity has recently grown to be a tremendous difficulty for organise management due to the widely spread use of encryption techniques in system applications. Due to these realities, it is critical to develop workable solutions for controlling system traffic. Much attention has been given to the development of novel methods for traffic classification because current methods are largely ineffective and easy to get around.

In a paper that was proposed [2], concepts like application-aware networking and network-aware application programming attempt to get around these restrictions. An improved interface between networks and applications is now possible because to the development of Software-Defined Networking (SDN). SDN is a modern, programmable networking architecture that represents how networks will develop in future. Several other network operations, such as security monitoring, accounting, QoS, and providing operators with meaningful projections for long-term provisioning, all depend on accurate traffic classification across SDN. Using an inter-disciplinary fusion of IP networking and data mining approaches, [3] released a paper that examines new research into the application of ML techniques to IP traffic classification. According to the ML methodologies they use and their main contributions to the body of literature, these works are categorised and reviewed.

Few important prerequisites for the use of ML-based traffic classifiers in operational IP networks, qualitatively assess these prerequisites. The fieldâŁ™s open problems and difficulties are all explored.

A article by Perera et al. [4] focused on network data analysis and ML implementation of a network traffic classification solution and integration of the model in Software-Defined Networking platform. Network automation is made possible by ML applications thanks to the introduction of Software-Defined Networking (SDN), a programmable and scalable networking solution. An SDN solution can address problems with conventional approaches to categorise network traffic and assign resources. The SDN controller will collect network data that data analytics techniques may evaluate and use ML models to tailor network management. A paper by Raikar et al. [5] introduced supervised ML as a method for classifying network traffic. Support Vector Machine (SVM), Nearest Centroid (NC) and Naive Bayes (NB) are three supervised learning models that are used to categorise the data traffic based on the applications in a Software-Defined Network architecture. In order to create features from the network traffic traces, which are then passed to the classifier for prediction. The closest centroid is 91.02%, NB is 96.79% and the accuracy for the SVM is 92.3%.

Using the statistical features of network traffic flow, Al Khater et al. [6] discussed how researchers utilise ML algorithms in a number of categorization techniques. It also describes the next step in our research, which is to examine several categorization methods (supervised, semi-supervised, and unsupervised) that make use of ML algorithms to handle actual network data. Network classification is a key component of network intrusion detection systems and the initial stage in network traffic analysis (IDS). Although classification techniques have advanced and their accuracy has increased, this area of study is still mostly unexplored due to the growing popularity of encryption and the persistence of application developers in finding novel ways to escape being filtered and discovered.

A paper [7] proposed introduces the various levels of network traffic analysis and the pertinent knowledge in the ML domain, analyses the issues with port-based and payload-based methods in traffic classification and makes recommendations that is available online. In order to assess the effectiveness and performance of the machine learning-based strategy, we experiment with unsupervised K-means. To identify the best feature set, we use feature selection, and to increase accuracy, we use log transformation. According to the experimental results on various datasets, the method can achieve an overall accuracy of up to 80%, and after a log transformation, the accuracy is increased to 90% or higher. Authors in [8] published an article in which they reviewed the methods currently used for categorising and analysing encrypted communications. We begin by outlining the most widely used encryption technologies on the Internet. It demonstrates how the start of an encrypted connection and the protocol structure reveal a lot of information for the classification and analysis of encrypted traffic. Payload and feature-based classification methods for encrypted traffic and categorise them using an established taxonomy is surveyed. The advantage of some of described classification methods is the ability to recognise the encrypted application protocol in addition to the encryption protocol. Finally, a comprehensive comparison of the surveyed feature-based classification methods and present their weaknesses and strengths is made.

3 Proposed Methodology

The SDN architecture is used for classifying the network traffic generated by the hosts. The OpenFlow protocol is used as the communication interface between the hosts and the controller. An overlay network is deployed so that the traffic generated by the hosts flows directly through the Switch VM. The hosts in the network used Open V Switch (OVS) with two interfaces. The first interface was internal network of the virtual box and the second connected towards the Switch OVS using VXLAN tunnelling. The Switch OVS had VXLAN interfaces towards the hosts and connected to the controller VM directly, i.e. using underlay IP. The Ryu controller was deployed on the controller VM as shown in Fig. 1.

After successful creation of topology the network is tested by sending data packets between any two host VMs. If connection established is successful there won't be any packet loss otherwise there will be packet loss. Logistic Regression and K-means clustering are the two ML models used for classifying the traffic (Fig. 2).

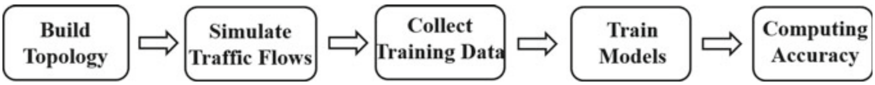


Fig. 1 Functional block diagram of the proposed methodology

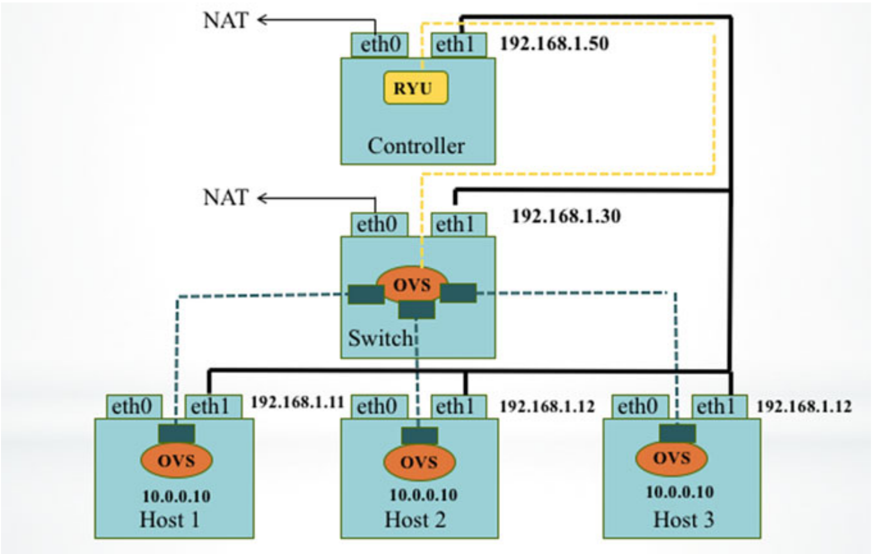


Fig. 2 Topology of the proposed methodology

- Topology: A custom topology is created consisting of three hosts, a switch and controller. Topology of any number of hosts and switches can be created as per user requirement.
- Controller: The controller is the central unit of the system. The SDN controller is used to direct the traffic flow between the hosts. The Ryu controller is used for forwarding the traffic within the network.
- Logistic Regression: The first ML algorithm is Logistic Regression model which is a supervise learning model. It is used to predict categorical target variables. Mostly, the outcome is a binary value; but in the case of multiple targets, it picks the target with the highest probability of occurring.
- K-means clustering The other model is K-means clustering which is an unsupervised ML model. It groups data points in multidimensional space together to form clusters and label each data point in the cluster the same. The initial desired number of clusters are chosen (in this case they are four types of traffic). The four clusters are used as an input parameter to the model.

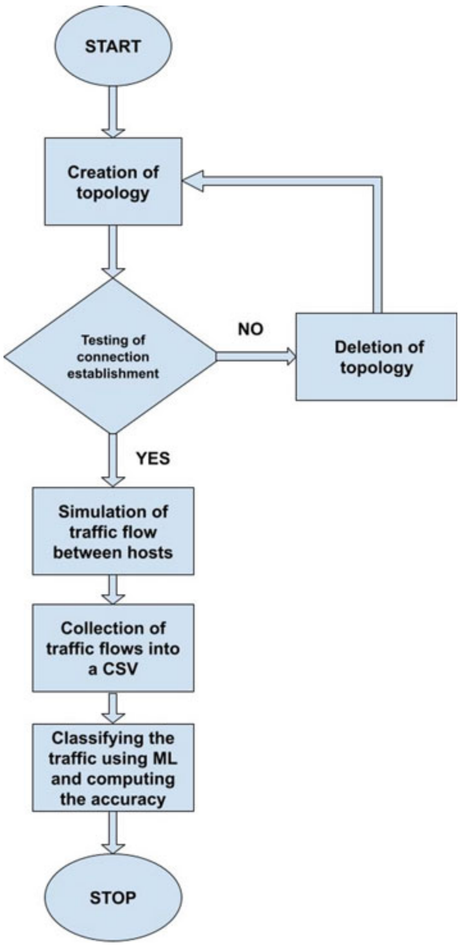


Fig. 3 Flow chart describing the methodology of the proposed model

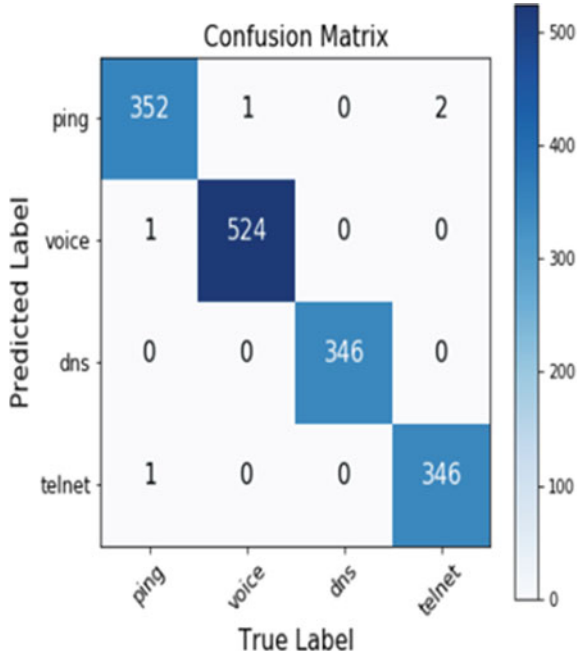


Fig. 4 Confusion matrix of Logistic Regression

The detailed flow chart for the proposed model is shown in Fig. 3. Firstly, a topology is created and testing of connection establishment is carried out, if there is a failure in connection establishment the topology is deleted and new topology is built. If the connection are right, then traffic flows between the hosts are simulated. These traffic flows are collected into a .csv file using Python scripts and are classified using ML models, and lastly, the accuracy is computed.

4 Result Analysis

Figures 4 and 5 show the difference between the confusion matrix (CM) of the two proposed models. The CM can help to pinpoint where the model is failing to accurately decide the target.

Figures 4 and 5 show the CM of Logistic Regression and shows the CM of K-means clustering. Comparing the CM of two models, it is clear that when the Logistic Regression is used, it predicts values more accurately compared to the K-means. The accuracy obtained for the Logistic Regression is 78.89 %, and the K-means clustering is 38 %. Figures 6 and 7 show the predicted results of both the models.

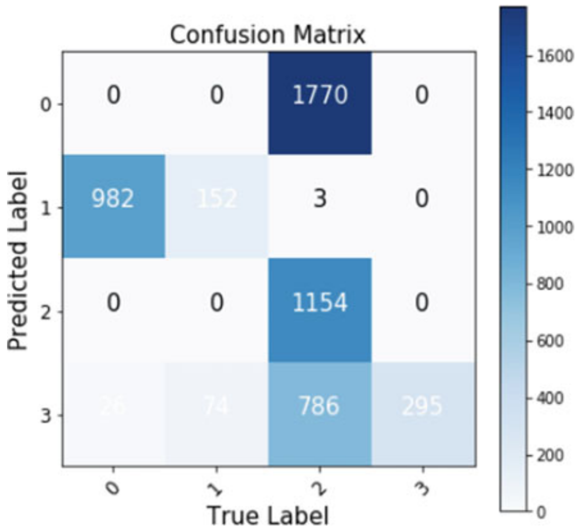


Fig. 5 Confusion matrix of K-means clustering

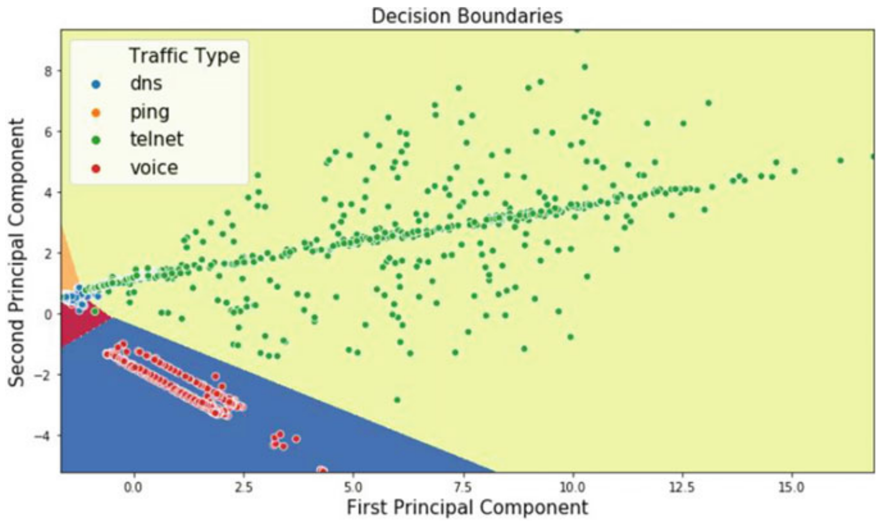


Fig. 6 Decision boundaries for the predicted results of the Logistic Regression

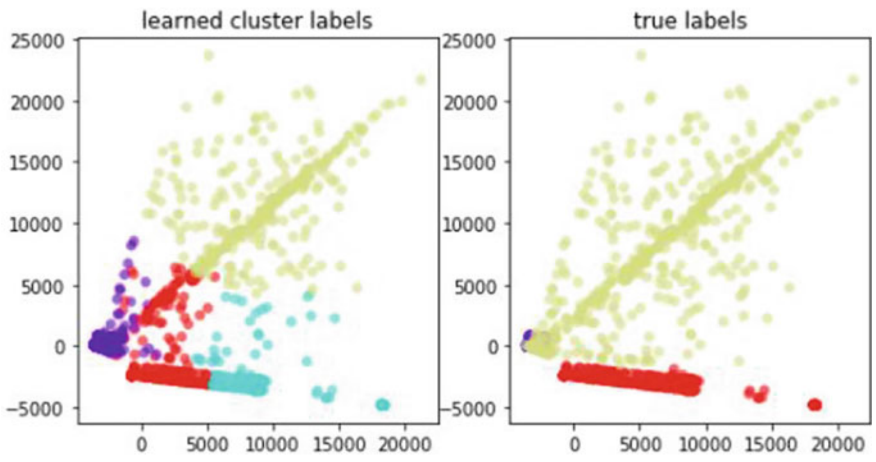


Fig. 7 Decision boundaries for the predicted results of the K-means clustering

5 Conclusion

The proposed method is used for classifying the traffic flows with the help of Software-Defined Networks and ML techniques, to compare ML algorithm. The captured traffic is classified using two ML algorithms, the experimental results show that Logistic Regression model is better in predicting the traffic with an accuracy of 78.89% compared K-means which has an accuracy of 38%. Higher accuracy rate was achieved using Logistic Regression which performed better compared to K-means clustering.

References

1. Gaofeng et al (2018) Mobile app identification for encrypted network flows by traffic correlation. Int J Distributed Sensor Netw 14(12)
2. Parsaei MR et al (2017) Network traffic classification using machine learning techniques over software defined networks. Int J Adv Comput Sci Appl (IJACSA) 8(7)
3. Nguyen TA, Grenville A (2008) A survey of techniques for internet traffic classification using machine learning. IEEE Commun Surv Tutorials 10(4):56–76
4. Amaral P et al (2016) Machine learning in software defined networks: data collection and traffic classification. In: IEEE 24th international conference on NP, pp 1–5
5. Al Khater N et al (2015) Network traffic classification techniques and challenges. https://doi.org/10.1109/ICDIM.2015.7381869
6. Hsu B-M (2020) Comparison of supervised classification models on textual data. Mathematics 8:851. https://doi.org/10.3390/math8050851
7. Liu Y, Li W, Li Y (2007) Network traffic classification using K-means clustering, pp 360–365. https://doi.org/10.1109/IMSCCS.2007.52
8. Velan P, Cermak M, Celeda P, Drasar M (2015) A survey of methods for encrypted traffic classification and analysis. Int J Netw Manage 25. https://doi.org/10.1002/nem.1901

9. Budihal SV et al (2016) Framework for intercell interference avoidance in MIMO OFDMA system. In: 2016 international conference on wireless communications, signal processing and networking
10. Hegadi R et al (2019) Performance evaluation of in-network caching: a core functionality of information centric networking. In: 2019 international conference on data science and communication (IconDSC), IEEE
11. Budihal SV, Banakar RM (2020) Evidence-based dynamic radio resource allocation to mitigate inter cell interference employing cooperative communication. IET Commun 14(12):1848–1857
12. Budihal SV et al (2019) User location-based adaptive resource allocation for ICI mitigation in MIMO-OFDMA, ICCNCT 2018. Springer, Singapore

Rice Leaf Disease Detection Using Different Models and Comparative Analysis

M. Anu Mohan, M. Sundharam, G. Shanmmukha Srinivasa Reddy, G. Ramesh, and L. V. Subbareddy

Abstract India is a land of agriculture and produces different types of crops in large quantities. And among those crops, rice is majorly produced in India. With that topic aside, Indian farmers are also uneducated, which makes the farmers liable and exploitable. They need to have an idea of the diseases which have affected their crops and what to do with the occurring problem. Without the proper guidance, the farmers who entirely depend on the crop will definitely get affected and some leads to death. After much research, the researchers came to the conclusion of using the modern technology called the deep learning model, which helps in finding the disease more efficiently. Among the model, the convolutional neural network (CNN) model shows more accuracy. Here, we have created our own dataset and have also used transfer learning where we have made a comparison of the results between two models, and moreover, we are giving remedy for each diseased plant. In this way, farmers will have an idea what to be done regarding the disease.

Keywords Deep learning · Transfer learning · Convolutional neural network (CNN) · Resnet50

1 Introduction

Rice is widely used as a staple meal in many countries, including India. It is vulnerable to numerous illnesses at various points in its life cycle. Due to the large areas under each grower's supervision, the wide variety of illnesses, and the frequent coexistence of many diseases on a single plant, early diagnosis and treatment of these diseases is not only desirable but essential for ensuring high yields and the best possible quality.

M. Anu Mohan (✉) · M. Sundharam · G. Shanmmukha Srinivasa Reddy · G. Ramesh · L. V. Subbareddy
New Horizon College of Engineering, Marathalli, Bangalore, India
e-mail: anumohan09123@gmail.com

M. Sundharam
e-mail: dr.meenakshis@newhorizonindia.edu

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_2

Complex because there are so many plant diseases to contend with. Specialized agricultural expertise is difficult to acquire in rural locations. This calls for some kind of automated mechanism. Research is being undertaken utilizing different machine learning algorithms to enhance plant disease detection accuracy and lessen farmers' workloads. Artificial neural networks [1] and support vector machines [2–4] are two such examples of algorithms.

The feature selection techniques used have a significant impact on the efficacy of these algorithms. Recent developments in convolutional neural networks for image-based recognition have made feature selection and pre-processing unnecessary. Moreover, it is challenging to obtain huge datasets for such circumstances. Using a model trained on a bigger dataset is recommended when working with a small dataset. Transfer learning is a technique for constructing a model that can be used again by discarding the last completely connected layer or modifying the final few layers to function more precisely for the current task at hand. Extractor of Features.

Since almost everyone in these days has access to a smartphone, we came up with the notion of an automated system in which farmers could snap images of damaged leaves and upload them to our server, where a neural network would then make a diagnosis. After that, the farmer will know how to proceed based on the disease's categorization. In this study, we present a framework for an automated system's illness categorization functionality. In this work, we used data on rice diseases gathered over the course of the last several months to construct a deep learning approach for error analysis. Analysis The research on convolutional neural networks presented in [5–9] and [5–8] serves as a major source of motivation for us. We utilized a VGG-16 model that has already been trained (on ImageNet Big Data) and adjusted the fully connected layers using transfer learning so that they better suited our data. Finally, we attempted to trace back the origins of the mistakes by doing an error analysis.

2 Related Work

Numerous studies have shown that the outcomes are totally reliant on the feature selection, methods, and dataset [10]. Due to its excellent identification accuracy, CNN is the sole technology that has caught the attention of numerous researchers.

A. *Plant Disease Detection Using CNN*

When training their CNN model, the authors of this work [11] had access to 87,848 pictures. There are 58 different kinds of healthy plants shown here among 25 different plant species. The best model developed by the authors had an accuracy of 99.53%. The authors' [12] CNN model for this job was trained on 54,306 pictures, including those of 13 crop kinds and 26 sick and healthy leaves. The key problem here is that they employed a fake dataset, reducing the accuracy of their findings from 99.35% to only 31.3%.

B. *Rice Disease Detection Using CNN*

A set of 227 images was used to train a convolutional neural network (CNN) [8]. Using AlexNet, the author's accuracy for this task is 91.23%. However, all it can do is predict whether a plant is susceptible to disease. Author [7] in this work designed a structure inspired by Le-Net and AlexNet. Using 500 images and the above structure, they were able to reach 95.48 percent accuracy. Here, images are scaled down to 512*512, and random pooling is used to avoid trapping instead of maximum pooling.

3 Types of Diseases and Dataset Description

The rice leaf photographs were gathered from farms in the villages of Peeleru and Chitvel in the Indian state of Andhra Pradesh, with additional images given by Syngenta of Hyderabad, Telangana (region: Hyderabad). We shot with a Sony A7 and iPhone 14 Pro equipped with a 28–77 mm lens. Disease and symptom data were given by the International Rice Research Institute (IRRI) and Syngenta. We employed augmentation techniques since our data collection was too tiny. We captured images of 3847 rice leaves with various diseases and defects, such as leaf blight, hypsa, brown spot, and leaf blight. There are 459 photos of leaves in good health. Single sickness or numerous diseases on a single sheet were both obstacles that needed to be overcome when gathering photos. To address the aforementioned issue, we amplified and adjusted the photographs. We also made advantage of zoom, horizontal positioning, rotation, and usability tactics when training the CNN.

A. leaf blast

The fungus Magnaporthe oryzae is responsible for the illness. The early signs are elliptical or spindle-shaped patches of white or gray–green with dark reddish to brownish edges. Some have diamond-shaped hubs that are wide in the middle and taper to pointed ends. Fig. 1a depicts a lesion with a spindle form, white centers, and a dark brown border.

B. leaf blight

Xanthomonas oryzae is the bacterium that causes it. The diseased leaves roll upward and become a grayish green until fading, straw-colored leaves appear, and lastly, death after withering. The lesions advance toward the base and have wavy edges. Bacterial slime that resembles morning dew drops can be seen on new lesions. Leaf blight-affected leaves are shown in Fig. 1b.

C. Brown Spot

A bacterial infection is defined as. Large, destructive blotches spread quickly over infected leaves. Small, brown to purplish brown lesions on the leaves are the first

Fig. 1 Top left to bottom right, from **a** to **d**. This epidemic might be caused by **a** leaf blast, **b** leaf blight, **c** brown spot, **d** hispa

visible sign. Gross lesions are ovoid to spherical in shape, with a brown or gray core and a reddish brown border caused by the fungus's toxic secretions. Brown spot, seen in Fig. 1c, causes tiny, dark brown lesions on leaves.

D. Hispa

The damage is brought on by the adults and larvae of the rice hispa, Dicladispa armiger. Adult beetles scrape the top layer of leaf blades off, only leaving the underside. Eggs are often laid in small cracks toward the tip of the delicate leaves. As observed in Fig. 1d, the grub is flat and pale yellow in color.

4 Proposed Methodology

The effectiveness of multi-layered networks like convolutional neural networks (CNNs) is determined on their structure. Convolution layers, pooling layers, and fully connected layers make up its architecture. The first two layers work together as a feature extractor, while the third serves as a classifier. When features are retrieved from a convolutional layer, their dimensionality may be decreased by using a pooling layer. To identify the photos with the help of the extracted feature, a fully connected layer is utilized, followed by softmax. In order to extract features from an input picture, a convolution layer uses a collection of trainable filters. To create a 2D feature map, we stack the raw image's pixels on top of each filter's output using a sliding window. Rectified linear unit (ReLU) is the most used activation function. Maximum pooling is a subsampling layer that reduces the size of the feature map. Object feature map is then fully linked to each one by the fully connected layer.

To categorize the photos, softmax creates a multiclass issue and assigns a decimal probability to each class3.2 million fine-tagged pictures from 5247 classes in the ImageNet database were used to pre-train Resnet50's weighted deep convolutional network [9]. The network's architecture is shown in Fig. 2. Thus, by applying translation learning to such pre-trained models, information in the form of weights, structure, and features learnt in one domain may be transferred to another domain. The attributes of the first few levels are more abstract, whereas those of the subsequent layers are more data-driven. Finally, our model incorporates a softmax layer of size 4 (as there are 4 classes) on top of a dense layer of size 128. Blocks of the first five convolutional layers are frozen in CNN to operate as a feature extractor, giving it an edge over more conventional approaches. To generate evaluations, we retrained and tweaked a pre-trained Resnet50 model. There are four different transfer learning approaches that may be used, each appropriate in different situations.

- The new dataset is smaller and more similar to the original dataset.
- The new dataset is larger and more similar to the original dataset.
- The new dataset is larger but not identical to the original dataset.
- The new dataset is smaller but not identical to the original dataset.

Due to the lower amount of classes in our dataset compared to ImageNet, we fixed the Resnet layers to be utilized as feature extractors until the final layers were constructed. Thus, the third scenario is the best match for our model. The layout

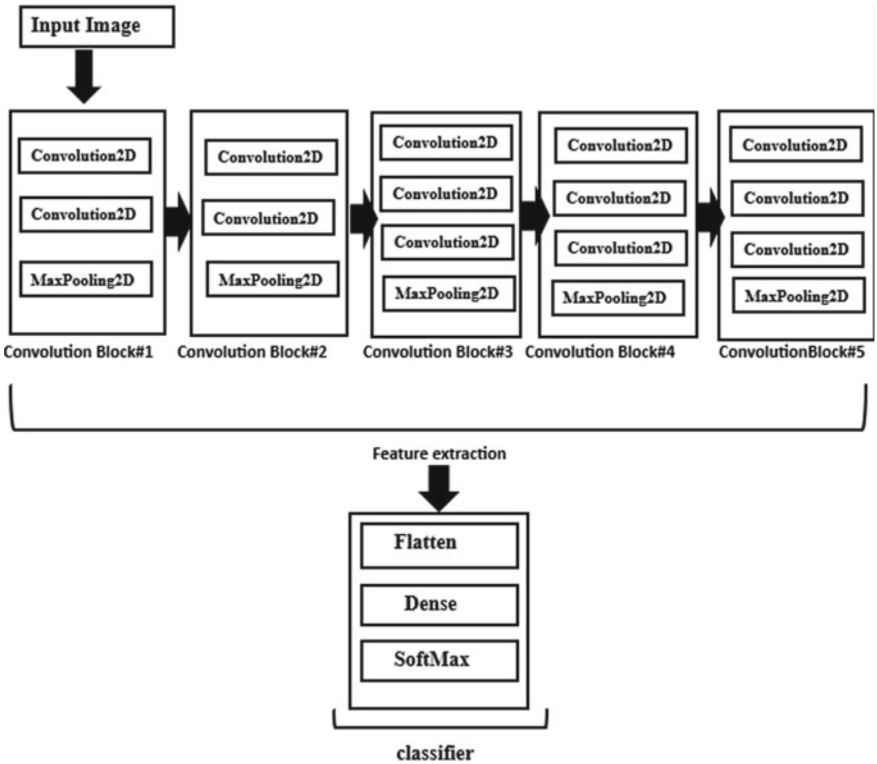


Fig.2 Resnet architecture tuned with last two layers with 64 dense FC layer and 5 dense layers as output

of the suggested model is shown in Fig. 2. In addition, we create a fully—convolutional neural network (CNN) model that omits transfer learning and comprises of four convolutional layers, ReLU, maxpooling, dropout layers, two fully—connected layers, and softmax. The performance, however, is not as good as the preceding model. The similar comparison is used in the discussion of findings.

5 Implementation

A. *Experimental Setup*

The simulation was run on a Windows 10 PC with a UHD graphics processing unit card and a 64-bit operating system. For this project, we deployed keras 2.2.4, a deep learning framework built on top of TensorFlow and Python 3.7.2.

B. *Image Acquisition*

The field and Syngenta photos are safe with us. Leaf blast, leaf blight, brown spot, and hispa are the five classifications represented in the dataset.

C. *Image Pre-processing*

The collected images are resized to 128*128 pixel of height and width, and batch size is 20. Using ImageDataGenerator in keras to generate new images, we have used horizontal and vertical shift, zoom, rotation as a augmentation techniques.

D. *Model Training (CNN)*

An image dataset is loaded for training and testing. Images and class labels are stored in the corresponding array for training. The train_teat_split function is used to divide the data into a training set of 70% and a testing set of 30%. From the remaining 30%, we're using 20% for validation purposes. Afterward, we move on to the Resnet50 model, for which we dump the model into keras after removing the completely linked layers. The subsequent layers are also rendered un-receptive to training. We flatten the feature extractor output, followed by a dense undo output layer with softmax. Our proposed model uses Adam optimizer and categorical_crossentropy as loss function. The number of epochs we used was 30, and we stopped after that because we did not see any improvement in the graph.

E. *Transfer Learning Model (Resnet50)*

Transfer learning defines as the situation what we have learned at one part and used to improve generalization in the next part. Generally, the main advantage of the transfer learning which all the people like is that it takes less time to the model when compared to other models. As we know, we usually need large data to train the model and takes more time to train the model. Since the model was already pre-trained, we employed transfer learning to reduce the data size as well as the training data Fig. 3.

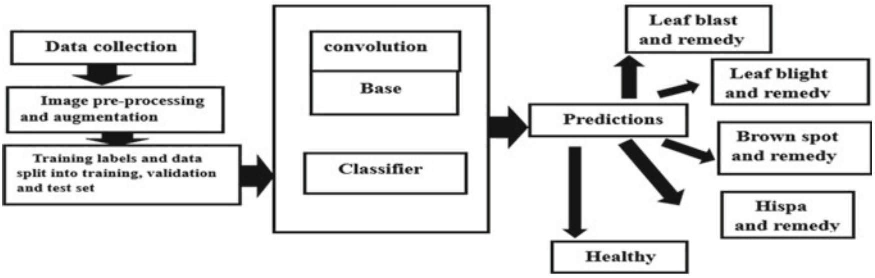


Fig. 3 Overview of the transfer learning

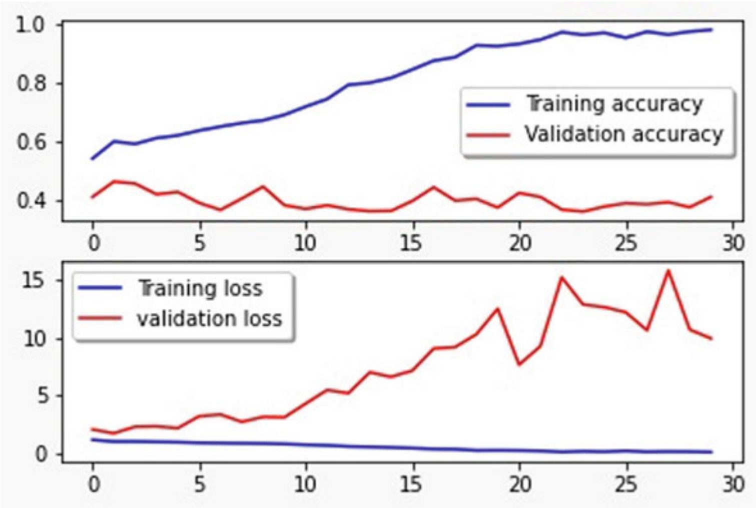


Fig. 4 Accuracy plot of CNN without transfer learning

6 Results

A. *Calculations*

The system that we suggested runs for 30 iterations using 2064 training and 881 test data, with an accuracy of 98.06% when using CNN, 81.68% when using Resnet50 as a transfer learning model, and 65.08% when using VGG16. The best accuracy was still 81.68% even if the batch size, number of epochs, and dropout 0.2 were all set to 20, 30, and accordingly. The CNN contains four convolutional layers, two fully connected layers, and softmax, but no transfer learning (Figs. 4 and 5; Table 1).

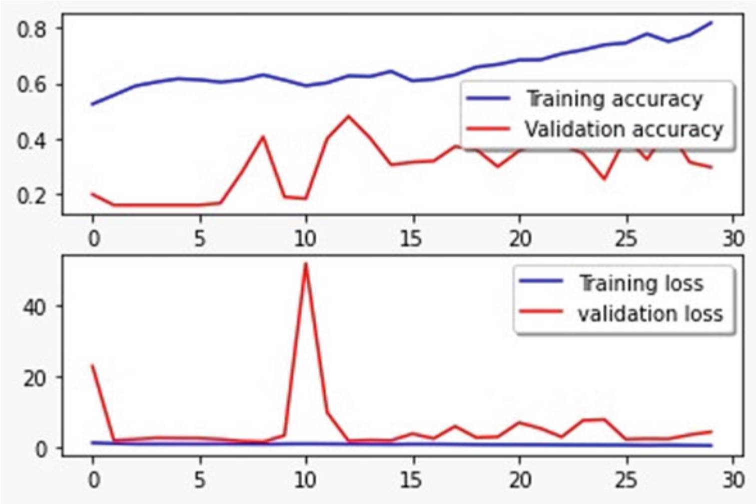


Fig. 5 Accuracy plot of CNN with transfer learning using Resnet50

Table 1 Accuracy comparison between CNN with transfer learning and without transfer learning

Model	Test accuracy (%)
CNN	98.06
CNN with transfer learning using Resnet50	81.68

7 Conclusion

With 2064 pictures used for training and 881 images used for testing, the deep learning system suggested in this study achieved an overall accuracy of 98.06% and a transfer learning accuracy of 81.68%. In this case, we capped the number of epochs at 30, since further iterations did not improve accuracy. During model training, only a small fraction of the training data is lost.

We want to enhance the transfer learning model and expand the dataset by include more photos in the near future. Our goal is to enhance our model so that it may be used more effectively in practical settings. In addition, we want to expand our model's detection capabilities to include additional, more consequential plant diseases in India.

Acknowledgements We would like to thank Mr. Rama Reddy and Mr. SubbaReddy who helped us during the process of collecting the rice leaf in the fields of Pileru and Chitvel in Andhra Pradesh state. We would like to extend our gratitude to Syngenta company in helping of classification of diseases and providing some images for the dataset.

References

1. Liu L, Zhou G (2009) Extraction of the rice leaf disease image based on BP neural network. In: International conference on computational intelligence and software engineering, pp 1–3
2. T Gupta 2017 Plant leaf disease analysis using image processing technique with modified SVM-CS classifier Int J Eng Manag Technol 5 11 17
3. Es-saady Y, El Massi M, EI Yassa M, Mammass D, Benazoun A (2016) Automatic recognition of plant leaves diseases based on serial combination of two SVM classifiers. In: 2016 international conference on electrical and information technologies (ICEIT), pp 561–566
4. Padol PB, Yadav AA (2016) SVM classifier based grape leaf disease detection. In: 2016 conference on advances in signal processing (CASP), pp 175–179
5. S Arivazhagan SV Ligi 2008 Mango leaf diseases identification using convolutional neural network Int J Pure Appl Math 120 6 11067 11079
6. Liu B, Zhang Y, He D, Li Y (2012) Identification of apple leaf diseases based on deep convolutional neural networks Symmetry. Niu XX, Suen C (2012) A novel hybrid CNN–SVM classifier for recognizing handwritten digits. Pattern Recogn 45:1318–1325
7. Y Lu S Yi N Zeng Y Liu Y Zhang 2017 Identification of rice diseases using deep convolutional neural networks Neurocomput 267 378 384
8. RR Atole D Park 2018 A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies Int J Adv Comput Sci Appl 9 1 67 70
9. Simonyan K, Zisserman A (2015) Very deep convolutional networks for largescale image recognition. Arxiv preprint arXiv. In: Proceedings of 2020 IEEE Calcutta conference (CALCON), pp 1409–1556
10. V Singh A Misra 2017 Detection of plant leaf diseases using image segmentation and soft computing techniques Inf Process Agric 4 1 41 49
11. Ferentinos KP (2018) Deep learning models for plant disease detection and diagnosis. Comput Electron Agricvol 145:311–318
12. Mohanty SP, Hughes DP, Salathé M (2016) Using deep learning for image-based plant disease detection. Frontiers Plant Sci 7:1419

5G Mobile System Drawbacks and Limitations

Jalal J. Hamad Ameen

Abstract Motivation of the researchers was toward higher data rate and lower latency motivated researchers to work forward fifth generation (5G) of mobile technology, this technology like other technologies not clear from drawbacks and limitations. In this paper, these drawbacks and limitations have been explained with evidence and proofs. Starting with the frequency bands given to the system and challenges to the system, these challenges as an architecture or world situation are suitable for the system completion, competition between companies in the world to finish the system design and then application, finally, is the 5G mobile system that will be really available as an applicable system like 4G mobile system, or because of the given drawbacks and limitations will be jumped to 6G system, details have been presented in this paper.

Keywords Mobile system · 5G frequency bands · 5G network architecture

1 Introduction

The evolution of mobile communication system starting with first generation (1G) the analog system and frequency modulation (FM) with its very low security, then second generation (2G) called Global system mobile (GSM) the digital system with low data rate, also, third generation (3G) called universal mobile telecommunication system (UMTS) with limited internet connection services leaded to fourth generation (4G) to be designed starting with long term evolution (LTE) to which multiple input multiple output (MIMO) has been used size 2×2 is continue still in process toward 6G because the more demands of speed and connectivity of devices as well as the drawbacks of 5G.

4G system evolved to LTE—advanced using 4×4 MIMO with improved data rate, and higher number of users, finally, LTE-pro, to which 8×8 MIMO is used, the indoor data rate is 1Gbps and outdoor with about 100 Mbps. From 1 to 4G evolution,

J. J. Hamad Ameen (✉)
Electrical Engineering, College of Engineering, University of Salahaddin—Hawler, Erbil, Iraq
e-mail: jalal.hamadameen@su.edu.krd

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_3

the system was developed by using different frequency spectrums with different coverage and specifications, at the same time the earlier system not affecting the later system.

Fifth generation (5G) system has been designed and tested in some countries, nowadays, some of these countries using it real time as a service. Generally, mobile cellular industry introduces starting with 2G mobile system, third-generation mobile system presented after 2000, 4G long-term evolution (LTE) ten years later, and now finally 5G since 2020 and still not confirmed finally. Within that evolution of mobile generations from 2 to 5G, the performance of G the radio access network for every technology cycle comes with less node-B coverage because of higher frequencies, this leads to higher capacity and throughput related radio access network. The air interface, including the physical layer and the protocol stack is the main part toward the design of the radio access network for any wireless network including mobile system. RAN is the connection of many base stations in GSM called BTSs, in 3G called Node-Bs, in 4G called eNode-Bs and finally in 5G called C-Node-Bs and connected to the core network, to obtain higher data rate and best signal quality.

5G mobile system like other mobile systems is not clear from some disadvantages, the problem is that 5G system as well as the disadvantages, according to this paper, have limitations and drawbacks, expected to stop developing this system and jump to sixth generation (6G) system, in this paper these limitations and drawbacks are presented with examples.

2 5G Mobile System

There are three goals and at the same time challenges in 5G mobile system, they are enhanced mobile broadband (eMBB), ultra-reliable low latency communications (uRLLC) and massive machine type communication (mMTC). In order to design the radio access network for the overall system and to reach the quality as an IMT-2020 radio interface and assigned spectrum, it has to fulfill certain technical requirements specified in Recommendation ITU-R M.2083 [1]. As shown in Fig. 1, the three goals of 5G system according to IMT 2020 applications would fall into the following three broad usage scenarios. The eMBB deals with enhancing the mobile broadband for human-centric scenarios which expected that these will spur innovation that will bring newer applications and services that are not available today. The second goal is to the lower latency with reliable communications (URLLC) deals with throughput, latency and availability requirements.

The third goal as shown in Fig. 1 is the massive machine type communications means more than 1 million users per square meters, are hardly anything equivalent as of now and it remains which deals with the use case of users [2]. As shown in Fig. 2 that the scenarios for the three 5G goals requirements according to IMT—2020 needs some other plannings like mobility, spectrum efficiency area traffic capacity, connection density [3].

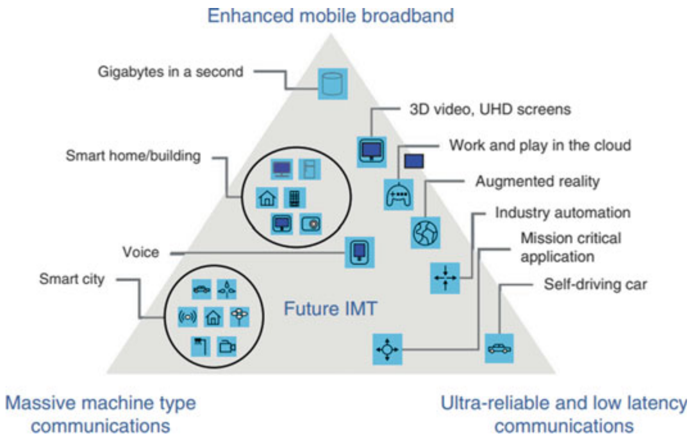


Fig. 1 IMT—2020 usage scenarios. *Source* ITU-R

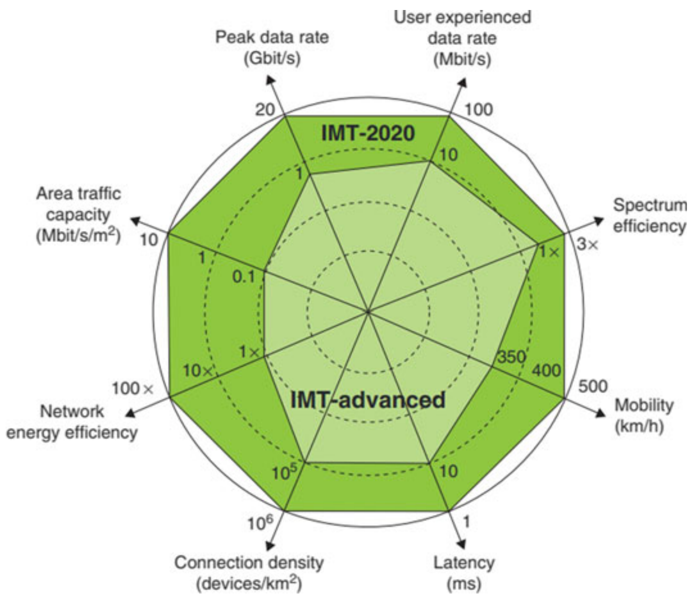


Fig. 2 IMT-2020 requirements. *Source* ITU-R

Table 1 gives the summary of the IMT—2020 requirements, these requirements needed to reach the three goals of this system of mobile technology.

5G mobile system started since 2017 with phase 1 and phase 2 as shown in Fig. 3 which named release 15, this release finished end of 2017 [1]. Continuously, release 14 for 4G release 15 phase 1 of 5G was processed, as shown in Fig. 3 during 2017 release 15 in three steps completed, then, phase 1 started [4].

Table 1 Summary of IMT-2020 requirements

Capability	Description
Peak data rate	10–20 Gbps
User-experienced data rate	100 Mbps–1 Gbps
Latency	1 ms
Mobility	500 km/h
Connection density	10^6/km^2

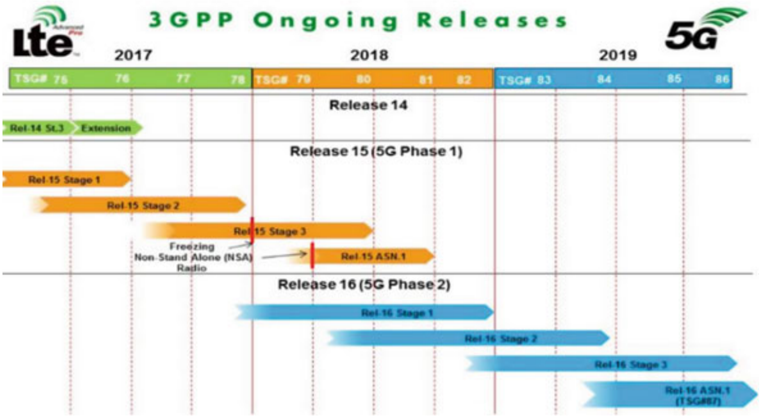


Fig. 3 Release 15 phase 1 and phase 2 timeline. *Source* 3GPP website

Release 16 was divided into quarters since 2018 quarters 3 and 4, year 2019 includes Q2, Q3 and Q4 for release 2016 which was an expansion items as shown in Fig. 4.

In Fig. 5 release 17 shows starting from end of Q4 2019 until 2020, as shown between 2021 there is a gap of evolution because of the pandemic corona COVID-19, this was one of the main drawbacks of 5G mobile system.

3 Aim of the Research

As mentioned in abstract above, the aim of this research paper is to present the drawbacks and limitations of 5G mobile system, the representation will be with evidence and proof, these lead to expect that this system will not be continued in development and there will be jump to 6G system in order to control these limitations and drawbacks in 5G system.

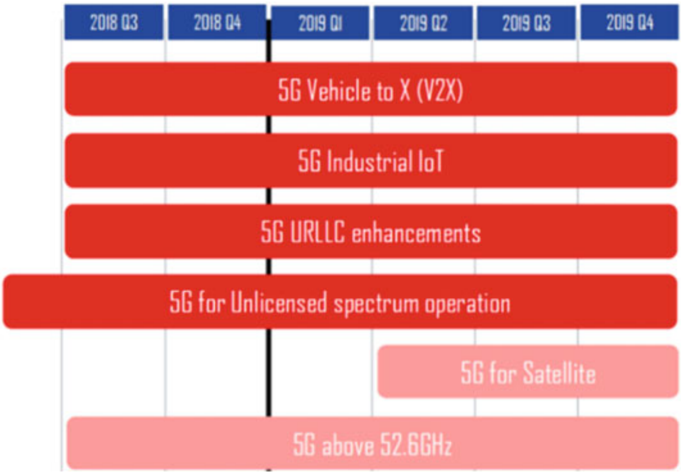


Fig. 4 3GPP release 16 expansion items. *Source* 3GPP website

Fig. 5 3GPP release 17 timeline. *Source* 3GPP website

4 5G Mobile System Drawbacks

As mentioned, the goal of this paper is to present the drawbacks and limitations of 5G mobile system, below are these drawbacks and limitations:

4.1 5G Frequency Spectrum

5G mobile system started and still uses sub-C band with 6 GHz, the plans toward using mmWave with 28 GHz are the main drawbacks of this system because of the risk causes from the electromagnetic wave and radiation against human, birds, and even trees and environments, the evidence for this is the wide demonstration throughout the world to stop 5G system specially using 28 GHz test.

4.2 Connectivity Obstrusions

The connectivity range of 5G is not great as the frequency range which is only for short distances, while 5G frequency will be interrupted by physical obstructions like trees, other towers walls and buildings. This affects the signal to be absorbed, this motivates the 5G industries to take into account this drawback, otherwise, the system signal quality will not be acceptable.

4.3 High Cost

To ensure the high-speed connectivity needs higher cost in the 5G infrastructure or even the adaptations between 5G and other mobile technologies like 4G because the 5G development to the existing cellular infrastructure will be high. Cellular operators are looking to minimize these costs by exploring alternative options in the form of network sharing, this is another drawback of the 5G system.

4.4 Rural Access Limitations

The real connectivity for urban areas is also another drawback for 5G, those living in the rural settings will not necessarily benefit from the connection. As a result, the coverage of 5G is small because the higher frequency spectrum then needs higher number of base stations that causes higher effect in these rural areas and only some of the population will benefit from 5G communication.

4.5 Drain of Batteries: Battery Used in Devices

The drain of batteries is also another drawback of 5G because the battery technology needs to design it for more life of the battery, while the users use more long time in the network. As a result the temperature of the device increases during the operation of 5G by the user.

4.6 Mismatch of Speed between Upload and Download Upload

Another drawback of 5G system is the speed mismatch between upload and download in some cases up to 1.9 Gbps. However, the upload speeds are rarely more than 100 Mbps. This problem is still not solved, even the companies started with 5G system.

4.7 Cellphone Towers

Another main drawback of 5G system is the towers, because the higher frequency leads to more affect the human as well as birds and even environment, specially the mmWave with 28 GHz, for example, in Nederland a test has been done with this frequency all the birds in the tested area died at once, therefore these cellphone. towers, is not welcomed by most communities. With all new technology, there are going to be some initial drawbacks as the system becomes more refined and the range becomes more extensive.

4.8 Political Causes

Political causes also will be one of the main effects of 5G mobile system and will be the drawback and limitation like situations, competitions and challenges between countries. Tracking the news before and after the 5G system proves this point, which leads to not finalizing it by the companies.

5 Conclusions

This research paper is an attempt to show the 5G mobile system drawbacks and limitations, as shown 5G mobile system expected not to be finalized and there will be a jump toward 6G mobile system, due to the 3GPP plan there are gaps between phases specially during corona pandemic 2020–2021, the competition between countries in 5G system design, also, the frequency bands assignments is another reason which conflicts with other generations of mobile systems, although heterogenous networks have been proposed to adapt between 5G and other mobile systems but is not successful method which causes another limitation for 5G mobile system.

References

1. Weerasinghe TN (2021) Enabling mMTC and URLLC in 5G: initial access, traffic prediction, and user availability
2. Hu RQ, Qian Y (2013) Heterogeneous cellular networks, Wiley
3. Asif SZ (2019) 5G mobile communications concepts and technologies. CRC Press, Tayler & Francis Group
4. 4G America (2015) The voice of 5G in the America, 5G spectrum recommendations
5. 5GPPP Architecture Working Group (2017) View on 5G architecture, 5G PPP, Version 2.0, December 2017, white paper
6. Karki RS, Garia VB (2016) Next generations of mobile networks. Int J Comput Appl (0975–8887), International conference on advances in information technology and management ICAIM
7. Ali KS (2018) Modeling, analysis, and design of 5G networks using stochastic geometry, PhD Dissertation, King Abdullah University of Science and Technology Thuwal, Kingdom of Saudi Arabia
8. Penttinen JTJ (2019) 5G explained security and deployment of advanced mobile communications, Wiley
9. Kanatas AG, Nikita KS, Mathiopoulos P (2018) New directions in wireless communications systems from mobile to 5G, CRC Press Taylor & Francis Group
10. Sirotkin S (2021) 5G radio access network architecture the dark side of 5G, Wiley
11. Luo H (2017) System capacity enhancement for 5G network and beyond, PhD Dissertation, University of Bedfordshire

Survey on Implementation of Machine Learning in Cloud Security

Divyesh Bhatnagar and Anil Gupta

Abstract The need to access data quickly without increasing the physical memory in devices like computers and mobile made big IT firms, IT persons and individuals move to cloud computing, but the confidentiality of data is the main concern. There are various stages of data travel while stored onto the cloud and many breaches between these stages of data storage. Many advanced technologies can secure data on the cloud; machine learning technique is one of them. The machine itself will manage the security with less human involvement. This paper is focused on cloud computing and the implementation of machine learning (ML) techniques and models in the field of cloud security to protect data from threats and malware.

Keywords Cloud computing · Big data · Security · Machine learning · Intrusion keyword

1 Introduction

Cloud computing is a way to build new infrastructure, train new individuals, or increase capacity without investing much in newly licensed software but as nothing goes without challenge, security of confidential data is a major concern. This paper presents a quickly enabled model of cloud computing which provides facilities like services, servers, applications and networks without increasing infrastructure covering implementation of machine learning and the management of data security with the help of classifiers for different kinds of service models (SaaS, IaaS and PaaS). There are features like *self-service on-demand*, *broader network access capability, resource pooling* and *rapid elasticity.* The cloud system is self-optimized, and the

D. Bhatnagar (✉) · A. Gupta
MBM University, Jodhpur, India
e-mail: divyeshbhatnagar@gmail.com

A. Gupta
e-mail: anilgupta@jnvu.edu.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_4

usage of resources is monitored and reported transparently by both of the provider and user ends.

The paper structure is as follows: Introduction section will discuss cloud computing and its essential characteristics [1], followed by security issues in cloud security models [2]. Later, a brief introduction to machine learning and different classifiers is given which will help secure the cloud.

2 Cloud Security

Cloud computing is now moving on to applying software and implementing databases in many data centers where the facilities are not trustworthy despite the easy deployment, low-cost management, low-cost disaster recovery and scalability. The newer models provide security, but many vulnerabilities must be cured like web applications (SQL injection, physical access, cross-site scripting), accessibility and virtualization.

2.1 Security Issues with the Service Models [2]

Three service models of cloud computing SaaS, PaaS and IaaS provide a platform for resources, and the application provides services to the users.

Security issues in Software as a Service (SaaS) [2]. Users cannot change the security as per their needs and are dependent on security measures provided to them. Despite SaaS provider storing data replicas (transactions, records, information, etc.) at multiple locations for maintaining availability, there is a lack of control. The main concern is about the breaches, availability and vulnerability of the data. SaaS security is shown in Fig. 1. SaaS providers work on three models:

Independent Identity Management stack: Provider maintains all the users' data like username and password.

Credential synchronization: Provider exchanges the user credential and information between the company and the application.

Federated identity management: The client's identity and attributes are modified on provider's request so that the company provides permission for access and login.

The identity management and login process of the SaaS provider are assessed by authentication weakness analysis and insecure trust configuration.

Security issues in Platform as a Service (PaaS) [2]. When using the enterprise service bus, weakness score and patch coverage are used for checking the security measurements and quality of the program's code but hackers may react to the new architecture on the cloud and attack the framework whose design code is easily accessible by broad black-box test.

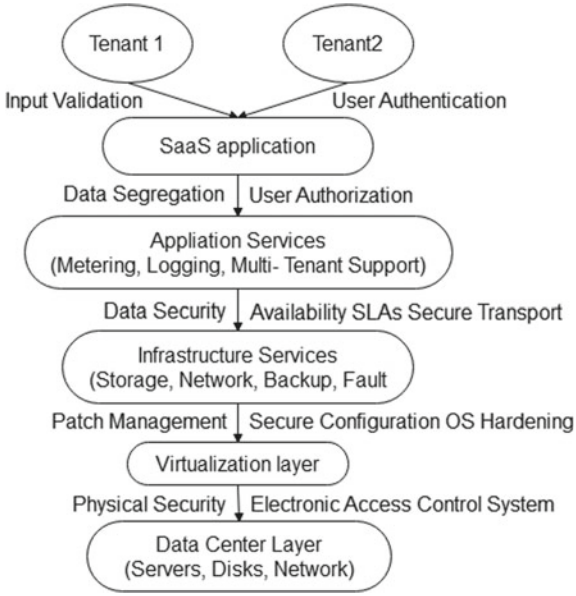


Fig. 1 Security for SaaS

Security issues with Infrastructure as a Service (IaaS) [2]. In IaaS, software developers get better security than others. Providers and consumers have different responsibilities regarding security where the consumer must manage security related to the operating system, data and application on their own.

Effect on Deployment Issues. When deployed through the cloud model, the public cloud suffers more making the private cloud a bit safer. Hackers can bypass the route of the data if encryption and security protocol used are not context-oriented and the requirement of proper precaution with the security threats does not meet.

3 Related Work

As per the paper [3], intrusion detection system is divided into three parts: network-based, host-based and hybrid detection system. In further work, paper [4] uses UNSW and ISOT datasets. Dataset statistics of USNW data are shown in Table 1.

Authors detect nine types of threats: fuzzers, analysis, backdoors, denial of service, exploits, generic, reconnaissance, shellcode and worms. For class labeling, Argus, BroIDS tools and 12 algorithms are used. ISOT dataset is two different datasets obtained from Honeynet projects (Storm and Waledac botnets) for the detection of defective traffic. Dataset characteristics of the ISOT dataset are shown in Table 2.

The techniques used are J48 decision tree, naïve Bayes, logistic regression, SVM-RBF, SVM-polynomial and SVM—linear on WEKA tool and comparison about overall performance of ML techniques as shown in Fig. 2. As observed, accuracies

Table 1 UNSW dataset [4]

Dataset	Total records	Normal records	Effected record
Training	175,340	56,001	119,340
Testing	82,330	37,001	45,334
Total	257,670	93,002	164,674
Percentage (%)	100	36.0	64.0

Table 2 ISOT dataset [4]

Traffic type	Unique flows	Percentage (%)
Defected	55,905	3.32
Normal	1,619,522	96.67
Total	1,675,427	100

with LR, J48 and SVM algorithm are respectively 89.26%, 88.67% and 68.06%. The prediction condition in ML is shown in Table 3.

The logistic regression is working best with 93.7% and identifies true positive anomalous packets and 6.3% of false negative erroneous cases. J48 shows 84.6% TP and 15.2% FN while naïve Bayes algorithm performs worst and gives 56.2% TP and 43.8% FN. Figure 3 shows true negative and false positive. The experiment has found that SVM-RBF kernel gives the best performance with 95.9% while true negative and linear regression gives 95.7%. WEKA with OCTAVE is used. The

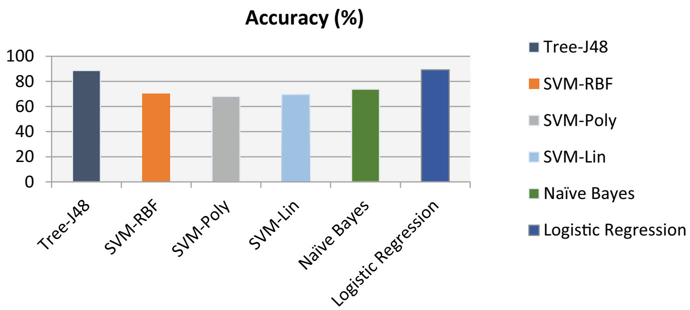


Fig. 2 Overall performances of ML techniques at UNSW [4]

Table 3 Prediction condition in ML [4]

Total population	Positive condition (defective)	Negative condition (Normal)
Positive (1, defected)	True Positive (TP)	False Negative (FN)
Negative (0, normal)	False Positive (FP)	True Negative (TN)

change probability threshold to categorize packets between normal and defective is set at the value 0.5. Value more significant than 0.5 makes it normal packet while lesser makes it defective (Figs. 4, 5, 6 and 7).

In conclusion, authors got the highest accuracy of 43% which is low compared to 93.7% making it very dangerous in real-time environment. In the paper [5], authors discuss service behavior by examining the metrics collected from all cloud infrastructure components from various layers. Various classification algorithms for detecting anomalies have been evaluated. The authors used a support vector machine and the

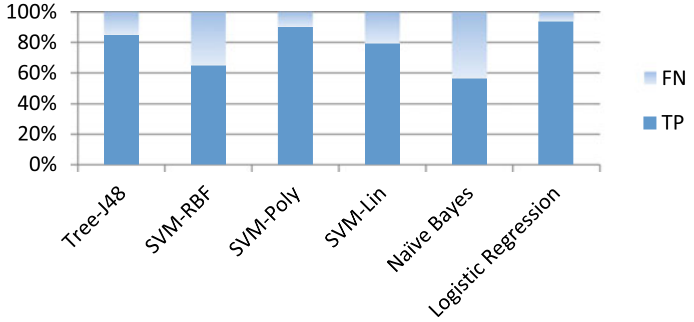


Fig. 3 Accuracy % of true positive and false negative with UNSW dataset [4]

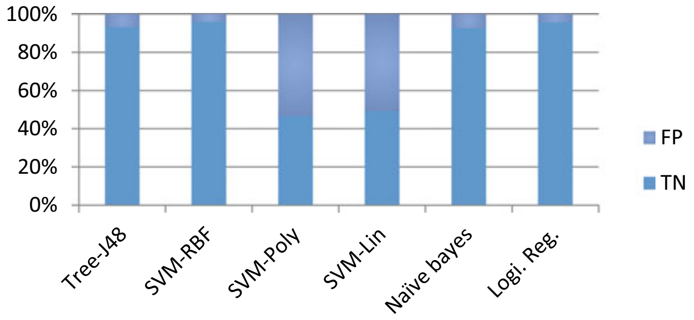


Fig. 4 Accuracy % of true negative and false positive with UNSW dataset [4]

Fig. 5 Overall performances of ML techniques on ISOT [4]

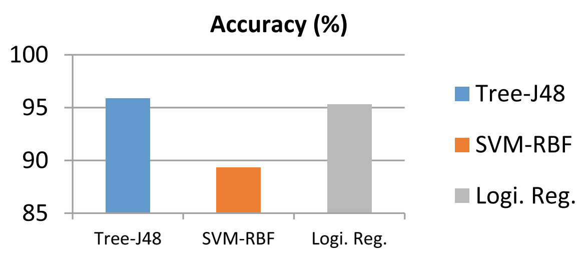


Fig. 6 Accuracy % of true negative and false positive with ISOT dataset [4]

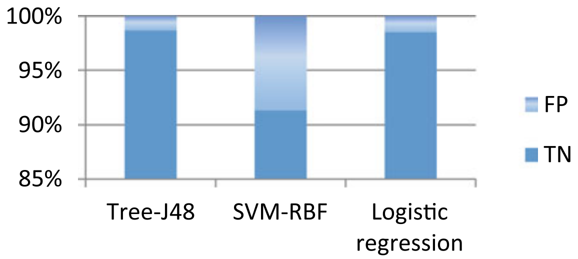
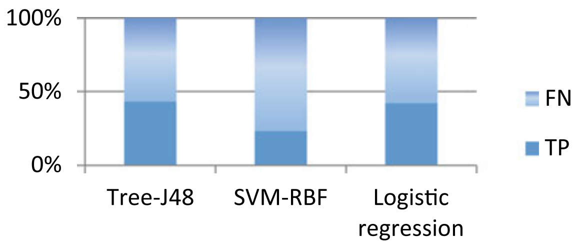


Fig. 7 Accuracy % of true positive and false negative with ISOT dataset [4]



decision tree. These algorithms use a common interface for testing and training the machine. The training function *train* $(S_1, S_2 \ldots S_n)$ where S_i belongs to $(\mathbb{R}^{\mathrm{m}} \times L)$, m is the sample metrics taken from the observations and L denotes the possible labels are $L = \{\text{normal, abnormal}\}$. As the train function, there is a classify function that works on the unlabeled data and converts it into predicted label data *classify* $\mathbb{R}^{\mathrm{m}} \rightarrow L$. This function works very fast on the real-time data as the training process is an offline phase. Finding out the abnormal samples is challenging, so the artificial injection of anomalies is used into the host. In the paper [6], authors discussed the ensemble-based multi-filter feature selection method that concentrates on classifying packets into signature and anomaly-based. In a signature-based attack, a signature is used in the knowledge database to find an attack and anomaly-based check that the fixed time slot deviates from the expectation. Signature-based is applicable in the detection of already known attacks while anomaly detection is used in the zero-day attack. The data is very large, so feature selection is divided into three parts: filter, wrapper and embedded. Ensemble-based multi-filter feature selection has combined output of information gain, gain ratio, chi-squared value and ReliefF for selecting the feature. The NSL-KDD dataset and WEKA tool were used for knowledge analysis. For raking the data, the information gain parameter is taken. In this, the entropy value of the distribution is calculated. The entropy is calculated by-

$$H(A) = -\sum_i P(a_i)\log_2(P(a_i))$$

Here $P(a_i)$ shows the prior probability of A. Now the entropy of the A after calculating another variable B is

$$H(A|B) = -\sum_j P(b_j)\sum_i P(a_i|b_j)\log_2(P(a_i|b_j))$$

$P(a_i|b_j)$ is the posterior probability of A and B. Now the information gain is

$$\text{Information Gain }(A|B) = H(A) - H(A|B)$$

The gain ratio for improving the bias of information gain.

$$\text{Gain Ratio}(b, a) = \frac{\text{information gain}(b, a)}{\text{intrinsic value}(a)}$$

where

$$\text{intrinsic value}(x) = -\sum \frac{|S_{i|}}{|S|} * \log_2 \frac{|S_i|}{S}$$

$|S|$ represents the number of possible values of feature x, and $|S_i|$ is the number of feature x.

The chi-squared statistic is used for testing the score measure of two independent variables.

$$X^2(r, c_i) = \frac{N[P(r, c_i)P(\overline{r}, \overline{c_i}) - P(r, \overline{c_i})P(\overline{r}, c_i)]^2}{P(r)P(\overline{r})P(c_i)P(\overline{c_i})}$$

where N represents the dataset, r represents feature presence and $\overline{r}$ represents the absence of the feature c_i denotes class. $P(r, c_i)$ represents the feature probability occurrence of r of class c_i and $P(\overline{r}, c_i)$ is the feature occurrence or not in class c_i. $P(r, \overline{c_i})$ and $P(\overline{r}, \overline{c_i})$ are the probabilities of the do and do not occurrence in the not-labeled in c_i. $P(r)$ is the probability of feature appearance in the dataset and $P(\overline{r})$ is the probability not appearing in the dataset. $P(c_i)$ and $P(\overline{c_i})$ are the probabilities of the dataset do and don't labeled to the c_i. The function ReliefF uses continuous sampling feature of nearest hit and nearest miss from the same class and from different classes. The calculator appends weights to features as per their ability to differentiate different classes. This overall ensemble-based multi-filter feature selection method joins the output of information gain of ranked feature, gain ratio, chi-squared and ReliefF. As a result, output is determined by the combination of each filter. The authors reviewed Snort misuse of intrusion detection and various anomalies based on statistical algorithms (ALAD, PHAD, LERAD and NETAD) [7]. Intrusion detection has three phases: monitoring, detection and response. The learning algorithm divides the normal design into two aspects: non-stationary model which depends on the time span of the last event occurring, another is the monitoring of intrusion detection protocol, collection of identifying the unknown malicious activity.

The existing anomaly detection methods are stuck to the plain byte of the network packet payloads [8]. Valuable information is provided when the syntactic context is introduced with anomaly detection. The issues are with the protocol context in payload anomaly detection. The new data representation c_n-gram has been introduced. The experiment on the text-based and binary application layer protocol shows excellent accuracy on many attacks on regular anomaly detection. Four stages are used in the experiment: data acquisition and normalization; feature extraction; the similarity computation between strings; and detection of the anomaly. The result found a method that gives protocol analysis and payload-based anomaly detection. The text-based and binary application layer protocol give excellent accuracy in the detection of unknown attacks. The identified attacks are reliable and boosted detection accuracy with 44% using unsupervised anomaly detection compared with plain sequential feature to 80% using the combined feature in a situation of attack.

The authors [9] discussed the local outlier factor (LOF) algorithm for outlier detection of network intrusion. Their method with two phases learns both the training and testing behavior of the application which reduces the effort to collect training data before the detection and gets the ability to detect contextual anomalies with low computational overhead. During the training phase, the system is logged until enough statistics periodical data logs are collected for detecting an anomaly. The logging interval should not be less than 5 s to reduce the noise. The statistics of CPU utilization, disk IO activity and network activity are collected. Time variant statistic of storage is used; memory-free is not suitable for the detection method. There are two interchangeable phrases "collecting statistics" and "sampling a point". These points are normalized and stored in an R tree data structure for testing. In the detection phase, they keep new sampled points periodically. Every new point will be according to the training provided. If the calculated value is lesser than the threshold, then that point will be considered normal or anomalous. For the detection, the K value is used to represent the number of neighbors involved in the nearest neighbor search. The detection of LOF is not good according to the value of K. Normally, the value taken for K is 30. The knowledge is stored in the R tree which provides quick access to the LOF algorithm for searching the neighbor. In the knowledge updating phase, the knowledge keeps increasing after the training is done until the capacity is full, and with it, the system gets the detection capability. The knowledge base behaves like a sliding window along with the history of normal behavior. When the new cluster comes, the oldest one gets removed. The point can be detected as anomalous even if the point belongs to the normal cluster. The recent normal behavior will keep in for further LOF calculation. Without applying normalization, LOF algorithm will be too sensitive for the dimensions and the larger variances can make the algorithm ineffective in detecting anomalies of dimensions with smaller variances. In the renormalization after updating phase, the updating procedures create new issues with the normalization of knowledge. It increases and decreases the minimum bounding rectangle of the knowledge base when the expected behavior is changing. It will have a negative effect on the LOF algorithm performance. Contextual anomaly is divided into three parts as per detection: spatial, sequential and graph.

Paper [10] discussed the two-phase classification method for intrusion detection method, phase one used the k means algorithm for clustering the set of patterns in the second phase the distance-based technique is used. In that phase, the outliers are created, and a label is assigned to each class. The knowledge discovery databases (KDD) cup 1999 dataset was used for intrusion detection. *K* means algorithm minimizing the objective function by the squared error function:

$$J = \sum_{j-1}^{k} \sum_{i-1}^{n} x_i^{(j)} - c_j^2$$

Here $x_i^{(j)} - c_j^2$ is used to measure distance between data points $x_i^{(j)}$ and the corresponding cluster center c_j. This point calculates the distance between the n data points from the respective cluster centers. In phase two, the outliers are the point that is out of the observation in a dataset. To detect the outliers, the point distance-based technique is used. This binary classification is used for classifying the members from the dataset into two different groups whether the members belonged to the same property or not. In this intrusion detection, the two classes are the normal connections and the attack situation.

In the paper [11], authors discussed anomaly threats. Their method detects all malicious activity with less than a 4.6% false positive rate. The real-world data is also in consideration for testing the malicious activities and the result showed that user profile detects 86% of suspicious behavior with less than a 1% false positive rate. Lightweight agents are used to collect the system process information. The obtained logs are converted into transactions by random resampling techniques. The transactions are having a set of active processes. Their prototype system collects data from both virtual machines and real-world data. Cloud service is the measured service which is an essential characteristic for controlling and optimizing resource usage. The measured service is aware of the dashboard or monitoring service in a cloud platform. It monitors and reports the resource usage of the cloud service. Lightweight agents are used to collect the workload of the process and send it to the database. The process is activated for 5 s. With the help of random resampling generating transactions to represent user behavior, each period contains 2521 active process records in a sequence of the data pool. For detection of the outlier factor, two techniques of frequent pattern outlier and longer pattern outlier were used. The open-source library system information gatherer and reporter (SIGAR) helps us to develop the data. In the learning module, the data collected by the data collector is parsed into categorical transactions. As a result, utilizing the process log and resampling frequent pattern build models predict anomalous behavior.

4 Conclusion

This paper overviews different kinds of machine learning techniques used in intrusion detection systems in cloud computing but still there is a need to identify attributes responsible for hacking of data. Related works mentioned many ML techniques but actual implementation of cloud is yet to be explored. So, with the help of machine learning techniques experiments, we can make intrusion detection into the cloud computing area more efficient and achieve cloud security. We suggest ML algorithms for a more expert intrusion detection system, which can make the detection of malicious activities easier and help in the prevention of important data. The future scope includes identifying challenges in the implementation and further modifying the techniques.

References

1. Mell P, Grance T (2009) The NIST definition of cloud computing. National Institute of Standards and Technology (NIST), Information Technology Laboratory. www.csrc.nist.gov
2. Subashini S, Kavitha V (2010) A survey on security issues in service delivery models of cloud computing. J Netw Comput Appl (Elsevier), 11 July 2010
3. Hajdarevic K (2017) Survey on machine learning algorithms as cloud service for CIDPS. In: 25th telecommunications forum TELFOR
4. Bhamare D, Salman T, Samaka M, Erbad A, Jain R (2016) Feasibility of supervised machine learning for cloud security. In: International conference on information science and security (ICISS)
5. Gulenko A, Wallschläger M, Schmidt F, Kao O, Liu F (2016) Evaluating machine learning algorithms for anomaly detection in cloud. In: 2016 IEEE international conference on Big Data (Big Data)
6. Osanaiye O, Cai H, Choo K-KR, Dehghantanha A, Xu Z, Dlodlo1 M (2016) Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing. EURASIP J Wireless Commun Netw
7. Garg A, Maheshwari P (2016) A hybrid intrusion detection system: a review. In: 2016 10th international conference on intelligent systems and control (ISCO), Coimbatore, India
8. Duessel P, Gehl C, Flegel U, Dietrich S, Meier M (2017) Detecting zero-day attacks using context-aware anomaly detection at the application-layer. Springer, Heidelberg
9. Huang T, Zhu Y, Zhang Q, Zhu Y, Wang D, Qiu M, Liu L (2013) An LOF-based adaptive anomaly detection scheme for cloud computing. In: 37th annual computer software and applications conference workshops
10. Songma S, Chimphlee W, Maichalernnukul K, Sanguansat P (2012) Classification via k-means clustering and distance-based outlier detection. In: 2012 tenth international conference on ICT and knowledge engineering
11. Chiu C-Y, Yeh C-T, Lee Y-J (2013) Frequent pattern based user behavior anomaly detection for cloud system. In: 2013 conference on technologies and applications of artificial intelligence

Deep Learning Approach to Predict Crop Yield Using IOT Sensor Data

Vimal Kumar, Pritee Parwekar, and Gunjan Gupta

Abstract Agriculture and farmers are backbone of any developing nation. But, there are certain hindrances that reduce the profit from agriculture. Moreover, with increase in global warming and temperature variations, it is becoming difficult day by day to predict the temperature that can be helpful for farmers. Also, the crop yield today depends mostly on chemicals such as fertilizers and pesticides. Such chemicals affect the pH levels of soil immensely. In such situations, it would be a blessing to have a device to predict the rainfall, temperature, humidity, and analyze the soil's pH levels to conclude which crop would give the best yield on that soil. This is exactly what our proposed system aims to do. We take as an input the soil moisture level of a particular place using sensors, the crop which the person wishes to grow and the month he/she wants to grow it in. As an output, we predict whether the person can grow a particular crop in the entered time frame at the mentioned place or not. This proposed system has a vast variety of use. It can be used by farmers, as well as a layman wishing to start their home garden. If it is made available at a cost-effective price to the many small scale agriculturist, it can prove to be really useful and will help in increasing the crop yield.

Keyword Deep learning · Agriculture · pH · Sensor

V. Kumar
School of Computer Science, University College Dublin, Dublin, Ireland

P. Parwekar (✉)
SRM Institute of Science and Technology, NCR Campus, Modinagar, Ghaziabad, UP, India
e-mail: priteep@srmist.edu.in

School of Life Sciences, JNU, Delhi, India

G. Gupta
A2G Enterprises Pvt. Ltd., Noida, UP, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_5

1 Introduction

Food security is a world challenge and 70% of the world population struggles for nutritious and hygienic food on a daily basis. Successful agriculture and prediction of crop yield therefore becomes important to foresee and address the food security challenge. The right tools and the right models might help the humanity achieve zero hunger, which has been recognized as a top UN Sustainable Development Goals for year 2030. Subsequently, a combination of methodological-structural approach such as variety inception, pesticide and fertilizer management, integrated cropping, rainwater harvesting, efficient irrigation techniques would be required. Moreover, it becomes essential to simulate and predict the crop yield under ambient conditions prior to the cultivation stage in order to make crop management more effective in achieving the desired results. In the recent years, many countries are employing precision farming practices to improve their crop yield.

Since the relationship is non-linear between weather, crop yield parameters, and non-weather factors. There are numerous challenges this proposed work has tried to address. Machine learning has been proven to be successful in predictive analysis and hence has been tried in crop yield predictions. This work aims to predict crop yield using a scalable and inexpensive method using machine learning on publicly available remote sensing data. The proposed method uses a deep learning approach which can help predict crop yield using soil moisture and previous years' data several months in advance. This solution would enable us to make better and informed planting decisions. Moreover, it would also help in setting an appropriate food reserve level, identifying low-yield regions, and improving risk management of crop-related derivatives.

Advances in sensing technologies, MEMS and machine learning are providing cost-effective and scalable solutions for a variety of environmental applications. Tapping them for agricultural decision making is therefore almost an obvious choice. Targeted applications of the sensor platforms and machine learning blending are being made using different sensor modalities and expert knowledge. These would aid in development of hybrid systems combining different ML and signal processing techniques to aid in precision agriculture (PA) in the near future. The need and concept of sensors communicating over Internet is nicely explained in [1] which can be applicable in agriculture. In the recent times, research has shown interest in automatic emotion classification by machine learning and deep learning. This can also be used in human–computer interfaces (HCI) with varieties of applications [2] [my human]. Machine learning models have also been used to predict the coefficient of consolidation [3] [my prediction].

The proposed approach collects information from the user regarding the crop being grown, the planned month for the crop, and the geographical area of the field. Using machine learning, the proposed system will be able to predict whether it would be possible to grow the said crop under mentioned conditions or not. The long short-term memory (LSTM) model has been used to predict the crops yield. LSTM models keeps a memory of the parameters learnt before the current input. The crop yield prediction depends on previous years' data, thereby establishing a dependency with preceding years data. Thus, LSTM network will be more helpful in predicting the crops yield in an accurate manner.

2 Related Work

The research paper [4] emphasizes on predicting the crop yield based on the existing data by using Random Forest algorithm. They focused on Kharif and Rabi crops (Indian summer and winter crops). The parameters taken into consideration were—rainfall, maximum temperature, crop production (tons), and perception. The results reveal that crop yield prediction can be improved by using the Random Forest algorithm. The limitation of this model was that it is only suitable for large data, and big farmers.

The global survey over the years 1961–2008 with over 2.5 million observations is traversed by paper [5].

Their study was on the Kharif and Rabi crops. They examined crop yields for rice, wheat, maize, and soybeans. Although yields continue to increase in many areas, it was found that across 24–39% of maize-, rice-, wheat- and soybean-growing areas, yields either improved very poorly or it collapsed. These statistics make it evident that a considerable investment in agriculture is required to be made in the coming decades to meet the growing demand for food and also maintain the ecological balance. In other words optimization of the resources is the need of the hour. Growth in agriculture, thus has to become an encompassing and sustainable solution.

Another facet of agriculture is weather forecasting. This process in todays times is based on atmospheric parameters such as temperature, rainfall, humidity. These parameters can be predicted using various mathematical models such as regression, fuzzy logic, neural network, and time series analysis. Finite Element Analysis on a global FE model is the state of the art which is used by all the latest meteorological departments across the world. The paper [6] has proposed a weather prediction model for IOT enabled polyhouse. This system has been developed on just two weather parameters namely, temperature and humidity. This can be enhanced by inclusion of other parameters in addition to existing parameters.

In [7] author has applied various data mining techniques for predicting the profit that can be achieved from crops. Later they proposed an approach that uses association rule mining on agricultural data. The association rules were designed for each of the regions. The system was found to perform better. Paper [8] has performed studies and research that are to assess patterns that can be seen in the yield production and to predict data of similar nature to overcome flaws in traditional agriculture such as high labor requirements, poor acquisition of real-time data, small monitoring area, and precision agriculture based on the Wireless Sensor Network. These nodes sense the environmental factors such as temperature, pH, humidity, and NPK. The analysis is done on the observed data that is captured from the field for further analysis and is stored in the server. The significant advantage of the proposed design is in development of an automated detection and alarming system.

3 Proposed Work

The proposed system is developed using IOT devices that can retrieve data from the sensory unit and transmit them to a firebase server. So that, the data retrieved is available for further processing and visualization. The same data can also be used for any other further extension that can be helpful in remote monitoring of the crops [9]. In the sensory unit, a moisture sensor is used for detecting the moisture of soil or judge if there is water around the sensor. The sensor values for different scenario are as shown in the Table 1. The analog output of soil moisture sensor is processed using analog to digital converter. The moisture content in terms of percentage is displayed on the serial monitor. The output of the soil moisture sensor changes in the range of 0–1023. This can be represented as moisture value in terms of percentage. The first step in solving our problem and predicting crop yield was integrating our soil moisture and soil temperature sensors with our ESP8266 NodeMCU Wi-Fi module [10]. This was done so that the sensors, after collecting data could pre-process and then, using the NodeMCU, upload that data to the Firebase server in real time. The data sent from the ESP8266 Wi-Fi module to the Firebase server is of no use to the user if they can't view it. Therefore, it needs to be visualized or processed according to the user requirement. The next step retrieves the data from the Firebase database and visualized on the user interface—Android as well as web (Fig. 1).

The final and most important step of this work is to predict the crop's yield which is performed in the data processing module. For processing of data, we obtained two datasets corresponding to the rainfall as well as temperature statistics in Indian states for the last few decades were collected. Next, this data was categorized into testing and training data. Subsequently, different machine learning models were applied such as regression, long short-term memory (LSTM), and stacked LSTM (Fig. 2).

Every model is used to predict the crop produced in a specific location and at specific time. Among all the models LSTM is the preferred one. Description of LSTM network is given in Fig. 3 which helps to preserve the error that can be back-propagated through time and layers. By maintaining a more constant error, it allows recurrent network to continue to learn over many time steps (over 1000), thereby opening a channel to link causes and effects remotely.

Table 1 Moisture sensor values

Moisture level	Sensor value
High	High
Low	Moderate
Very low	Low

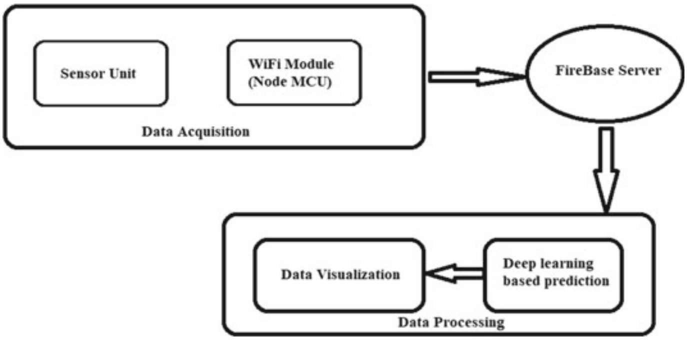


Fig. 1 Proposed crop yield prediction system

Fig. 2 Overall system architecture

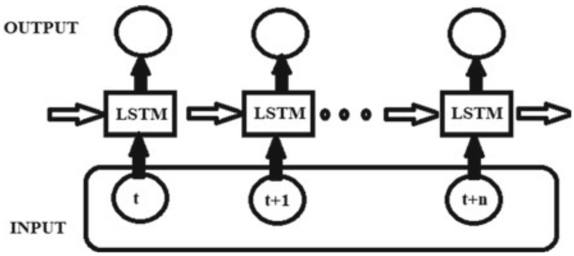


Fig. 3 LSTM network

To improve the accuracy of overall system, a stacked LSTM network-based weather prediction was employed. Based on the analysis of dataset, it is found that weather data is spread over different timestamp. The information at each timestamp has a contribution in predicting the overall weather parameter. Thus, usage of LSTM network will be more beneficial in such a scenario. But, the relation between each data at different timestamp has an influence on the final prediction. This can be captured by making use of stacked LSTM layers instead of a simple LSTM network.

4 Results

The proposed system collects the real-time data using the data acquisition module which has a sensory module and Wi-Fi module. The Wi-Fi module transmits the data to a firebase server so that, it can be used by any user according to their need. The data processing module makes use of various machine learning algorithms such as linear regression, long short-term memory (LSTM), and stacked LSTM networks. The Fig. 4 depicts the comparison of results in each of these machine learning models. It is evident from the results that the stacked LSTM network outperforms over the other two machine learning algorithms (Table 2).

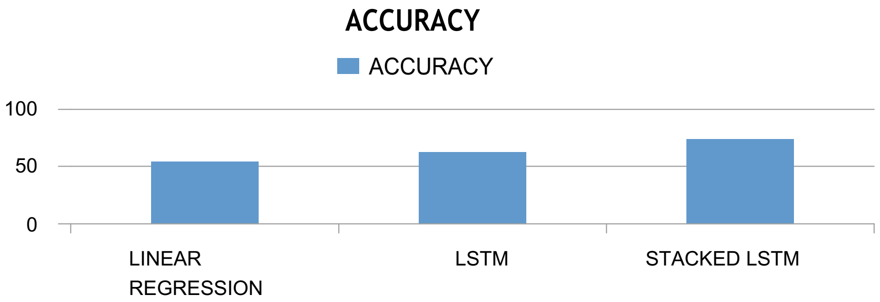


Fig. 4 Result analysis of proposed system

Table 2 Comparison of accuracies of 3 classifiers with dataset of 1300 samples and 2500 sample

Classifiers	Accuracy with dataset = 1300 samples (%)	Accuracy with dataset = 2500 samples (%)
Linear regression	49.58	52.8
LSTM	55.78	64.57
STACKED LSTM	70.59	74.94

5 Conclusion

In this research, the real-time monitoring features of IOT and the powerful deep learning LSTM models are clubbed to achieve the aim of better crop yield for poor farmers in their crop season. The LSTM deep neural network technique was successful in giving around 75% accurate results for rainfall and temperature at different parts of India across all the months. The LSTM network consists of 2 hidden layers that take the input from the underlying input layer, and gives the output to the output layer. This result is later used to draw the comparison between the requirements of the crop which the farmer has chosen. The proposed system has an upper edge of using real-time cloud storage database that can be accessed by anyone from any part of the world with an Internet connection. This facilitates the users with a remote monitoring system so as to make their work easier. The node MCU featuring ESP8266 is used for communicating sensor values over the cloud to the FIREBASE database that is finally retrieved through android applications. Initially the soil moisture and atmospheric temperature are monitored that can be further extended to the essentials of a healthy crop. It is found that the model is successful in achieving its aim using above mentioned techniques and it could be more helpful if developed further.

6 Future Work

In the proposed work, only two of the parameters are considered for the monitoring of crop growth namely soil moisture and atmospheric temperature. There are several other factors that are left and are important as part of crop monitoring systems such as pH level of soil and soil temperature. In future, it could also be used to create a safety system to prevent the crops from heavy rains as well as microbial attacks that are the two main causes of poor crop yield. The work could be extended to verify as well as to make the model learn from the upcoming weather conditions and consider the same for future predictions so that the predictions could be made more accurate and works more efficiently. The relation of crop's yield with all the parameters can be used to extracted and used for more accurate prediction as part of future work in this research.

References

1. Parwekar P (2011, Sep) From internet of things towards cloud of things. In: 2011 2nd international conference on computer and communication technology (ICCCT-2011). IEEE, pp 329–333
2. Khan MA, Kadry S, Parwekar P, Damaševičius R, Mehmood A, Khan JA, Naqvi SR (2021) Human gait analysis for osteoarthritis prediction: a framework of deep learning and kernel extreme learning machine. Complex Intell Syst 1–19
3. Mittal M, Satapathy SC, Pal V, Agarwal B, Goyal LM, Parwekar P (2021) Prediction of coefficient of consolidation in soil using machine learning techniques. Microprocess Microsyst 82:103830
4. Priya P, Muthaiah U, Balamurugan M (2018) Predicting yield of the crop using machine learning algorithm. Int J Eng Sci Res Technol 7(4):1–7
5. Ray DK, Ramankutty N, Mueller ND, West PC, Foley JA (2012) Recent patterns of crop yield growth and stagnation. Nat Commun 3:1293
6. Surendra Prasad Babu M, Li WZ, Sneha Deepika M, Sravani B (2018) Polyhouse predictions on meteorological parameters using time series analysis and precision agriculture through CLOUD & FOG interfaces. Int J Innov Res Sci Eng Technol 7(9):10139–10143
7. Manjula E, Djodiltachoumy S (2017) A model for prediction of crop yield. Int J Comput Intell Inf 6(4):298–305
8. Meeradevi MRM (2019) Automated control system for crop yield prediction using machine learning approach. Int J Appl Eng Res 14(2):480–484
9. Importance of crops in India, 5 April 2018. [Online] Available: https://www.mbarendezvous.com/essay/importance-of-crops-in-india. Accessed: 13 Dec 2018
10. The future of predicting yield production, 3 May 2018. [Online] Available: https://www.agriculture.com/technology/crop-management/the-future-of-predicting-yield-production. Accessed 21 Dec 2018

Classification of Pomegranate Fruit Disease Using CNN

Apoorva Thorat, Apurva Kulkarni, Damayanti Chavan, Manasi Joshi, and Ashwini Jarali

Abstract Agriculture can have a great impact of change in the climate, use of extensive pesticides, and number of other factors. Various traditional methods are used to examine and detect the plant diseases, and these are performed by the domain experts. The identifying and detecting task can be costly, and it requires more time and effort. For this, some solution can be developed to identify and detect the diseases in crops and plants. Deep learning and image processing techniques can be used to develop a system for identifying and detecting the crop diseases which can be much faster and cost-effective solution. This paper proposes a pomegranate fruit disease classification model using convolutional neural networks and transfer learning methods, which can be useful for detecting and identifying crop and fruit diseases. The research in pomegranate fruit disease is not that explored. For implementing this system, the dataset is created by capturing the images directly from the farms in Solapur. The proposed system uses TensorFlow as its framework for development of the deep learning model. VGG-16 which is a popular transfer learning algorithm is used for extracting the features from the images, and fully connected layer is added externally which is used for classifying the image into respective classes. We have created a web portal, where the end user can upload the image of pomegranate fruit and the system can accurately predict the disease of that fruit.

Keywords CNN · Deep learning · Transfer learning · Image processing · VGG-16 · Pomegranate fruit diseases

A. Thorat (✉) · A. Kulkarni · D. Chavan · M. Joshi · A. Jarali
Department of Computer Engineering, International Institute of Information Technology (IIT), Pune, India
e-mail: apoorvabthorat@gmail.com

A. Jarali
e-mail: ashwinij@isquareit.edu.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_6

1 Introduction

Plant disease is abnormal state in the plant that has overall negative effect on structure and functioning of all parts of the plant. The symptoms are used to classify the diseases into its respective classes and predict the disease that the plant has. The crop diseases can have great impact on production in agricultural sector and can incur significant economic loss. Diagnosis of plant disease through visual examination is very complex and highly time consuming. For best fruit formation, supplemental irrigation is used, as humidity and overly wet soils can cause a variety of pomegranate fruit diseases. Overall, the management and prevention of crop diseases and pests are critical for the success of the agricultural sector. Early detection and appropriate treatment are essential for minimizing economic losses and ensuring optimal plant health. Pomegranate fruit diseases:

1. **Anthracnose**. The most common symptoms of Anthracnose are spots, blotches, defoliation, shoot blight, twig cankers, and dieback [1]. Fruit symptoms consist of sunken lesions in which the pathogen produces massive amounts of spores that may be visible to the naked eye [1]. Anthracnose is a group of diseases that affect a variety of crops in hot and humid regions [2].
2. **Bacterial Blight**. Bacterial blight disease reduces the production of fruit by 30–50%. Pomegranate disease is triggered by the fungal pathogen Xanthomonas axonopodis. The pathogens enter the plant's fruit through wounds and stomata, affecting the plants. Relative humidity of > 50% and temperature range between 25 and 35 °C favor the disease development [3] (Figs. 1, 2 and 3).
3. **Cercospora**. The plants affected by cercospora show small, dark brown spots on fruits that are initially circular, but they eventually become irregular as they grow [3]. On the leaves, the lesions are dark, reddish brown to almost black and show a faint halo [3]. The spots on fruit resemble bacterial spots, but they are darker, of various sizes, without cracks, and no stickiness [3].

The research on pomegranate leaf diseases is not that explored. We have captured the images from the farms, and a database is created. We have used TensorFlow as our framework for creating the CNN model. For classification, we have use convolutional

Fig. 1 Anthracnose

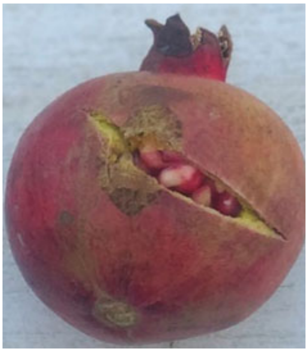


Fig. 2 Bacterial blight

Fig. 3 Cercospora

neural network (CNN) algorithms and compare accuracies of both. We created a web portal, where the end user can upload the leaf image and the system would accurately predict the disease of that leaf.

2 Literature Survey

The authors Dhakate and Ingole [4] have used artificial neural networks for classifying the pomegranate leaf disease. They have created the database by capturing images directly from the farm. They used *K*-means clustering for creating different clusters of different diseases. They have created a multilayer perceptron for building the neural network. They have got 90% accuracy through this model for classification.

The authors Bhange and Hingoliwala [5] have developed a web-based portal using which farmers can upload fruit images and get the disease in real time. The clusters were formed by using *K*-means clustering method, and support vector machine was used for classification of diseases. They extracted three features out of which the morphology proved to be the best one scoring accuracy of 82%.

The authors Deshpande et al. [5] have worked on both leaf and fruit disease identification of pomegranate. It specifically focuses on bacterial blight disease on pomegranate leaf and fruit. They identified the disease of the leaf or the fruit by the yellowish border over it. For segmenting and extracting this feature, K-means clustering algorithm was used.

Kantale and Thakare [6] have presented a review on classifying pomegranate fruit disease using machine learning techniques as well as image processing techniques. They used KNN for feature extraction with 87.02% accuracy. They have also used SVM for classification.

Shaikh and Manjramkar [7] worked on apple fruit and leaf disease identification. The three classes that they defined for fruit images were blotch, rot, and scab. The dataset they used was of size 280 in total, and there were 70 images in each class. They used SVM for multiclass classification of the fruits. They got accuracy of 96%.

Sharath et al. [8] worked on various image processing techniques for detecting the diseases in pomegranate fruit. They considered color of the fruit and edge details for classification basis. They have developed a system that provides the percentage of the fruit infected and also suggests the measures of precautions. The system gives the real-time results where the images are captured through phone camera which undergoes image processing steps and finally is classified. They have used OpenCV and Python language for developing the algorithm.

Sharath et al. [9] have developed a system which detects the bacterial blight using image processing methods. They have used Gaussian filters in order to detect the edges. For segmentation, they used grab cut segmentation method which deals with 2D segmentation.

Mangena Venu Madhavan, Dang Ngoc Hoang Thanh, Aditya Khamparia, Sagar Pande, Rahul Malik, and Deepak Gupta [10] have worked on developing a framework for detecting and classifying diseased pomegranate plants accurately. They have used MATLAB and used SVM for classification purpose. The disease identification or recognition is done using K-means clustering algorithm. The accuracy they achieved over classifying pomegranate leaf disease is 98.39%.

Banupriya, Rajaneni Deepila sai chowdry, Yogeshwari, and Varsha [11] have studied and worked on various image processing and machine learning techniques used for plant disease detection. K-means clustering was used for clusters of images that shared similar properties. Convolutional neural network was used for learning the features in images and predicting the results. They achieved the accuracy of 89%. They have considered potato and pomegranate leaves for their study.

Akshai and Anitha [12] have proposed a system that identifies the plant disease and classifies it correctly into its respective classes. They have collected the dataset from plant village. The dataset includes 14 different species of plants and 38 classes out of which 12 are healthy and 26 are diseased leaf. They have trained VGG, ResNet, and DenseNet model over the dataset, and DenseNet has achieved the highest accuracy of 98.31%.

Li et al. [1] have proposed a review on the progress made in the sector of plant disease detection. They have explained traditional image processing techniques along with various advancements in this sector and applied all these techniques on leaf disease detection.

Saleem et al. [2] have proposed a review on plant disease detection and explained various DL models used to visualize various plant diseases. They have compared various data visualization techniques and results obtained through various transfer learning algorithms and also with other research papers.

Venkataramanan and Agarwal [3] have worked on plant disease detection and examined the leaf at various stages. A YOLOv3 is used as an object detector for extracting a leaf image from the input image. The ResNet18 model is used to extract features of the leaf image. This also identifies the disease of the leaf.

3 Algorithms

3.1 Convolutional Neural Network

CNN is a type of deep neural network that can efficiently learn the features from the given data, either image or textual. They consist of an input layer which helps in taking input of the data. The next layers are convolutional layers which are responsible for extracting the features. The next layer is fully connected layer, which is used for classification purpose. For this layer, widely used activation functions are sigmoid (for binary classification) or softmax (for multiclass classification) (Fig. 4).

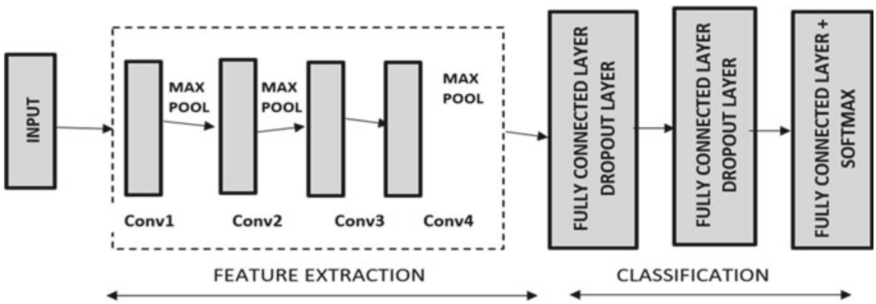


Fig. 4 CNN architecture

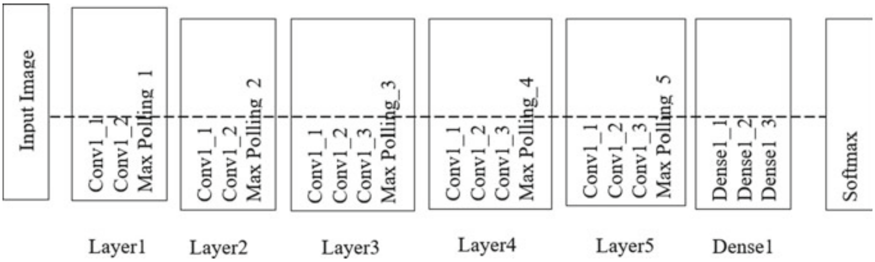


Fig. 5 VGG-16 architecture

3.2 *VGG-16 Model*

VGG-16 is the one easiest to understand and simple to build models. Key features of VGG-16 are

- It is also called the Oxford Net model, named after the Visual Geometry Group from Oxford.
- Number 16 refers to a total of 16 layers that have some weights.
- It only has Conv and pooling layers in it.
- It always uses a 3×3 Kernel for convolution.
- 2×2 size of the max pool.
- It has a total of about 138 million parameters.
- It is trained on ImageNet data (Fig. 5).

4 Methodology

(See Fig. 6).

4.1 *Preprocessing*

During the preprocessing phase, several operations were performed to optimize the image dataset for training the model. First images were resized. Another important step was adjusting the contrast of the images to ensure that the minute details were clearly visible and distinguishable. This helped to improve the accuracy of the model and enable it to detect even subtle signs of disease.

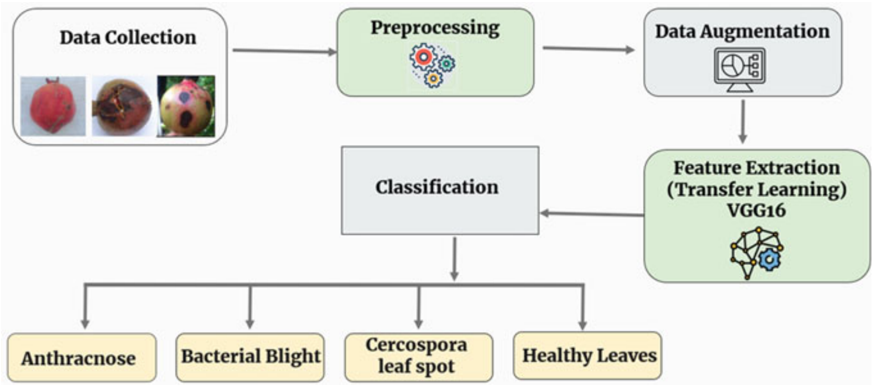


Fig. 6 Architecture diagram of system

4.2 Data Augmentation

In this proposed system, we used data augmentation techniques to increase the size of our dataset and improve the accuracy of our classification model. Overall, data augmentation is a critical component of our system, which allowed to create a more robust and reliable model for disease detection.

4.3 Feature Extraction

In this proposed system, we will be using VGG-16, a transfer learning algorithm to build the model for disease detection. The convolutional layers of this model are specifically designed to extract features from images.

4.4 Classification

For classification of leaf diseases, we had decided to use convolutional neural network (CNN). In the CNN approach, we used fully connected layers for classification. This technique had proven to be highly effective in image recognition tasks.

4.5 Application

For the ease of use of the system, we have developed a web application using Flask framework. The application includes a homepage which asks for input of the image.

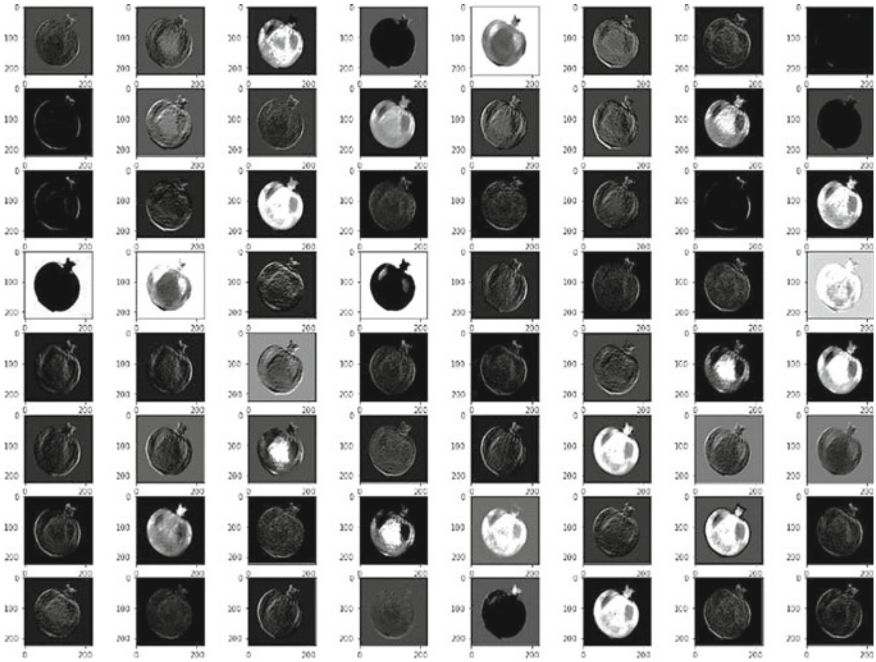


Fig. 7 Output after first layer of CNN

After uploading the image, the system undergoes image processing steps at the backend, predicts the class of the given input image, and displays on the web page.

5 Results

For the experimental work, the dataset was created by clicking images directly from the farm. We also took the help of domain experts from pomegranate research institute of Solapur for classifying the images with naked eyes. Their guidance helped a lot in developing out project, specifically out dataset. The images after applying various filters in VGG-16 layers are as follows (Fig. 7).

As we mentioned, we used total across 700 images for training the CNN model.

The training and testing accuracies of our models are given below (Figs. 8 and 9, Table 1).

The output after uploading the image of pomegranate is given in Fig. 10.

```
print("Training Accuracy: ", history.history['accuracy'][-1]*100, "%")
Training Accuracy:  99.67741966247559 %

print("Testing Accuracy: ", history.history['val_accuracy'][-1]*100, "%")
Testing Accuracy:  96.59090638160706 %
```

Fig. 8 Training and testing accuracies

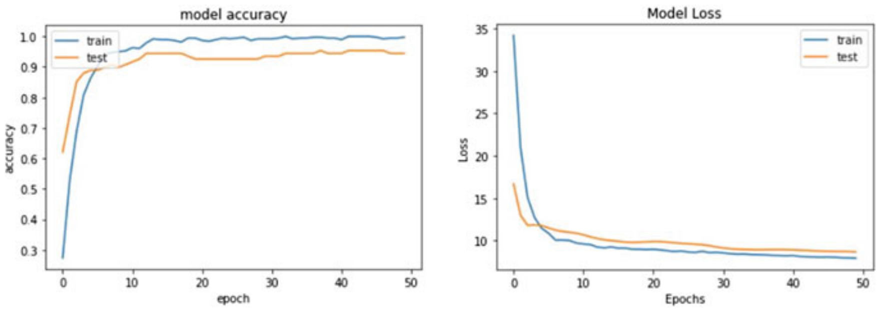


Fig. 9 Model accuracy versus epoch and model loss versus epoch

Table 1 Comparison of existing pomegranate diseased fruit detection system

Sr. No	Title of paper	Author Name	Accuracy achieved (%)	Technology used
1	Diagnosis of pomegranate plant diseases using neural network [4]	Dhakate and Ingole	90	They have used *K*-means clustering for extracting features and ANN for classification
2	Smart farming: pomegranate disease detection using image processing [5]	Bhange and Hingoliwala	82	They used image segmentation for extracting features and SVM for classification
3	Disease detection in pomegranate using image processing [9]	Akhilesh et al.	80	For the detection of plant disease, they used grab cut segmentation
4	A review on pomegranate disease classification using machine learning and image segmentation techniques [6]	Kantale and Thakare	89.23	They used KNN for feature extraction and also used SVM

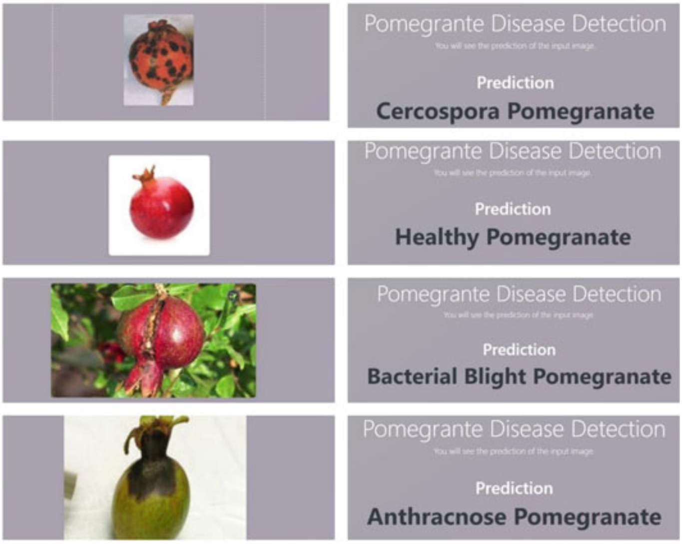


Fig. 10 Output displayed in web app

6 Conclusion

Historically, the identification of fruit diseases has been dependent on the visual inspections carried out by professionals and experts in the field. However, accessing expert consultations at remote locations can be difficult and costly. Furthermore, there has been limited research done on pomegranate fruit disease prediction. To address these issues, we have developed a system that can assist farmers, experts, students, and others in detecting diseases early on. The system allows users to upload images of pomegranate fruits, which are then processed and their features extracted. These features are compared to a training dataset, and the system makes a prediction based on the results.

Moving forward, we aim to focus on preventing diseases by providing users with suggestions on how to avoid them. Ultimately, our goal is to help farmers and other stakeholders make informed decisions that can improve crop health and yield.

References

1. IFAS Extension Homepage. https://edis.ifas.ufl.edu/
2. Britannica Homepage. www.britannica.com
3. Bighaat Homepage. https://www.bighaat.com/
4. Dhakate M, Ingole AB (2015) Diagnosis of pomegranate plant diseases using neural network. In: 2015 fifth national conference on computer vision, pattern recognition, image processing and graphics (NCVPRIPG). IEEE, Patna, India, pp 1–4

5. Deshpande T, Sengupta S, Raghuvanshi KS (2014) Grading and identification of disease in pomegranate leaf and fruit. Int J Comput Sci Inf Technol 5:4638–4645
6. Kantale P, Thakare SS (2020) A review on pomegranate disease classification using machine learning and image segmentation techniques. In: 4th International conference on intelligent computing and control systems (ICICCS), pp 455–460
7. Shaikh R, Manjramkar A (2019) Fruit disease classification and identification using image processing. In: Third international conference on computing methodologies and communication. IEEE, pp 754–758
8. Sharath DM et al (2019) Image based plant disease detection in pomegranate plant for bacterial blight. In: 2019 international conference on communication and signal processing (ICCSP), pp 0645–0649
9. Sharath DM, Akhilesh RMG, Kumar SA, Prathap C (2020) Disease detection in pomegranate using image processing. In: 2020 4th international conference on trends in electronics and informatics (ICOEI) (48184), Tirunelveli, India, pp 994–999
10. Akshai KP, Anitha J (2021) Plant disease classification using deep learning. In: 2021 3rd international conference on signal processing and communication (ICPSC). Coimbatore, India, pp 407–411
11. Li L et al (2021) Plant disease detection and classification by deep learning—a review. IEEE Access 9:56683–56698
12. Saleem MH et al (2019) Plant disease detection and classification by deep learning. Plants 8
13. Bhange M, Hingoliwala HA (2015) Smart farming: pomegranate disease detection using image processing. Proc Comput Sci 58:280–288

NICE: Navagraha Iconography Classification Engine

Aayush Mehta, C. Sindhu, Saksham Dube, Meghana Kanneganti, and T. Kavya taneesha

Abstract India is a country with a very rich cultural legacy. Its religion plays a significant role in shaping its customs and traditions. Navagraha gods are a group of nine celestial bodies that are revered in Hinduism. They are believed to have a significant impact on human destiny and are worshiped to seek their blessings and mitigate their negative effects. Throughout recent years, numerous architectural sculptures, figurines, statues, and inscriptions of these Navagraha gods have been unearthed and discovered in India. Recognizing entities in these sculptures presents a highly challenging and complex task in the domain of image recognition and classification. Our idea is based on a manually compiled database of images of the nine Navagraha gods in separate folders that are to be classified. Each image has a unique feature that sets it apart from others, and its orientation, angle, size, and color are crucial in the model's processing. The images were captured from different angles, sizes, colors, and orientations. The model will be trained on the image dataset. Deep learning plays a very important role in image classification and recognition. The effectiveness and accuracy of the suggested model can be evaluated using Resnet50 which is a convolutional neural network architecture. Additionally, many other important libraries such as NumPy, Matplotlib, Keras have also been used in this process of image recognition and classification. The model which this project has helped trained has propelled an accuracy of 100% on similar images with the help of Resnet50 model.

Keywords Navagraha · Iconography · Resnet50 · Multiclass · Deity · Cultural recognition · Classification

A. Mehta · C. Sindhu (✉) · S. Dube · M. Kanneganti · T. K. taneesha
Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India
e-mail: sindhucmaa@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_7

1 Introduction

The digital age has revolutionized the way we store and access information, making it more efficient and less cumbersome. Visual mediums, such as images, videos, and gifs, play a critical role in assimilating information and have led to an increase in research projects aimed at extracting insights from data. This includes projects like sentiment analysis, car model detection, human detection, and Navagraha Iconography Classification research. With technological advancements, we can leverage vision to our advantage and continue to advance our understanding of the world around us.

Hinduism is a complex and ancient religion with diverse beliefs and practices. The Navagraha gods are a group of nine celestial bodies of paramount importance in Hinduism. Each planet is believed to have a profound impact on human fate and is worshiped to seek blessings and mitigate adverse effects. There are numerous sculptures, statues, and small idols related to Navagraha gods, each with its unique form and characteristics. Lord Surya represents vitality, courage, and authority, often depicted as a man with three eyes, holding lotus flowers and riding a chariot drawn by seven horses. Lord Chandra, representing emotions, nurturing, and intuition, is often depicted as a young man with a white complexion, holding a lotus flower, and riding a chariot drawn by ten white horses. These Navagrahas are depicted in stone sculptures, bronze statues, or small idols made of various materials like wood, clay, or brass. The variations in these sculptures are nearly impossible to detect with the naked eye, and every image has a key characteristic that differentiates it from the other.

The use of machine learning technology is crucial in processing images of artifacts such as sculptures, but several factors such as orientation, angles, size, and color make the task complex. The study focuses on developing a highly efficient and sophisticated machine learning model based on the Resnet50 architecture, which can detect and categorize Navagraha gods in images. The model aims to achieve optimal precision in identifying the depicted entity [1] within the sculpture image. The objective is to recognize the Navagraha god based on its unique characteristics, and the study aims to draw comparisons and ultimately institute an efficient model. The use of deep learning and convolutional neural networks such as CNN [2, 3] is fundamental to achieve the project's goals.

The research aims to develop a machine learning-based classification model that can accurately identify all nine Navagraha gods Fig. 1, which includes Surya, Brihaspati, Shukra, Mangala, Budha, Shani, Chandra, Rahu, and Ketu. The dataset used in the study comprises images with different angles and orientations, and it is available in formats such as Jpg, Png, or Jpeg. The data is divided into nine distinct folders for training and testing, respectively, to ensure the highest level of precision output. However, collecting the dataset is a challenging task, and data augmentation is necessary to expand the available dataset. The machine learning model is trained using the training data and tested with the remaining data to evaluate the model's efficiency and accuracy. The research paper is structured into literature survey, methodology, implementation, and results sections. The model has potential applications in various fields, such as archaeology, astrology, art, and tourism. For instance, it can save significant time and effort in identifying and categorizing Navagraha gods in Indian sculptures and artifacts. In astrology, the model can help individuals understand the astrological influences associated with each god. In art, it can aid in better understanding Indian art and mythology. Moreover, it can be used to educate people about Indian culture and mythology, promoting greater understanding and appreciation. The study contributes to ongoing research in computer vision, image recognition, and artificial intelligence. To ensure accuracy, the dataset was thoroughly curated and segregated into folders for easy access. Data augmentation was used to expand the dataset further. During the testing process, a 10% split of the data was used, with the remaining reserved for training purposes. This ensures that the model is well-trained and can make accurate predictions.

In conclusion, the research study developed a well-curated database of images for accurate classification using machine learning. The model can be used in various fields, such as archaeology, astrology, art, and tourism, and can help individuals better understand Indian culture and mythology. The study contributes to ongoing research in computer vision, image recognition, and artificial intelligence (Fig. 1).

2 Literature Survey

The table provided in the research paper compares various image recognition/ detection works and the most used architectures include CNN, Euclidean distance method, and Eigenfaces method. CNNs are deep learning algorithms that analyze images and differentiate between them by utilizing biases and weights. The network consists of a convolutional layer to extract features and a pooling layer to select data with the highest value inside an area.

Convolutional neural networks (CNNs) excel in categorizing and identifying images, and reduce computation in comparison to regular neural networks. CNNs use raw data for learning and recognition without any processing and do not require an understanding of face [4, 5] geometry or reflection. However, CNNs require a substantial amount of training data, and do not encrypt data. On the other hand, the Eigen Face algorithm effectively extracts the pertinent data from an image and accommodates variations by using a collection of images for this purpose. The Eigen Face algorithm is sensitive to scale and requires low-level pre-processing, but is insensitive to variations in lighting and position.

The "Thai Buddhist Sculpture Recognition System" was a model for sculpture detection developed by a team at IACSIT. This system's primary objective was to identify Buddhist sculptures found in Thailand. Acquisition of image, pre-processing of images, extraction of features, recognition of images (implementing the Euclidean Distance technique), and output display through a graphic user interface make up the five components of this system.

Fig. 1 Navagraha Gods

A novel method for identifying deities' faces has been proposed in the "Deity Face Recognition using Schur Decomposition and Hausdorff Distance Measure" paper [6]. The paper utilizes a combination of Schur decomposition and principal component analysis (PCA), along with a doubly modified Hausdorff distance calculation and a *k*-nearest neighbor classifier for classification purposes. The output of the experiments showed that the prospective method performed is superior than the other methods.

The cifar10 dataset is used in the "Image recognition using convolutional neural network combined with ensemble learning algorithm" research. The paper describes a simulation of a model that incorporates both an integrated learning algorithm and a

Table 1 Literature survey table

S. No	Project title	Architecture used	Accuracy (in %)	Merits
[7]	TBuSRS	Euclidean distance method	90.00	High accuracy
[6]	Deity face recognition using schur decomposition and hausdorff distance measure	PCA, M2HD	–	Less training time, more training images
[8]	Image recognition using convolutional neural network combined with ensemble learning algorithm	ELA-CNN	98.89	Very high accuracy, Enhanced accuracy, and dependability in identifying the intended target
[9]	State of the art	Adaboost, Haar, LB P, SVM, HOG	92.35	Overcom es datasets complexity
[10]	A face recognition approach using Zernike moments for video surveillance	Eigenfaces, Fisherface, Zernike Moments	88.33	Better accuracy

convolutional neural network. The results of the study indicate that augmenting the number of parameters in a convolutional neural network that relies on an integrated learning algorithm can enhance the accuracy and dependability of target recognition (Table 1).

3 Methodology

In the methodology section, we put forth a project flow diagram and a flow chart that will brief us through the various steps of our process. As you can see below in the Fig. 2, the overall procedure can be summarized in four steps—which is shown in blocks.

The project flow diagram in Fig. 3 provides an overview of the major stages of basic image processing, and it highlights four primary steps that have been indicated. These include gathering the images, pre-processing them, using a model for classification and recognition, and presenting the results and evaluations to represent statistical data related to the project. By breaking down the process into these four main steps, we can better understand the procedures involved in our project and ensure that we are able to carry out each step effectively.

The block diagram in Fig. 4 illustrates the major steps involved in the image classification process for Navagraha gods. It consists of four main components: Image Acquisition, Image Pre-Processing, Image Recognition and Accuracy, and Classification Analysis. Resnet50, a popular pre-trained deep convolutional neural

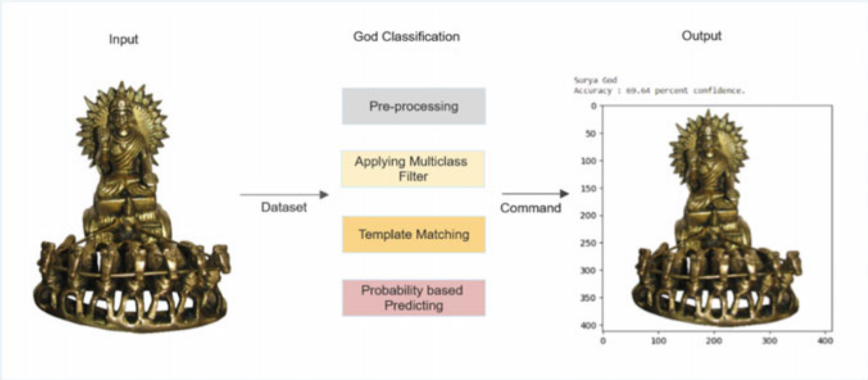


Fig. 2 Project flow diagram

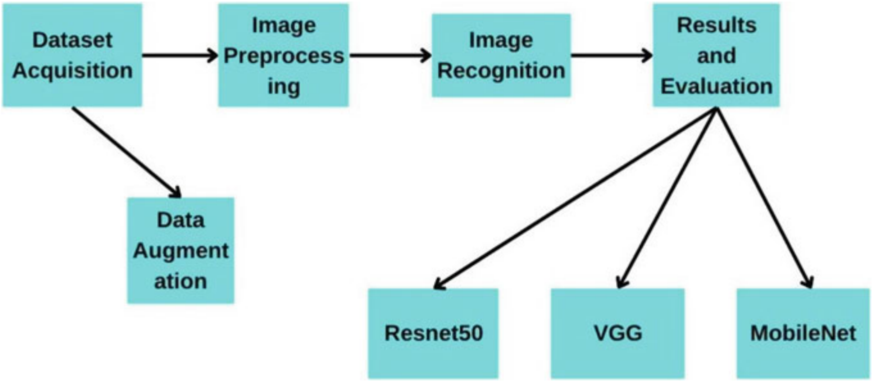


Fig. 3 Architectural block diagram

network architecture, is used for image classification. The methodology involves several steps, including loading the Resnet50 model, preparing the data, training the model, and evaluating the accuracy of the model.

3.1 Dataset Acquisition

To train the image recognition model for Navagraha gods, the first step is to collect a diverse dataset of images of the nine gods, including different poses, lighting conditions, and orientations. Due to the lack of available data, data augmentation techniques such as flipping, rotation, and scaling are employed to artificially increase the size of the dataset. The images can be obtained from various sources, such as online repositories or captured using a camera. About 10% of the images are used

Fig. 4 Data augmentation—RGB to grayscale

for testing, and the remaining are used to train the model. Data augmentation aims to capture variations that exist between different images in the dataset and improve the performance of the classification model.

3.2 Image Pre-processing

Data pre-processing is a crucial step in Navagraha gods image classification, and it involves resizing images to a uniform size and augmenting them to create new images to reduce overfitting. Additionally, splitting the dataset into training, validation, and testing subsets is necessary to instruct the model, fine-tune hyperparameters, and assess its efficacy on unseen data. The preferred size for Resnet50 images is 224 $\times$ 224. Techniques such as flipping, rotation, and changing brightness can be applied to the dataset for image augmentation.

3.3 Classification of Images

During the recognition phase, various architectures such as Resnet50, VGG, MobileNet can be used for building, recognition, and testing. The work discusses the use of deep learning with sequential models or CNN for sorting photos into their appropriate categories. There are nine types of representations to be categorized and identified: Ketu, Rahu, Saturn (Shani), Venus (Shukra), Jupiter (Brihaspati), Mercury (Budha), Mars (Mangala), Moon (Chandra), and Sun (Surya).

4 Analysis

The accuracy and classification analysis phase involves evaluating the accuracy of the model's prediction, determining the loss or error, and plotting graphs using libraries like matplotlib and NumPy. The results are analyzed, and the output is in the form of an image with the name of the god and a percentage indicating the confidence level of the model's prediction. The results of this phase are shown in Figs. 5 and 6, which provide an overview of the accuracy and loss of the training and testing data, as well as the classification results for the Navagraha gods dataset.

```
Epoch :  1
/tmp/ipykernel_9188/2922417473.py:3: UserWarning: `Model.fit_generator` is deprecated and will be removed in a future version.
Please use `Model.fit`, which supports generators.
  loss_history += [my_new_model.fit_generator(train_generator, steps_per_epoch=4,
4/4 [==============================] - 1s 346ms/step - loss: 0.2420 - accuracy: 0.8750 - val_loss: 0.3101 - val_accuracy: 0.750
0
Epoch :  2
4/4 [==============================] - 1s 384ms/step - loss: 0.2802 - accuracy: 0.8750 - val_loss: 0.5679 - val_accuracy: 0.750
0
Epoch :  3
4/4 [==============================] - 1s 333ms/step - loss: 0.0698 - accuracy: 0.9286 - val_loss: 0.3429 - val_accuracy: 0.750
0
Epoch :  4
4/4 [==============================] - 1s 355ms/step - loss: 0.0470 - accuracy: 1.0000 - val_loss: 1.2316 - val_accuracy: 0.750
0
Epoch :  5
4/4 [==============================] - 1s 330ms/step - loss: 0.0676 - accuracy: 0.9375 - val_loss: 1.2086 - val_accuracy: 0.500
0
```

Fig. 5 Accuracy and loss values for Resnet50

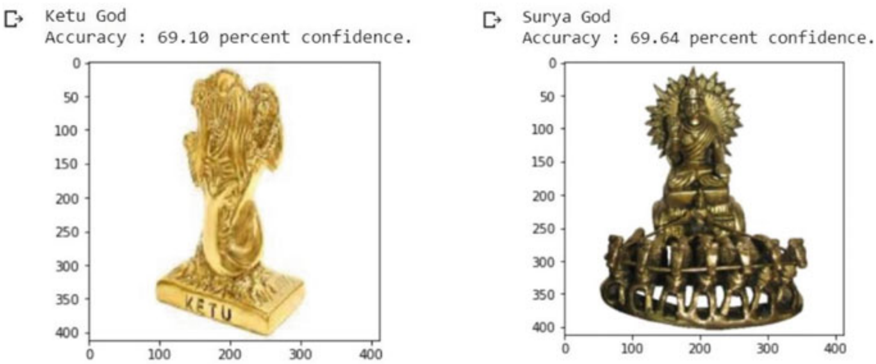


Fig. 6 Sample outputs for two input images

5 Implementation

This section of the research paper involves utilizing deep learning techniques, specifically for image processing and classification. The Resnet50 architecture was selected as the primary architecture for the convolutional neural network (CNN), as it was found to be the most accurate in achieving the desired outcomes. The Resnet50 architecture is known for its ability to facilitate the training of deeper neural networks, which is especially useful in the case of image processing. Aside from Resnet50, the study also explored other CNN architectures such as VGG and MobileNet. However, they did not perform as well as Resnet50, resulting in a lower accuracy rate. This was likely due to differences in the architectures and the way that they process information.

5.1 Libraries Used

To train a model for image classification using Resnet50 architecture for Navagraha gods dataset, multiple machine learning libraries are used, including NumPy, Pandas, Matplotlib, OS, Keras, and TensorFlow. NumPy is used for performing complex numerical computations, Matplotlib for data visualization, Keras for constructing and instructing complex deep learning models, and TensorFlow for performing dataflow and differentiable programming with a focus on scalability and ease of use. These libraries provide pre-built models and layers, as well as options for customization and fine-tuning models for specific applications, facilitating the process of image classification.

5.2 Resnet50

Resnet50 is a popular convolutional neural network (CNN) architecture used for image classification. It consists of 50 layers divided into several parts, each with a specific function. The input layer prepares the input image for processing, while the convolutional layers extract key features using compact filters to identify patterns in the image. Max pooling layers downsample the feature maps, capturing only the most important features. Residual blocks, a unique feature of Resnet50, consist of multiple convolutional layers and shortcut connections, allowing gradients to propagate efficiently during training, making it easier to train very deep networks. The fully connected layers take the extracted features and perform the final image classification. Softmax is used in the output layer to produce a probability distribution over classes, while ReLU is used in hidden layers to introduce non-linearity and reduce the likelihood of vanishing gradients, allowing the network to model complex functions.

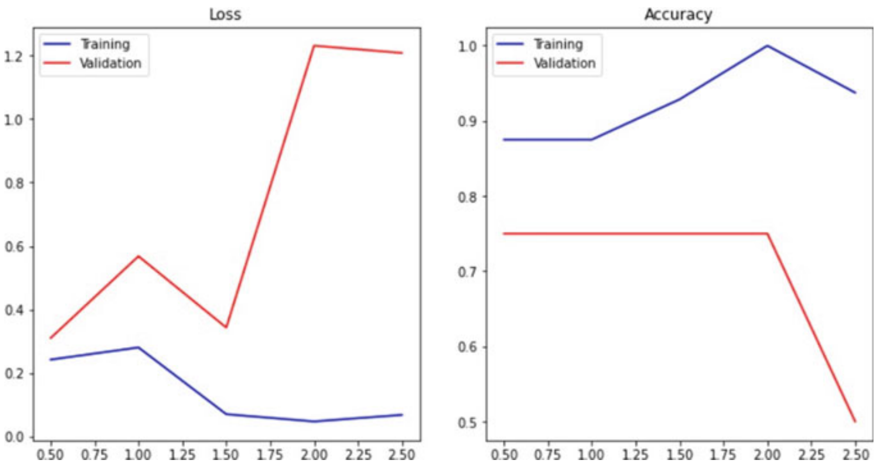


Fig. 7 Loss—accuracy graphs for Resnet50

5.3 *Other Architectures*

VGG16 is a convolutional neural network with 16 layers that extract features and generate class probabilities for image classification. It achieves this by using small convolutional filters with a stride of 1 and padding of 1 to capture detailed information. Additionally, some convolutional layers use up to 512 filters, enhancing the model's representation capabilities. In contrast, MobileNet is a lightweight and efficient convolutional neural network designed using depth-wise separable convolutions and linear bottlenecks, which helps to reduce computational complexity and model size while increasing efficiency. These methods filter input channels individually, followed by point-wise convolutions to combine output feature maps with a 1 × 1 filter (Table 2).

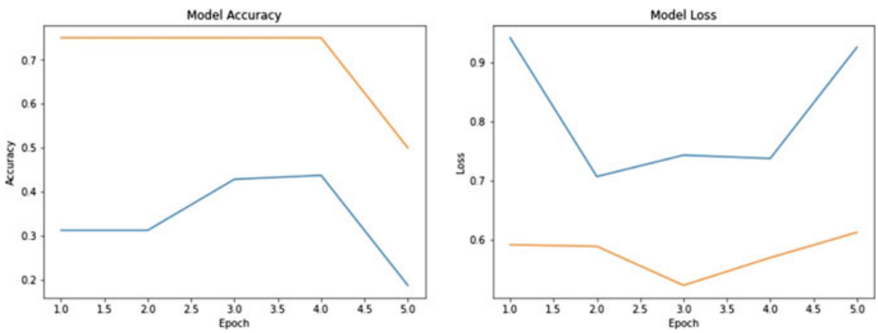


Fig. 8 Loss—accuracy graphs for VGG16

Table 2 Architecture comparison table

Sr. No	Model/architecture	Accuracy	Loss
1	Resnet50	0.9375	0.0676
2	MobileNet	0.3125	1.1120
3	VGG16	0.1875	0.9253

6 Conclusion

This project sought to create a system for detecting Navagraha gods and evaluate the efficacy of various image classification techniques to determine the best method. To assess the success of the project, we measured the accuracy of several image detection methods. The results showed an accuracy of 0.9375 and a loss of 0.676 for Resnet50, 0.03125 and a loss of 1.1120 for MobileNet, and 0.1875 and a loss of 0.9253 for VGG16. Based on these findings, we identified Resnet 50 as the most suitable model for our project. Our research indicates that Resnet 50 is an effective and reliable method for detecting deities and sculptures [11] with exceptional accuracy. In conclusion, this project successfully developed a Navagraha god detection system and identified the most effective image classification technique. By utilizing Resnet50, we achieved impressive accuracy in detecting deities and sculptures. Overall, this research provides valuable insights into image classification methods for recognizing deities and sculptures, and the findings can inform future studies in this field. The results of this study can have significant implications for a wide range of fields, including archaeology, art history, and cultural preservation. With the ability to accurately identify and categorize Navagraha gods, researchers can better understand the cultural and historical significance of Indian sculptures and artifacts. Additionally, the system can contribute to ongoing efforts to digitize and preserve cultural heritage.

References

1. Dalara A, Sindhu C, Vasanth R (2022) Entity recognition in Indian sculpture using CLAHE and machine learning. In: 2022 First international conference on electrical, electronics, information and communication technologies (ICEEICT), pp 1–12. https://doi.org/10.1109/ICEEICT53079.2022.9768565
2. Kaur G, Sinha R, Tiwari PK, Yadav SK, Pandey P, Raj R, Vashisth A, Rakhra M (2022) Face mask recognition system using CNN model. NeuroSci Inf
3. Saini A, Gupta T, Kumar R, Gupta AK, Panwar M, Mittal A (2017) Image based indian monument recognition using convoluted neural networks. In: 2017 international conference on BigData, IoT and data science (BID), pp 138–142. https://doi.org/10.1109/BID.2017.8336587
4. Almadan A, Rattani A (2021) Towards on-device face recognition in body-worn cameras. IEEE IWBF
5. Cheng EJ, Chou KP, Rajora S, Jin BH, Tanveer d M, Lin CT, Young KY, Lin WC, Prasad M (July 2019) Deep sparse representation classifier for facial recognition and detection system. Patt Recog Lett
6. Balakrishnan R, Kannan R, Kannaiyan S, Swaminathan S (2013) Deity face recognition using schur decomposition and hausdorff distance measure. IEEE

7. Pornpanomchai C, Arpapong V, Iamvisetchai P, Pramanus N (Aug 2011) Thai Buddhist Sculpture recognition system (TBuSRS). IACSIT Int J Eng Technol 3(4)
8. Weilong MO, Xiaoshu LUO, Zhong Y, Jiang W (2019) Image recognition using convolutional neural network combined with ensemble learning algorithm. IOP Conf Ser J Phys Conf Ser 1237:022026
9. Faizan A, Najam A, Ahmed Z Image-based face detection and recognition: State of the Art
10. Wiliem A, Krishna Madasu V, Boles W, Yarlagadda P (2007) A face recognition approach using Zernike moments for video surveillance. RNSA Secur Technol Conf
11. Shalunts G, Cerman M, Albertini D (2017) Detection of sculpted faces on building facades. In: 2017 Asia-Pacific signal and information processing association annual summit and conference (APSIPA ASC), pp 677–685. https://doi.org/10.1109/APSIPA.2017.8282119

Autograder: A Feature-Based Quantitative Essay Grading System Using BERT

Roopchand Reddy Vanga, C. Sindhu, M. S. Bharath, T. Charandeep Reddy, and Meghana Kanneganti

Abstract Automated essay grading has become an important area of research in natural language processing. In this paper, we present a new approach for essay grading using BERT language model with Convolutional Neural Networks and Long Short-Term Memory networks on the Automated Student Prize Assessment dataset. The proposed essay grader evaluates essays based on several writing traits such as grammatical mistakes, semantics, coherence, prompt relevance, and others and provide a score for each trait. We evaluate the performance of our model on a fairly large dataset of essays and compare its performance with existing state-of-the-art models. Our results demonstrate the effectiveness of our proposed approach for automated essay grading, achieving high accuracy and improving on the performance of existing models.

Keywords Automated essay grading · Automated student prize assessment convolutional neural networks · Long short-term memory · Automated student prize assessment

1 Introduction

Automated essay grading is an important task in natural language processing because it reduces the time and effort required to grade a large number of essays. For automated essay grading, machine learning and natural language processing techniques have been proposed. However, due to the complexity of natural language and the subjective nature of writing, the accuracy of such systems remains limited.

In this paper, we present a novel method for automated essay grading that combines the Bidirectional Encoder Representations from Transformers (BERT) language model with Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks as it is known to provide best results [1] of slightly better

R. R. Vanga · C. Sindhu (✉) · M. S. Bharath · T. C. Reddy · M. Kanneganti
Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur 603203, India
e-mail: sindhucmaa@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_8

than the reliable models [2]. BERT is a cutting-edge language model that has been shown to perform well in a variety of natural language processing tasks. BERT is used as a feature extractor, and its output is fed into a CNN and LSTM combination. To capture contextual information and the structure of the essays. Our proposed essay grader assesses essays based on a variety of writing characteristics such as grammatical errors, semantics [3], coherence, prompt relevance, and others. We provide the user with more specific feedback on areas that need improvement by using a set of scores to represent each writing trait. For example, if an essay receives a score of 6/10 in prompt relevance, it is clear that the essay is unrelated to the prompt [4]. However, the essay may be strong in terms of other characteristics such as spelling errors and semantics. These scores are already included in the dataset for each essay type, for example. Each essay in Essay Set 7 has a different score, each representing an aspect of the essay.

We assess our model's performance on a large dataset of essays and compare it to existing state-of-the-art models. Our findings show that our proposed approach for automated essay grading is effective, achieving high accuracy while outperforming existing models.

The remainder of the paper is structured as follows. Section 2 goes over related work in automated essay grading. Section 3 goes into great detail about our proposed approach. Section 4 presents the experimental results as well as a comparison of our approach to existing models. Finally, in Sect. 5, we summarize our contributions and plans for future work.

2 Related Work

The paper covers the use of Automatic Essay Scoring (AES) models and makes the case that in order to improve the efficiency of AES, both feature-engineered and end-to-end models are required. The Long Short-Term Memory (LSTM) neural network is used in the paper's two-stage learning framework (TSLF) [5, 6], which incorporates both types of models and determines semantic score, coherence score, and prompt-relevant score. On the Automated Student Assessment Prize dataset, the article also incorporates the Grammar Error Correction (GEC) system into AES models and shows how successful TSLF is (ASAP). The findings reveal that TSLF performs better than baselines on five-eighths of the prompts and, on average, reaches new state-of-the-art performance. Moreover, TSLF has outstanding robustness against adversarial samples, supporting the coherence model, and prompt-relevant model. The paper highlights the importance of handcrafted features such as grammar errors, which are often overlooked in current AES models. To assess written communication skills, automated essay scoring (AES) systems are commonly used. However, when grading essays from various other prompts/topics these systems may encounter accuracy issues for example this model gave a subpar accuracy of around 60% when tested for GRE essay answers.

Phandi et al. [7] proposed a novel domain adaptation technique based on Bayesian linear ridge regression to address these issues. The proposed method reduces the need for a large amount of new training data for each prompt, making the process more efficient and cost-effective. The technique is tested on the Automated Student Assessment Prize (ASAP) dataset and produces competitive AES results. This research provides an intriguing solution for increasing the accuracy and adaptability of AES systems The proposed technique could help reduce the time and cost of generating new training data for each prompt, making AES systems more efficient and effective tools for assessing written communication skills. The novelty of this technique is the use of Bayesian linear ridge regression for prompt-specific adaptation of the AES system, which is a promising approach for domain adaptation [8] in natural language processing tasks.

Hongbo Chen and Ben He's proposed rank-based approach for automated essay scoring (AES) [9] uses listwise learning and linguistic and statistical features to optimize agreement between human and machine raters. Listwise learning is a method used in web search ranking that incorporates rater agreement into the loss function. This method eliminates the need for manual feature engineering and domain expertise, which were previously required. Using the Automated Student Assessment Prize (ASAP) dataset, the proposed approach achieved high agreement with human raters while outperforming existing classification, regression, and preference ranking-based approaches. To aid the learning algorithms, the approach employs linguistic and statistical features such as readability, sentence structure, and lexical features. The success of the proposed method suggests that it could also be useful for other natural language processing tasks. This model like all other reliable models lack feature-based scoring, i.e., it only provides one score which encompasses every feature/trait of the essay such as grammatical errors, prompt relevance in one score.

WordNet is a lexical database (created by Ted Pedersen, Siddharth Patwardhan and Jason Michelizzi) [4] that organizes words into synsets based on is-a hierarchies, enabling the development of similarity and relatedness measures that go beyond surface-level similarities. It differs from traditional grading methods due to its ability to use additional relations and short glosses that provide usage examples. Natural language processing tasks such as information retrieval, text classification, and sentiment analysis have all made use of WordNet-based measures. Despite the benefits of other approaches, using WordNet provides a one-of-a-kind and powerful tool for measuring similarity and relatedness based on a rich semantic structure. This paper emphasizes the benefits of WordNet-based measures and their significance in natural language processing tasks.

3 Methodology

In this section we introduce and explain the methodology behind the autograder. To elaborate, the use of different sets of self-defined features and other prompt wise features provided in the dataset for calculating multiple scores [10], each embodying

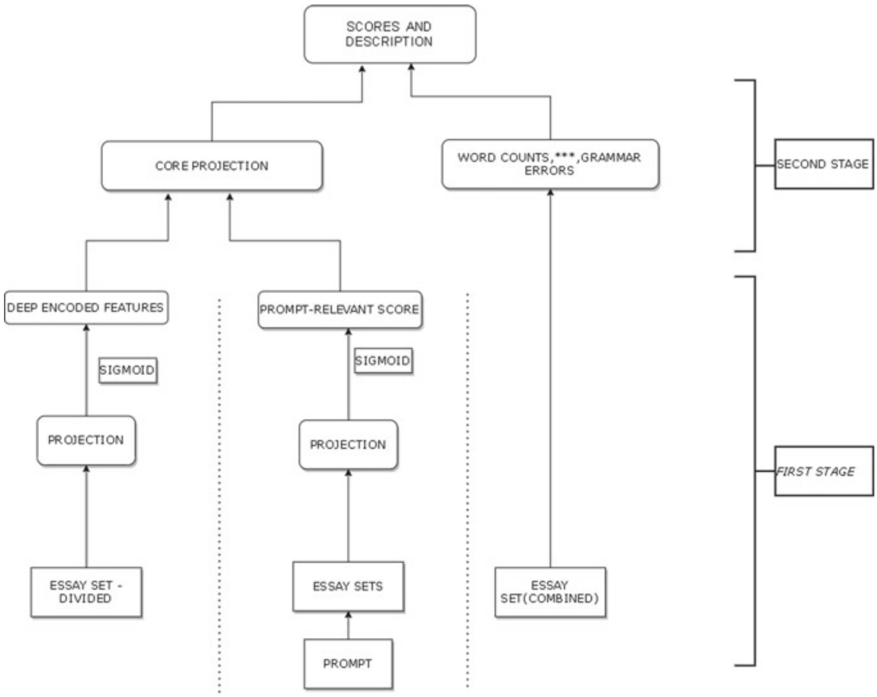


Fig. 1 Flow diagram of autograder methodology

a trait combined with a short description, that are reliable along with the process of adapting a set of features based on the prompt that are to be used in calculating the scores will be explained (Fig. 1).

ASAP dataset is divided into 8 different parts as shown in Table 1 each with its own prompt and a set of scores as shown in Table 2 essay set 1 has 3 scores and no "Rater3" score unlike some other essay sets. Since each essay set has a unique set of scores and provided the ample number of essays per set, we split the dataset into 8 different sets while keeping their respective scores. For example, Essay set 1 has 1 unique score by 2 different raters and a sum of both as shown in the Table 2 similarly set 7 has 4 unique scores each embodying a unique trait of the essay such as structure, coherence.

3.1 Data Cleaning

As mentioned earlier the dataset used in the project is Automated Student Assessment Prize (ASAP). Essays in this dataset has tagged labels such as "@acb" which must be removed to make the data conducive for finding handcrafted features. The scores or the "y" label are different for each set of the essays causing inconsistencies hence

Table 1 Essay set descriptions

Prompt	Essays	Average length	Score range	Number of scores
1	1783	350	2–12	2
2	1800	350	1–6	3
3	1726	150	0–3	4
4	1772	150	0–3	3
5	1805	150	0–4	3
6	1800	150	0–4	2
7	1569	150	0–30	4
8	723	650	0–60	6

Table 2 Split dataset number 1

Essay set	Rater1 domain 1	Rater2 domain 1	Rater3 domain 1	Domain 1 score
1	4	4	NaN	8
1	5	4	NaN	9
1	4	3	NaN	7
1	5	5	NaN	10

all scores must be normalized using Sigmoid function and must be projected back to actual scores before printing the output. Once this data is partially cleaned and essays are used to find the self-defined features for scoring, the data will be further cleaned by using techniques like stop word removal, tokenization, removal of repeated entries, etc.

3.2 *Handcrafted Features*

In the Table 3 are few handcrafted features used in [11]. Some of the 8 essay types in ASAP data such essay set-2 are not only scored based on the writing applications but also the language conventions which encapsulates features such as spelling mistakes, punctuation errors, grammatical errors, paragraphing, word count. Generally, low scored essays either contain very few words and sentences or be very lengthy. This implies writer's subpar skills in writing concisely, features like vocab size, i.e., the number of unique words in the essay helps us understand the writer's vocabulary skill. For grammar, punctuation and spell check it only makes sense to use some existing grammar error correction systems or tools as they check for spelling errors, uncased letters, context, punctuation word repetition, etc. In this model language tool python, as its one of the most accurate systems [12], is used to collect and provide grammatical errors.

Table 3 Selected handcrafted features of essays

No.	Features
1	Character's average and range of word lengths
2	Characters average and range of sentence lengths in words
3	Essay length in words and characters
4	Prepositional and comma usage
5	Word count for original essays
6	Average amount of clauses per sentence
7	Mean clause length
8	Maximum number of clauses of a sentence in an essay

3.3 Sentence Embedding/vectorization

Google's Bidirectional Encoder Representations and Transform (BERT) gave state of the art embedding results and proved itself to be better than word2vec in NLP tasks such as text classification [1]. Given the robustness of BERT, sentence embedding is performed on the stop words removed texts using $\text{BERT}_{\text{base}}$[2]. $\text{BERT}_{\text{base}}$ is a pretrained BERT model trained that has 768 dimensions and 12 encoder layers. "Pooled output" of the BERT result will be used to represent text into a multi-dimensional embedding.

Pooled output: Is representation/embedding of CLS token passed through some more layers BERT pooler, linear/dense, and activation function. It is recommended to use this pooled output as it contains contextualize information of whole sequence.

3.4 Deep Encoded Features

The following explains the model created that can be used to train the datasets with every score provided in the dataset as a target label. The dataset is altered accordingly to fit the case to calculate the score for example if the prompt relevance score isn't provided in the dataset we can combine prompt with every essay before training the model [11, 13].

For every essay $E_x = \{\text{se}_1, \text{se}_2, \cdots, \text{se}_m\}$, where se_n denotes the nth ($1 \leq n \leq m$, $s_n \in R^d$) embedded sentence of the essay and $k = 768$ means the length of sentence embedding. The encoding process of the LSTM is described as follows:

$$j = \sigma\left[M_j \cdot p_t + R_I \cdot h_{t-1} + b_g\right]c \tag{1}$$

$$G_t = \sigma\left[M_g \cdot p_t + R_g \cdot h_{t-1} + b_g\right] \tag{2}$$

$$D_t^{\sim} = \sigma\left[M_d \cdot p_t + R_d \cdot h_{t-1} + b_d\right] \tag{3}$$

$$D_t = it^\circ c_t^\sim + f_t^\circ c_{t-1} \tag{4}$$

$$Q_t = \sigma\left[M_q \cdot p_t + R_q \cdot h_{t-1} + b_q\right] \tag{5}$$

$$u = Q_t^\circ \tanh(c_t) \tag{6}$$

u_t—hidden state of sentence p_t. (M_y, R_y) where $y = (j, D, G, Q)$ are the weight matrices for the input, forget, candidate, and output gates, respectively. b stands for the bias vectors of the specific gates. Sigmoid function $= \sigma$ and $\circ$ means element-wise multiplication. Hence, for every essay, we will get the hidden state set $H = \{u_1, u_2, \cdots, u_m\}$. H_m is passed through the dense layer to convert to scalar value. The values from dense layer output are then projected back to their respective ranges according to ASAP dataset as has different sets of essays with different scoring range mentioned in Table 2.

3.5 Prompt Relevancy Score

This score shows how relevant the essay is to the question/prompt. Just like semantic score Sequential and LSTM model is used. Prompt $= \{s_{1,\ldots\ldots,}s_k\}$ and essay $= \{s_{1,\ldots\ldots,}s_k\}$ where s refers to sentence are combined into one set of sentences. Now the exact same procedure followed for semantic score will be followed here, i.e., after LSTM hidden layer data is fed to dense layer which gives a scalar output scaled to 0–1 range and score 0 is received by essays that are irrelevant. The scores are projected back to their actual score ranges according to the dataset.

3.6 Model Adoption

For the essays that have a different prompt or type and writing instructions it is the deep encoded models gives subpar accuracies as the model is trained on a specific set of data with limited types of essays. To counteract this problem, we apply conditions to test the type of essay, weather it comes with a prompt or not, etc. to categorizes the essay making it easier to choose a model to be more prompt specific or a general model that works well with all types of essays. For example, an unseen essay model, let’s say, essay set 7 from the dataset, won’t work, as it has specific type of scores on which the model is trained, so it would be better to use a simple model that scores the essays based on hand crafted features.

3.7 Training and Evaluation Metrics

For consistency, all scores in the training dataset are scaled down to a 0–1 scale and rescaled to original values before the testing phase, since essay sets in ASAP dataset have different ranges of scores [14]. A monolayered LSTM neural network with a hidden layer size of 1024 is used in our model. It is known that the performance of multi-layered and bidirectional LSTM model is subpar from [11]. A dropout proportion of 0.5 is set in order to avoid overfitting of the model. The epochs are set to 50 and are compiled 5 times per fold of cross validation.

4 Results and Discussion

For evaluation of the results, we use kappa score as the metric. In quantitative grading agreement between scores is more important than equivalence between scores. No two graders are expected to provide similar scores but what's expected is agreement, i.e., for example in case of 3 grader's scores such as 7.3/10, 7.8/10, and 7.6/10 are expected as they all are considered as positive scores but scores such as 8/10, 4/10, 9/10 are not in agreement so it may lead to ambiguity. To put it in perspective the AWE is expected to provide scores that are in agreement with actual human grader's scores to prove accuracy. Since kappa score gives us the degree of agreement between given and predicted scores, we use it for the evaluation of our model.

As we provide multiple scores for each essay in Fig. 2, we cannot compare them like-to-like, so we calculate the arithmetic mean of the accuracy and compare it to other models. For comparisons, we used results from two-stage learning (TSL) [11], which used CNN, CNN + LSTM, and LSTM algorithms, and Bayesian Linear Ridge Regression (BLRR) and Support Vector Regression (SVR), both of which utilize domain adaptation methods [15], as well as other handcrafted features such as parts of speech. We use these models as they are similar to the methods used in our model and are proven to be reliable [16]. Based on their high kappa values and wide array of grading features, these models were selected for comparison. Our model is compared with other reliable models in Table 4 and Fig. 2.

Developing an AWE that works consistently for all prompts of the essays is vital. BERT proved itself to be the best when it comes to contextual text embedding and the pretrained $\text{BERT}_{\text{base}}$ model is vast enough. Handcrafted features were definitely necessary for both scoring and providing mistakes in the output. Provided the right set of features are chosen in the model with right tools to extract those features. The used method is unique and very useful as is provides many different scores to each essay which is unlike any model that exists.

The average kappa score of our model has shown us good results Fig. 3, a score of 0.81, but not ahead of the existing highest accuracy of 0.821 of TSM. The main focus of our model is to provide multiple scores and descriptions that are useful and reliable. Each score representing each aspect of the essay such as organization,

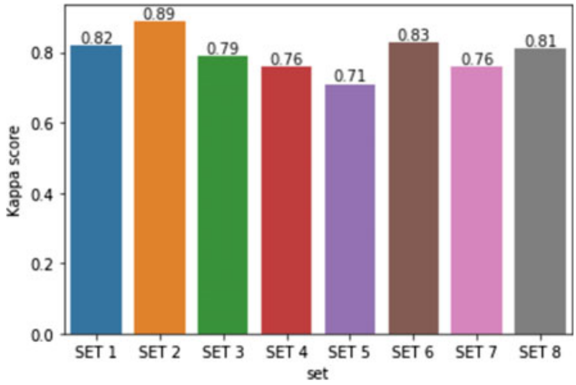


Fig. 2 Set wise kappa score split

Table 4 Comparison of different results from different papers

MODEL	SCORE
EASE(SVR)	0.699
EASE(BLRR)	0.705
CNN	0.726
LSTM	0.756
CNN + LSTM	0.761
TSM	0.821

prompt relevance, structure, style combined with description explaining the quality of the essay would make self-evaluation much easier to the user so a good accuracy is expected for each and every type of score provided by the model.

For the description part, since the scores are provided for different aspects of the essay the model provides some default description for each type and range of the score. For example, essay set 7 has 4 different types of scores each grading style, organization, punctuation, and prompt relevance of the essay. For an organization

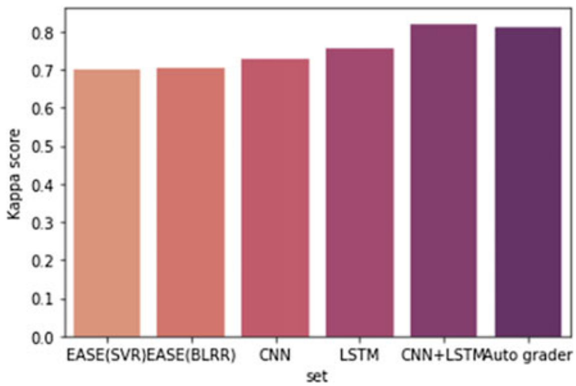


Fig. 3 Comparison of different results from different research works

score of 3/4 the description provided would be "Organization and connection between sentences, ideas, and events is mostly logically and clearly sequenced with very few issues".

5 Conclusion

We recommend our model as it provides multiple scores and descriptions with great accuracy. It uses a robust set of handcrafted features combined with prompt relevance and a deep encoded feature model for scores provided in the dataset for evaluation. The model also adapts to the type of essay to only give the specific scores for that specific type of essay making it useful for any type of essay unlike other models that are biased toward the type of essays provided in the dataset. The results of the studies demonstrate how effective our model is on the ASAP dataset; our model outperforms numerous trustworthy baselines and performs similar to the existing best in class models with a kappa score of about 0.81. In conclusion, both manually created and vector-encoded characteristics and adaptability are used to provide our system its strength and capabilities to work with any type of essay irrespective of the availability of the essay in the dataset the model is trained upon.

References

1. Taghipour K, Ng HT (2016) A neural approach to automated essay scoring. In: 2016 EMNLP. https://aclanthology.org/D16-1193
2. Attali Y, Burstein J (2006) Automated essay scoring with e-raterR v. 2. J Technol Learn Assess 4(3)
3. Yang Y, Xia L, Zhao Q (2019) An automated grader for Chinese essay combining shallow and deep semantic attributes. IEEE Access 7:176306–176316. https://doi.org/10.1109/ACCESS.2019.2957582
4. Pedersen T, Patwardhan S, Michelizzi J (2004) Wordnet: similarity: measuring the relatedness of concepts. In: Demonstration papers at HLTNAACL 2004, Boston, Massachusetts. Association for Computational Linguistics, Stroudsburg, PA, pp 38–41
5. Prabhu S, Akhila K, Sanriya S (2022) A hybrid approach towards automated essay evaluation based on Bert and feature engineering. In: 2022 IEEE 7th international conference for convergence in technology (I2CT), pp 1–4. https://doi.org/10.1109/I2CT54291.2022.9824999
6. Liu J, Xu Y (2019) Automated essay scoring based on two-stage learning. https://doi.org/10.48550/arXiv.1901.07744
7. Phandi P et al (2015) Flexible domain adaptation for automated essay scoring using correlated linear [regression]. In: 2015 EMNLP. https://aclanthology.org/D15-1049
8. Heilman M, Madnani N (2013) ETS: domain adaptation and stacking for short answer scoring. In: Proceedings of the seventh international workshop on semantic evaluation
9. Chen H, He B (2013) Automated essay scoring by maximizing human-machine agreement. In: Proceedings of the 2013 Conference on empirical methods in natural language processing
10. Larkey LS (1998) Automatic essay grading using text categorization techniques. In: Proceedings of the 21st annual international ACM SIGIR conference on research and development in information retrieval. ACM, pp 90–95

11. Alqahtani A, Alsaif A (2019) Automatic evaluation for arabic essays: a rule-based system. In: 2019 IEEE international symposium on signal processing and information technology (ISSPIT), pp 1–7. https://doi.org/10.1109/ISSPIT47144.2019.9001802
12. Näther M (2020) An in-depth comparison of 14 spelling correction tools on a common benchmark. In: 2020 LREC. https://aclanthology.org/2020.lrec-1.228
13. Sindhu C, Vadivu G (2021) Fine grained sentiment polarity classification using augmented knowledge sequence-attention mechanism. J Microprocess Microsyst 81
14. Wang Y et al (2022) On the use of Bert for automated essay scoring: joint learning of multi-scale essay representation. In: NAACL 2022. https://aclanthology.org/2022.naacl-main.249
15. Li X, Liu J (2016) Automatic essay scoring based on Coh-Metrix feature selection for Chinese English learners. SETE@ICWL
16. Shermis MD, Burstein JC (2002) Automated essay scoring: a cross-disciplinary perspective. Lawrence Erlbaum

Bluetooth Automated Wheelchair

K. Iniyan, I. Aravindhan, J. K. Ajin, S. Kirivarshan, and K. Malathi

Abstract The wheelchairs may all move inside and around manually. But a smart wheelchair gives a person independence and ease. A smart wheelchair is a mechanically controlled vehicle that can move on its own when a simple head or hand command is given by the user. As a result, using the wheelchair's wheels requires less effort from the user. Moreover, this enables people who are physically or visually handicapped to move between two locations. Even if a person's body is completely or partially paralyzed, only the movement of his head or hands will allow a wheelchair to go forward, backward, left, and right. The wheelchair and the human are in wireless communication. He keeps the apparatus that detects human movement on his head or in his palm. The chair will automatically move forward if the device detects a head swing. The process is comparable to another such as left, right, and backward. If the head is stationary and not moving, it is regarded as having stopped. The experimental version of our product works by pairing the mobile Bluetooth with the Bluetooth module installed in the wheelchair. The Arduino Uno supports Bluetooth and wheel movement and detects human movement.

Keywords Wheelchair · Bluetooth · Automatic · Arduino

K. Iniyan (✉) · I. Aravindhan · J. K. Ajin · S. Kirivarshan · K. Malathi
KCG College of Technology, Karapakkam, Chennai 600097, India
e-mail: 19it16@kcgcollege.com

I. Aravindhan
e-mail: 19it04@kcgcollege.com

J. K. Ajin
e-mail: 19it02@kcgcollege.com

S. Kirivarshan
e-mail: 19it47@kcgcollege.com

K. Malathi
e-mail: Malathi.it@kcgcollege.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_9

1 Introduction

1.1 A Bluetooth Automated Wheelchair

This document is intended to make life easier for those who are unlucky enough to lose the use of their legs as a result of serious paralysis, an accident, or growing age. Most persons with disabilities rely on others especially when transferring from one location to another in their daily lives. The people who use wheelchairs require ongoing assistance from another person to move the wheelchair. Because their wheelchairs lack an easy-to-use control mechanism that enables autonomous movement so that their lives are rendered more challenging. When a person with a physical disability uses an electronic wheelchair because they are unable to walk or use a motorized vehicle, they become significantly more independent.

Those with physical disabilities who are unable to walk or use a mechanical wheelchair on their own gain a great deal of independence by using an electronic wheelchair to use, smart wheelchair solutions are still not always widely available.

The suggested methodology, which is covered in this article, will be presented as the easy-to-use, low-cost option for voice-controlled platforms. It is easy-to-use, entirely linguistically adjustable, and will promote the user's independence in movement. Another relatively new and highly successful technology is Bluetooth, it has changed wired. Digital devices have evolved into wireless ones, changing how people use them at home and at work. This study is based on the design and implementation of Bluetooth-enabled mobile platforms and wireless remote controls for voice-controlled wheelchairs. The project also uses ultrasonic sensors to stop the wheelchair from further instruction and to identify obstructions within a 4-m range. This study shows how a smart wheelchair may be controlled using an Android application and an Arduino Uno microcontroller. A wheelchair is necessary for a disabled person to do tasks that involve moving around when carers cannot constantly actively monitor patients.

2 Literature Review

In this paper [1], they submitted a design for an affordable smart wheelchair that combines amenities to support their ability to move freely and safely. There are two operating modes for this design: automatic and manual. Voice recognition is turned on when in automated mode. The user can use the wheelchair using the AMR voice application. The stop instruction, along with the commands for left, right, forward, and reverse, will all be handled by the microcontroller. The processed signal from the gathered signal will subsequently be received by the L289N module motor driver. The wheels will then be propelled when necessary, by two DC-geared motors. In manual mode, an Arduino Nano is attached to a dual-axis joystick that controls the wheelchair's movement. Frontal HC-SR04 ultrasonic sensors attached behind the chair measure the echo wave distance and signals a buzzer and light to indicate a potential obstruction. MPU605, an emergency voice message call, delivered to the user's kin or the closest hospital.

In this paper [2], they represented the wheelchair with health monitoring system. The various sensors present in the prototype along with health monitoring system makes module which is very reliable and helpful. In the future, voice controlling system or IR sensor glasses can be installed in the existing prototype to increase the mobility level of chair to a very high standard. Gear box can be installed to increase the speed of chair and handling. Solar power panel can be installed for promoting eco-friendly charging. A detachable metallic staircase can be installed to climb slopes and small hurdles.

In this paper [3], they represented several smart wheelchair. The severity of the handicap, the person's morale, attitude, and the cost all affect how much attention is paid to features like the human–machine interface and navigation techniques and tools when utilizing high-tech smart wheelchairs. Future research should put more emphasis on add-on approaches that allow for flexible sensor, interface, and input device configurations tailored to the needs and financial constraints of each individual user. By including added functions like health monitoring, first aid, muscle relaxants, and rehabilitation tools, wheelchairs can be improved.

In this paper [4], they provided an overview and suggested areas for further investigation. Smart wheelchairs have been developed for disabled people with manual support. Technical issues such as cost versus accuracy trade-off and standard communication protocol. Clinical acceptance and reimbursement remain even after the technical issues are overcome. Due to the high cost and complexity of smart wheelchairs, considerable infrastructure and resources are needed for familiarization and training efforts. For use in enclosed locations to prevent drop-off access, modern smart wheelchair technology is available. Sensor technologies will help to further improve the circumstances under which intelligent wheelchairs can function securely.

In this paper [5], in 2014, Andrej Škraba et al., presented a prototype speech-controlled cloud-based wheelchair platform. The control of the platform is implemented using low-cost available speech WebKit in the cloud. Besides the voice control, the GUI is implemented which works in the web browser as well as on the mobile devices providing live video stream. The odometrical data was used to locate the mobile when the trajectory was automatically carried out. There are two problems while using mobile platform. One involves teaching the user how to utilize a complex technology, and the other involves having enough time to react when a person has a significant physical impairment. During mobile navigation, trajectory errors are difficult to correct with local servo-controls. The machine will need a specific process to enhance local servo-control and add some reasoning capabilities.

3 Proposed Model

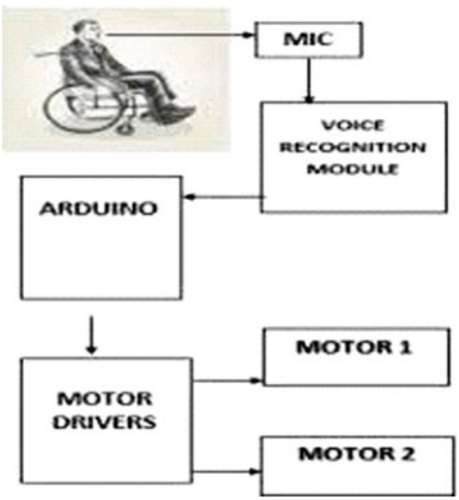
The voice-controlled wheelchair is the subject of the suggested model. In this project, a robot chair is constructed first, followed by the installation of several devices for the old and ailing people. By utilizing the voice command, this may be moved anywhere, and the button will be utilized to control it when sat in the chair. The elderly and sick were the original target audience for the wheelchair. This model (Fig. 1) represents the architecture diagram.

The spoken commands or button presses of the wheelchair user operate this system. Since the user does not need a second person to help move the wheelchair, the technology is entirely self-sufficient. Typically, there are five commands. The wheelchair will move accordingly in response to the user's commands. In Fig. 1 architecture diagram, in the first stage, the voice command is acknowledged in the initial step. Voice recognition and motor driving are the two main programs and components that make up this system. A speech recognition program is used to carry out the voice recognition.

This Bluetooth module's output is delivered to Arduino, which uses a motor driver IC to power the motors. The speech-controlled wheelchair is propelled by motors, an Arduino, Bluetooth, motor driver module, and a unilateral voice recognition program. The system receives input through the phone application. It can obey human voice commands while ignoring outside noise. The location of the mobile application will be determined by the output preference of the user.

According to the user's preferences, the mobile application will be placed. Voice recognition software, which acts as a bridge between the Bluetooth module and the Arduino (9–12), receives speech signals as an output. The output from the speech recognition program is then transferred to binary code by Arduino. Only binary code may be understood by the system.

Fig. 1 Architecture diagram



In the first stage, the voice command is acknowledged in the initial step. Voice recognition and motor driving are the two main programs and components that make up this system. A speech recognition program is used to carry out the voice recognition.

This Bluetooth module's output is delivered to Arduino, which uses a motor driver IC to power the motors. The speech-controlled wheelchair is propelled by motors, an Arduino, Bluetooth, motor driver module, and a unilateral voice recognition program. The system receives input through the phone application. It can obey human voice commands while ignoring outside noise (Fig. 2).

The location of the mobile application will be determined by the output preference of the user. According to the user's preferences, the mobile application will be placed. Voice recognition software, which acts as a bridge between the Bluetooth module and the Arduino (9–12), receives speech signals as an output. The output from the speech recognition program is then transferred to binary code by Arduino. Only binary code may be understood by the system.

It is illiterate in all other languages. The resulting voice command is then transformed into a format that is machine-readable. This arrangement makes use of an Arduino Mega. It has forward, left, and right movements which are the three basic movement directions. The user who will use the wheelchair must shout out words for the voice recognition processor to learn how to recognize them. Both switch control and voice command are employed in this project.

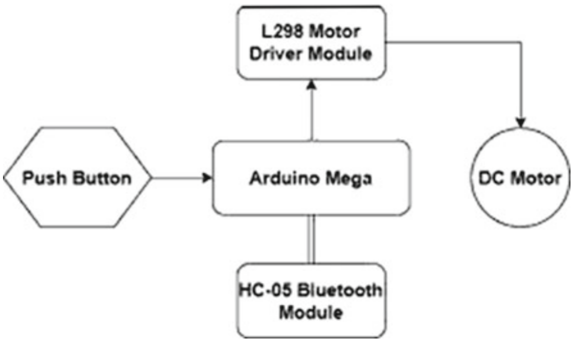


Fig. 2 Block diagram

4 Requirement Specification

4.1 Hardware Requirements

Arduino Uno Board

A microcontroller called Arduino Uno contributes to the accessibility of interactive object applications. A bootloader that is pre-programmed into the onboard microcontroller chip enables uploading program into the memory of the microcontroller without the need for a chip or device programmer.

L298 Module

The L298 module is a dual H-bridge motor driver integrated circuit (IC) that is commonly used in robotics and other motor control applications. The module includes two H-bridges, which can control the direction and speed of two DC motors or a single stepper motor. The L298 module is popular because it can handle high current and voltage loads, making it suitable for controlling larger motors.

Wheels

Wheels reduce friction, making it far simpler to drag a weight on a cart than it is to drag it on the ground. Instead of just rolling over the ground, the wheels sink down and rotate around sturdy rods known as axles.

Module Bluetooth

It serves as a prototype Bluetooth Serial port module. It is simple to program as both a master and a slave. The Bluetooth Module HC-05 makes serial wireless data transfer straightforward. The 2.4 GHz ISM frequency band corresponds to its operating frequency. It complies with the Bluetooth 2.0 + EDR standard. The time it takes for various Bluetooth 2.0 devices to transmit their signals is 0.5 s. The use of second intervals reduces the stress placed on the Bluetooth chip and extends Bluetooth's sleep time. With its serial interface, this module is simple to use and takes less time to construct than other modules.

4.2 Software Requirement

Arduino Software

The Arduino software, often known as the Arduino Integrated Development Environment (IDE), includes a text editor for writing code, a message box, a text console, a toolbar with buttons for essential operations, and a variety of menus. A connection is created so that program can be uploaded and talk to Genuine and Arduino hardware. Using the Arduino IDE, sketches are computer programs. These images were created with a text editor, and the files with the extension include them. The console shows comprehensive error messages, extra data, and text generated by the Arduino software (IDE).

The board that has been configured and the serial port are displayed in the window's bottom right corner. You may make, view, edit, and validate sketch. Verify the application, then upload it. Using the toolbar button, you may also launch serial monitor and other program. Installation and configuration of the Arduino software: The computer has the Arduino software installed. The arduino.exe program was then installed once we accessed the software file. There are two crucial program settings that need to be looked at. (a) Choose the board you want to connect in the software. The "Arduino/Genuine Uno" is also known as the "Arduino Uno." (b) We must select the proper "Serial Port" in order to inform the computer to which port the board has been connected.

Arduino Blue Control App

With the additional functionality offered by the Arduino Blue Control app, you may use Bluetooth to control your Arduino board and create amazing projects that are completely unique.

5 Result and Discussion

A Bluetooth automated wheelchair for those with physical disabilities. Despite GPS back and sensor, the robotized wheelchair framework focuses on characteristics such as line adjustment, speed, and organically moving wheelchair (Fig. 3).

Once the flag-conditioning microcontroller has acquired the data, analyzed the clients' input, and prepared the study, individual has their flag-conditioning sheets.

Fig. 3 Bluetooth automated wheelchair

6 Conclusion

In conclusion, aim of the project's implementation is to assist everyone whose movement depends on a wheelchair. Wheelchairs are easy-to-use and do not require any outside assistance. This is a design for an android application that can control how a wheelchair moves. It was created to aid the physically impaired and elderly maneuver their wheelchairs on their own, as well as to give the elderly the capacity to do so using Android smartphones. The developed system has successfully completed some testing and the fundamental performance.

References

1. Al Shabibi MAK, Kesavan SM (2021) IoT based smart wheelchair for disabled people. National University of Science and Technology, Muscat, Oman
2. Nayak SS, Gupta P, Upasana AWB (2017) Wheelchair with health monitoring system using IoT. Int Res J Eng Technol (IRJET) 04. e-ISSN: 2395-0056
3. Simpson RC (2005) Smart wheelchairs: a literature review. J Rehabil Res Dev (JRRD) 42(4):423–436
4. Bourhis G, Moumen K, Pino P, Rohmer S, Pruski A (1993) Assisted navigation for a powered wheelchair. Systems engineering in the service of humans: proceedings of the IEEE international conference on systems, man and cybernetics, France, pp 553–558
5. Skraba A et al (2014) Prototype of speech controlled cloud based wheelchair platform for disabled persons. In: 2014 3rd Mediterranean conference on embedded computing (MECO). IEEE

SCADA in Healthcare

Vanishree Pabalkar and Rajkumar Verma

Abstract The Supervisory Control and Data Acquisition (SCADA) systems play a significant role in supporting and ensuring the remote access, along with monitoring and controlling of the systems through critical infrastructures. SCADA stands for Supervisory Control and Data Acquisition. The rapid growth in the communication system protocols in the e-hospitals that comes with a sensor infrastructure and convenience of remote access that SCADA provides invites few challenges as well. The invasive form of action on SCADA can evoke major threatening repercussions. At the same time, SCADA is significant tool that plays a vital role in assessing threats in the security system. This enables the early detection of potential malicious that causes within the SCADA device network. Supervisory tools and techniques for SCADA invasion and detection have been proposed by many researchers. SCADA consists of RTU—remote terminal unit, PLC—programmable logic controllers, central host operator controllers. SCADA basically procures the information, transfers the info to central site and further alerts the home station about the data breach in the central system; this could be a leakage in the data, like patient's data, patient's history in a hospital. This data can be analyzed to further detect the consequences of the data breach. SCADA can control the healthcare systems and can achieve centralized monitoring systems. In the future, in case of pandemic like calamities, SCADA will be of great help through monitoring the healthcare system like the health of patient, public health, early detection of ailment.

Keywords SCADA · Classification · Classification infrastructure · Supervised learning · Industry control system · Healthcare · e-hospital · Sensor infrastructure

V. Pabalkar (✉)
Symbiosis International (Deemed University), Pune, India
e-mail: vanishree.p@sims.edu

R. Verma
Imagineer Private Limited, Bangalore, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_10

1 Introduction

1.1 SCADA Systems

Supervisory Control and Data Acquisition (SCADA) has a vital role to play in several areas, like hospitals, manufacturing plants, power systems [1]. The proposed topic is about SCADA in the healthcare systems. Though this comprises of complexities in the process, the innovative idea is proposed to be executed subsequently. The timely detection of SCADA threats has been posing as challenges. The advancement in technology and presence of Big Data that can be extracted through SCADA sensors, it becomes imperative to devise tools that will enable the inconsistent models in the SCADA. This will further enable to arrive at desirable solutions in the prediction and detection with SCADA. SCADA is the computer system-based control system that enables procure real-time data. This real-time data is then assessed in order to have the access to the systems through remote. The aim is to monitor and control equipment, healthcare systems, diagnosis, and labs.

SCADA refers to in SCADA environments and industries, they are utilized to manage a range of equipment and several servers. Multiple communication protocols may be used by a single SCADA system. SCADA systems are frequently employed to automate complicated operations when human control is impractical. SCADA applications include the difficulties of providing round-the-clock monitoring. Basic SCADA control system implementation can quickly generate value through staffing logistics and expenses. Modern SCADA systems are useful for industrial operations with a lot of control elements. System operators can benefit from the careful construction of rule sets by having help with common administration tasks. Data will be processed automatically, and directives will be sent. You can concentrate on systems that use human–machine interface (HMI) software (Fig. 1).

SCADA—growth year-wise

With a projected CAGR of 6.3% from 2021 to 2030, the global Supervisory Control and Data Acquisition (SCADA) reveals to grow from its estimated value of US$ 35.38 billion in 2021 to more than US$ 61.22 billion by 2030.

A SCADA system combines the hardware with the software. Through the software, the automation process starts with the help of real-time data. SCADA gets connected through sensors. These sensors monitor equipment like the machines, tools, motors, pumps, and valves at the remote server.

- SCADA System architecture can be categorized as follows:
- First generation termed as the monolithic SCADA.
- Second generation termed as distributed SCADA.
- Third generation termed as networked SCADA.

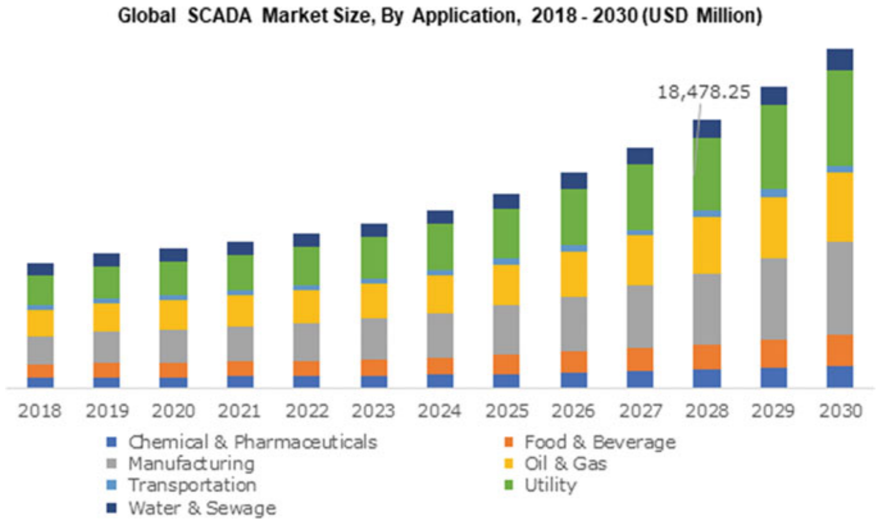


Fig. 1 SCADA—growth year-wise

The SCADA system empowers the following functions for an organization.

- SCADA ensures to control of the processes either in a local environment of through remote locations
- SCADA acquires, assesses, and produces the real-time data
- SCADA interacts with the industrial equipment like machines, tools, sensors, valves, pumps, and motors
- SCADA records and archives the history that can be used for future references.

The initial SCADA systems were equipped with statistics and theories. The SCADA systems come with very complex nature. Therefore, the technological advancements are expected in line with data-driven technology and machine learning algorithms (Fig. 2).

In the SCADA hardware system, the remote terminal units (RTUs) and programmable logic controllers (PLCs) act as the points for local connection. The crucial and sensor information is procured through these local connection points. There are several edge workloads that use computer hardware in different ways.

- Gateways are responsible for surpassing the data from PLCs to servers or to the edge.
- Edge computers are considered to be near to the main source of the data. They can as well be considered as the gateway.
- Human–machine interfaces (HMIs)—Through this, a local touchscreen is provided that acts as the interface for machine monitoring and control.
- The server itself—This can be the central control for the local SCADA system. Location under that the action happens.

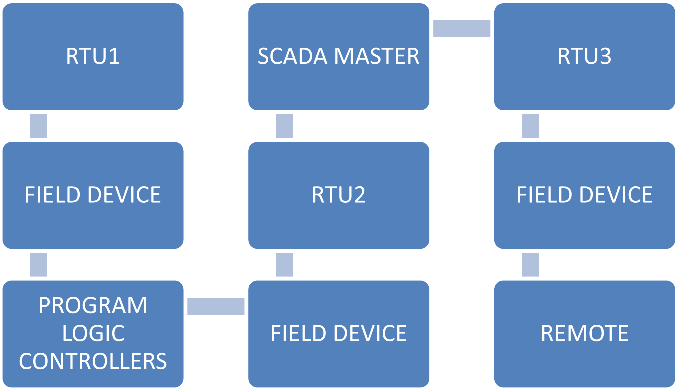


Fig. 2 SCADA system hardware

Historically, the method to examine alerts and techniques has been SCADA host software. Control tools for remote access were established, and the control from the SCADA host itself is possible. These systems belonged to operators, technicians, and engineers and were shut off from the outside world. Monitoring, maintaining, and engineering processes and SCADA components fell within their purview. Information technology (IT) improvements have changed this; therefore, it is no longer true. The data that the SCADA host software creates now needs to be accessible in real time to a wide range of different stakeholders.

- *SCADA Host Platform*

SCADA Host Platform—With the help of data derived from the SCADA system, accounting, maintenance management, and material purchasing requirements are performed, at least in part.

1.2 Materials and Methods

In the current research, study has been on the Supervisory Control and Data Acquisition (SCADA) systems that play a significant role in supporting and ensuring the remote access, along with monitoring and controlling of the systems through critical infrastructures. The emphasis here is to use SCADA in the healthcare systems. Through the remote access, the lives of patients could be saved. The study was conducted by use of Scopus databases, Google Scholar and web of science databases. The methodology in the study comprises of referring to published work on similar lines. For the current study, the emphasis was done on papers that were published in recent times on the similar lines. After referring through the research papers on the topic, the focus of this paper is on understanding the role of SCADA in healthcare systems.

1.3 SCADA in Healthcare

SCADA control systems in healthcare enable the medical experts and the tools, machines and other equipment to be operated via remote locations. There are several merits by adapting this method by SACDA.

They include the following:

Cost saving
Fast, accurate, and enhanced service levels
Coordination of healthcare experts and auto systems
Dynamic process as per the requirements and specifications of the patient
Use of digitalization
Remote access enables using of the existing Infrastructure and saves cost
The monitoring system is through the web browsers, which is considered to be safe and sound.

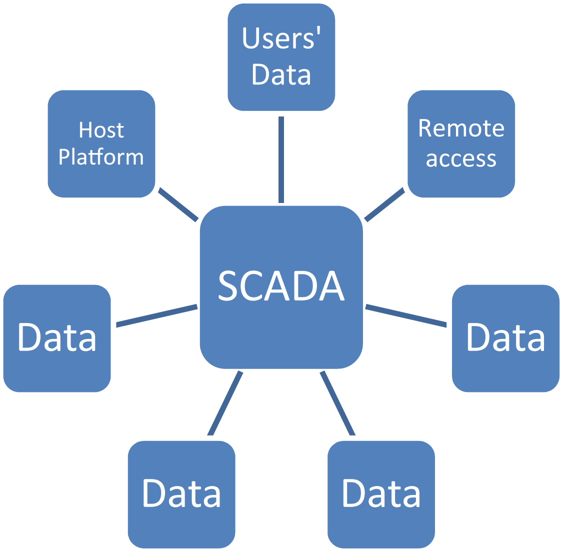
1.4 Background of SCADA Architecture

SCADA system comprises tools and plethora of software applications. Field devices, RTUs, MTUs, programmable logic controllers (PLCs), intelligent electronic devices (IEDs), HMIs, historians, etc., are important architectural systems of a contemporary SCADA system [2].

The streamlined architectural layout of a SCADA system is shown in Fig. 3. In a typical SCADA information is gathered through field tools like sensors and actuators that interact with the CI's actual physically present tools like valves and machines [3]. The field devices, which direct and supervise the activities of the CIs process, are controlled and monitored by the RTUs, PLCs, and IEDs. RTUs, IEDs, and other modern devices are technologically sophisticated. SCADA refers to cyber-physical systems, which means an activity in the cyber-world may cause a change or action in the real world. They are typically employed to keep an eye on and manage physical elements. Unplanned changes or outages can therefore become extremely harmful when discussing networks, algorithms, and computing technology in the context of healthcare, like a hospital. We are still processing the most recent Industrial Revolution as we transition to Industry 4.0 through the evolving of cyber-physical systems, through this, humans and machines tend to collaborate more closely as equals and humans will delegate an increasing amount of their work to machines.

The following Industrial Revolution will see an increase in human–machine cooperation as humans delegate an increasing amount of their work to machines. Even if it might take some time, the cyber-security sector needs to start working now to better protect SCADA and healthcare settings before the next Industrial Revolution. Unfortunately, there are numerous cybersecurity issues that SCADA and Health Information Systems (HIS) have in common, including outdated hardware, the usage of

Fig. 3 SCADA host platform



unsupported operating systems, and the exposure of sensitive and important devices to the internet.

Companies have needed 280 days, or more than 9 months, on average across all industries to discover and contain a breach. These figures worsen when only healthcare is taken into account, with lifespan of 329 days, or almost 11 months, with 48% of breaches coming from hostile attacks, 27% from human mistake, and 23% from system flaws [4]. However, we are already noticing an increase in the number of e-clinics. To safeguard the SCADA section of healthcare facilities, a range of procedures must be employed, recognizing that a single security product or technology cannot properly protect a SCADA. It consists of a combination of security rules and a set of security controls that have been correctly set up to match the technologies used in the environment. A SCADA's operation may be significantly impacted by the choice and application of security controls.

- Which security measures are required to effectively reduce risk to a level which provides help and assistance supports the organizational missions and operational needs?
- Has this chosen security control exposed to practice, or is there a workable implementation strategy in place?
- To what extent must the chosen security controls be assured?

The essence of data in today's world is crucial and sensitive at the same time. It therefore becomes imperative to protect this data. With SCADA, the healthcare systems can be made more effective. Unlike traditional systems, monolithic monitoring system exists. The data might be not monitored. But with SCADA, the data monitoring is possible in real time in both ways. The both ways include the patient and the hospital for instance. These data points help in enhancing the patient's life in a better way. This also aids in enhancing the public safety, organization safety, and compliance can be maintained and can be retrieved as when required. SCADA, with the help of PMC, HMI, RTU can be a HITRUST framework that can be adapted throughout hospitals for HIPPA compliance.

HIPPA deals with compliance, and HITRUST is a framework. The hospitals deal with HIPPA to a great extent. A tool for assisting healthcare organizations and service providers in adhering to HIPAA's privacy and security requirements is HIPAA compliance software. It functions by putting HIPAA regulations into practice to stop breaches and unauthorized PHI sharing. There will be no alternate option for an active, integrated, fully effective HIPAA compliance Program that is firmly anchored and a well-built Culture of Compliance, according to our Team of Info Sight Inc. specialists. HIPAA establishes the bar for safeguarding private patient information. Protected health information (PHI) handling organizations are required to implement physical, network, and procedural security safeguards, or else they risk severe penalties and fines. The requirements for continually securing electronic protected health information. Providers, payers, clearinghouses, and entities all need to secure EPHI. InfoSight Inc. provides guidance for HIPAA Compliance Assessment to ensure policies, technical security mechanisms, and physical safeguards, ensuring HIPAA Compliance and Assessment which catch new violations.

Provider Information—HIPPA

According to the US Department of Health and Human Services, InfoSight is a compliant hosting provider with administrative, technical safeguards in place. InfoSight is aware of HIPAA Compliance Assessment and Regulations and will host your data. These precautions, together with information on what qualifies as a HIPAA compliant data center, Audit Reports, Tracking Logs, Technical Policies, Network, and Transmission, are pertinent to the services offered by our HIPAA compliant hosting. A supplement to HIPAA strengthens the enforcement of the privacy and security rules and increases the fines for health organizations that breach them. In response to the advancement of health technology and the growing usage, storage, and transmission of electronic health information, the HITECH Act.

HITRUST framework is used and HEPA—health info trust alliance

HITRUST—A nonprofit organization called The Health Information Trust Alliance (HITRUST) provides data protection standards and certification programs to assist organizations in protecting sensitive information, controlling information risk, and achieving compliance objectives.

As per HITRUST framework—understanding of business contexts, scope of systems, GAP analysis and risk assessment, strategy policy documentation, monitoring measurement and number of domains can be maintained by HITRUST. The domains that retain patient's data and organization data. This can be maintained through HITRUST domain. Information Protection Program—Security for SCADA has changed significantly in recent years. Before the invention of computers, the only method to keep track of a SCADA network was to send many personnel to each station, who would then return with a report on the status of each system. Technicians were permanently stationed in busier stations to manually run the network and communicate over telephone connections.

Previous studies indicate that provide evidences of all the critical aspects in the field of cyber security of SCADA networks. Studies have also revealed the SCADA system along with the properties of SCADA communication protocols. SCADA protocol analyses, cryptography the remote operating system security, and saves time and cost. The critical infrastructures are always accompanied by external threats like data leakage and cyber-attacks [1]. Cyber-risk management comprises of detecting the leakages in the systems and assess the ways and means to minimize these leakages [5]. The role of SCADA is to detect the vulnerabilities in the system and to ensure that these are attended by SCADA [6]. SCADA systems comprise sensors, remote sensing and remote terminal units. These are vulnerable to various intrusions and obstacles [7]. SCADA-based energy transmission system and RTU transcend the data of node and include warnings in cases where the data is extreme [8]. Machine learning also is very efficient in securing the SCADA systems [9]. SCADA system gathers the information through sensors and involves the critical infrastructure systems that include pumps and valves [10].

Wireless security—The industrial environment is paying more attention to wireless communication. To offer hassle-free access for IT networks, several organizations have switched from wired to wireless networks. The user can access the network practically anywhere thanks to wireless technologies. Wireless networks are more vulnerable to attack due to connectivity. This article will examine wireless attacks on the environment of industrial control systems (ICS), which frequently result in business interruption.

Configuration—Any set of components can be chosen to serve the load in this design. Compared to separating components into redundant systems, this offers more flexibility, but it necessitates shared control over all parts, making the PLC a potential single point of failure.

Vulnerability—Keep your private information out of the wrong hands. Forcepoint is available to assist. Data security need not be challenging. Allow Forcepoint to simplify things for you. Set up a demo. Prevention of Data Loss. Cloud Safety. Intelligent Cybersecurity.

Transmission protection—The protection of electric power systems for power generation, transmission, and distribution is the focus of the electrical power engineering subfield known as protection and controls (P&C). P&C engineering and design is the art of foreseeing abnormal power system conditions and protecting the power grid while simultaneously launching some corrective activities. P&C entails safeguarding substantial and pricey power equipment, such as:

Power circuit breakers, generators, transformers, transmission lines (TLs), electrical buses, and conductors.

Access control—Access control for critical infrastructure requires moving the perimeter to workloads and managing access based on context. This zero-trust approach ensures access based on user/device characteristics, target workloads, and associated risk. In this article, I give an example of one of two general approaches to achieving zero-trust ICS networks.

Password Management—Identity and Password Manager Solutions: Protect Data from Cyberthreats and Secure Private Information. Passwordless Authentication Can Streamline Employee Access to Digital Records and Tools. Cost-effective pricing. Password Stress Must Go.

Incident management—system for incident management. Supervisory Control and Data Acquisition (SCADA) systems are frequently used to manage industrial control systems (ICS), which offer a graphical user interface for operators to easily view a system's status, receive any alarms indicating out-of-band operation, or enter system adjustments to manage the process.

Risk management—SCADA systems may be at danger from malware, including as viruses, spyware, and ransomware. Malware can still be a threat to the critical infrastructure that supports the management of the SCADA network even though it may not be able to target the network directly.

Business continuity and Disaster recovery—What are the business continuity management's three essential elements? Business continuity, disaster recovery, and SCADA image as an incident renders the SCADA hard drive or computer hardware unusable, disaster recovery is a way to recover. Failures can occur for a number of causes, including the structure floods. The motherboard or hard drive gets destroyed by a power quality event (such a spike).

Physical environmental security—preventing physical access by unauthorized people. Prevention of system removal or other physical alteration, destruction, or theft. Preventing the observation of information assets and sensitive data.

Data protection and Privacy—The practice of designing and defending Supervisory Control and Data Acquisition (SCADA) networks is known as SCADA cybersecurity.

Cyber security framework will be used to maintain this. With SCADA, the data is available at one centralized system. Cyber security framework can be used accordingly. Data would be authentic, through cyber security. This will enable for behavioral analysis that aids in not only maintaining the health of the patient, but also in having control and access to secured data with the help of SCADA.

SCADA system protocol T101, DNP3 do not fire in cyber security. But in framework or in compliance, these protocols can be enhanced so that the data is channelized and is stored appropriately so that it can further be analyzed. With these protocols, cyber-security framework can be associated with SCADA and it can overall help in the healthcare systems.

DNP3 and T101—BLOCK DIAGRAM

DNP3 MONITORING SYSTEM and Interface with HMI (Fig. 4):

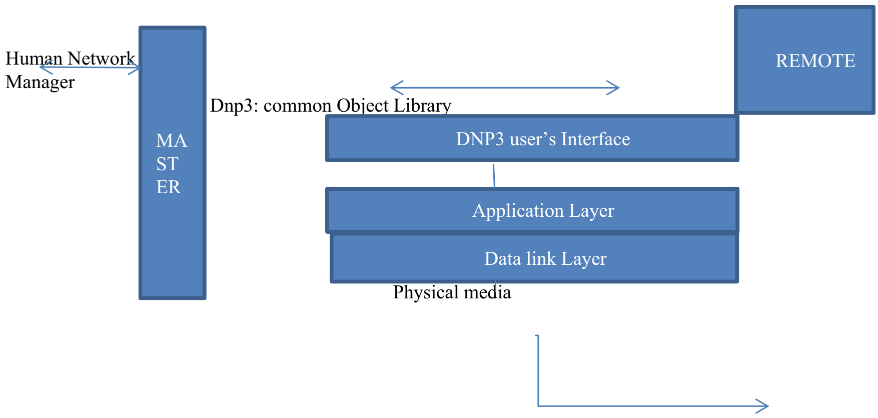


Fig. 4 DNP3 protocol layer

Comparison of DNP3 and T101

	DNP3	T101
Organization	DNP user's group	IEC TC 57 WG 03
Standard	Open industry specification	IEC standard
Architecture	Support 4-layer architecture and OSI and TCP/IP model	3-layer EPA architecture
Main coverage	Application layer (services and protocol)	Application layer (services and protocol)
Device addressing	Link based on source and destination address both 16 bits in size	Link address could be 0,1,2 bytes Unbalanced link contains slave address Balanced link is p2p
Parameter setting control	Dynamic and secure controls	Few
Dominant market	North America	Europe
Open for other encoding solutions	Yes, like XML	NO

RTU, PLC—RTU, often known as a remote telemetry or remote-control unit, is an acronym for remote terminal unit. A RTU is a microprocessor-based device that links to Supervisory Control and Data Acquisition (SCADA) or plant control systems and controls and monitors field devices.

INFERENCE OF BLOCK DIAGRAM—DNP3 and T101—communicate with other sensing objects when data is collected, cyber-security framework is also attached, and hence, data is secure, authentic, and authentic.

Limitations: SCADA is an expensive system for projects since the platform used in Windows interface. This can face cyber-attacks and threats. The system also demands regular maintenance and is costly.

2 Discussion

In order to ensure to provide functional benefits from the SCADA host to the instrumentation, not only in terms of monitoring data but also in terms of engineering, implementing, operating, and maintaining these assets, modern SCADA systems are driven to create and implement open standards using best practices established by open groups rather than a single manufacturing body to facilitate the integration of assets inside a SCADA system. As a result, the cost of owning SCADA will go down.

In healthcare systems, the info is either with hospital or with patient. The continuum of data analysis had loopholes since doctors were not continuously having the real-time data of the patient round the clock. Patient's data therefore was volatile. This delays the curing process of the patients. With SCADA, there are many learnings to the doctors and the patients. This helps in fast recovery of the patient. Compliance is not thoroughly followed, since hospitals had scattered data. SCADA is the way out to follow the compliance appropriately and can maintain all data at one point that is authentic, can be retrieved at any point and is secured data.

SCADA system has various protocols like DNP3 and T101 and there could be more protocols that can be introduced. Natural calamities like Pandemic can be controlled with centralized systems like SCADA. Health sector can therefore prove to be efficient in maintaining the data of patient healthcare systems. SCADA can help in curtailing such natural calamities and save many lives in the world. Customer's ID can be produced and can remain static across the Globe in the healthcare systems.

Cyber-attacks were common in monolithic system. SCADA curtails these attacks through cyber-security framework to maintain HITRUS, which as well helps to maintain the health compliance. Public health info—SCADA systems can be used to monitor and manage medical equipment, healthcare systems, diagnostic facilities, and laboratories, gathering and analyzing real-time patient data. SCADA systems can be used to remotely manage and access medical equipment. Asset health data are all the systems that can be protected with SCADA. Real-time data is accessible through this.

3 Conclusion and Future Scope

SCADA is fiction turned into reality. Subsequent to the progress that SCADA has made till day, the road ahead of human–machine interface includes the following:

Keyboards with LED projections, eye moment and eye tracking through human–machine interface. Displays of control room, Realistic picture of the remote tools and machines relating to data equipped with technology. The other potential domains comprise of drones at domestic level vehicles with no drivers for perennial surveillance, leak detection through pipelines, and other infrastructures. SCADA experts suggest extensive and rampant use of SCADA, though many industries yet take the conservative approach to SCADA. Cutting down on costs and saving of costs is one of the major factors that drive the use of SCADA to a great extent. Upcoming generation SCADA/HMI.

References

1. Tariq N, Asim M, Khan FA (2019) Securing SCADA-based critical infrastructures: challenges and open issues. Procedia Comput Sci 155:612–617 [Google Scholar]
2. Alimi AM, Ouahada K, Abu-Mahfouz AM (2020) A review of machine learning approaches to power system security and stability. IEEE Access 8:113512–113531
3. Samdarshi R, Sinha N, Tripathi P (2015) A triple layer intrusion detection system for SCADA security of electric utility. In: Proceedings of the 2015 annual IEEE India conference (INDICON), New Delhi, India, pp 1–5
4. IBM (2020) IBM report: compromised employee accounts led to most expensive data breaches over past year. https://newsroom.ibm.com/2020-07-29-IBM-Report-Compromised-Employee-Accounts-Led-to-Most-Expensive-Data-Breaches-Over-Past-Year
5. Cifranic N, Hallman RA, Romero-Mariona J, Souza B, Calton T, Coca G (2020) Decepti-SCADA: a cyber deception framework for active defense of networked critical infrastructures. Internet Things 12:100320 [Google Scholar]
6. Upadhyay D, Sampalli S (2020) SCADA (supervisory control and data acquisition) systems: vulnerability assessment and security recommendations. Comput Secur 89:101666 [Google Scholar]
7. Phillips B, Gamess E, Krishnaprasad S (2020) An evaluation of machine learning-based anomaly detection in a SCADA system using the modbus protocol. In: Proceedings of the 2020 ACM southeast conference, Tampa, FL, USA, pp 188–196 [Google Scholar]
8. Ahmed M, Anwar A, Mahmood AN, Shah Z, Maher MJ (2015) An investigation of performance analysis of anomaly detection techniques for big data in SCADA systems. EAI Endorsed Trans Ind Netw Intell Syst 2:5 [Google Scholar] [Green Version]
9. Rakas SVB, Stojanovic MD, Markovic-Petrovic JD (2020) A review of research work on network-based SCADA intrusion detection systems. IEEE Access 8:93083–93108
10. Alimi AM, Ouahada K (2018) Security assessment of the smart grid: a review focusing on the NAN architecture. In: Proceedings of the 2018 IEEE 7th international conference on adaptive science and technology (ICAST), Accra, Ghana, pp 1–8

Data Security in a Cloud Environment Using Cryptographic Mechanisms

Shailesh Gahane, Reshma Pohankar, Komal Ugemuge, and Dnyaeshwari Nakhate

Abstract Internet-based host service delivery is referred to as "cloud computing". The essential necessity of such applications, where data are being exchanged or transmitted between the servers and their users, is the secure data storage on cloud environments. In this research, we use cloud-based cryptography, one of the best methods for secure communication. Writing or generating codes to allow for information confidentiality is known as cryptography. This research proposed a method of plain text and key encryption to protect data security in cloud computing. The suggested system offers data concealing and information-theoretic security. Due to its benefits, including cost-savings, resource sharing, widespread network access and simplicity of management, cloud computing has become quite popular among various enterprises. By maximising shared use, it improves the capabilities of physical resources. In a shared environment where several customers are joined together, the valuable objects of the clients (data and applications) are transferred outside of the purview of regulatory oversight. Unfortunately, this procedure raises security issues like the theft of confidential information and the disclosure of personally identifying information. By creating a range of technologies to safeguard cloud data, including encryption, several academics have helped to lessen the issue of data security in cloud computing.

Keywords Cloud computing · Internet · Cryptography · Encryption · Information security

S. Gahane (✉) · R. Pohankar · K. Ugemuge · D. Nakhate
Faculty of Sciences and Technology, School of Allied Sciences, Datta Meghe Institute of Higher Education and Research, Wardha, India
e-mail: shailesh.gahane@dmiher.edu.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_11

1 Introduction

Advanced computing is now more accessible to both users and service providers thanks to a new technology called cloud computing. The user will always have access to the data they have provided by using the cloud. In this research, we use cloud-based cryptography, one of the best methods for secure communication. To protect data utilised or stored in the cloud, cryptography employs encryption techniques. This method of communication with a server is most frequently utilised by client programmes. Web browsers and web servers are two examples, as are email clients and email servers. The best example is web encryption, which enables us to choose between an unencrypted or encrypted version of a website by changing the URL from HTTP to HTTPS. Like any data that is encrypted by cloud providers, it transforms data into an unreadable format for an unauthorised user while simultaneously allowing users to access shared cloud services simply and securely [1].

The main applications of cryptography are the following four essential information security services.

- A security service called **confidentiality** guards against unauthorised access to data.
- **Data integrity** is the responsibility of the security service, which must detect any alterations to the data.
- **Authentication** assures the recipient that the received data has only come from a certain, verified sender.
- **Non-repudiation** is legal concept ensures that a party cannot claim ownership of a previous undertaking or activity.

2 Research Background

The cloud computing describes the remote provision of computing services, much as how the majority of users obtain power, water and other utilities. This paradigm enables unified, realistic, whenever needed the on-demand network access to a shared pool of reconfigurable computing devices or resources such as networks, servers and storage. Cloud computing refers to both services and its applications. It enables prompt provisioning and release of those resources without the need for management effort or service provider intervention [2].

Cryptography has advanced more quickly as a result of the greater value placed on knowledge and, more recently, data. The oldest known encryption, which had been used to protect religious or commercial knowledge, was then affected by the need for secret military communication also [3].

The following important improvements were brought about by the necessity for secure commercial and private communication. Since the 1980s, computers and the Internet have gradually become a part of every aspect of our lives. Therefore, it is best for all communications in the Information Age to be encrypted, whether they are between humans and machines or between humans and machines [4].

3 Cloud Deployment

Private, public and hybrid clouds are the three categories under which cloud computing is divided. Private clouds are solely controlled and supervised for a single business, and the resources are not utilised by other clients, indicating that unauthorised individuals secure them from access. Organisations and the other public can use public clouds. Each client receives a portion of the assets. Depending on the service received and the resources used, the customers pay the cloud owner. Cloud service providers are in charge of managing the physical infrastructure, which is far from the clients. Public and private clouds are combination to become a hybrid clouds [5].

Similarly the services, the three main services that cloud computing offers are:

1. **Platform as a Service (PaaS),**
2. **Infrastructure as a Service (IaaS) and**
3. **Software as a Service (SaaS).**

IaaS: This term describes the CSP's hardware, which includes networks, storage, memory, CPUs and a variety of other computer resources Systems that have been virtualised and are accessible online serve as the resources. The CSP is in charge of the necessary resources. **PaaS**: It offers operating systems, platform layer resources, middleware, integrated development environments and other resources through a third-party supplier who offers consumers online access to hardware and software tools. PaaS simply gives clients control over the programmes that are migrated to the cloud; it does not offer them control over the underlying cloud infrastructure [6].

SaaS: It enables users to access applications as a service online. Users do not need to purchase, install or maintain software to access it; they can do so just using the internet. Instead of purchasing the software outright, customers pay for consumption. The front end and the back end of the cloud computing systems are separated into two parts.

A website: The side that cloud clients see is called the front end [7].

Customers are generally unaware of the back-end component, which consists of the network connection, cloud servers and their applications.

4 Service Models

From simple computing activities to the delivery of complex applications, cloud computing can offer a variety of services. Although there are numerous methods to classify these services. According to the National Institute of Standards and Technology (NIST) definition, the usages of following three fundamental service models are:

Service-based software: Instead of utilising software that is installed on a local desktop or server, customers that utilise the SaaS model use programmes that the supplier makes available and remotely accessible on demand. In terms of comprehension and accessibility, the SaaS service paradigm is the most user-friendly. The bulk of the time, SaaS applications are accessed by "thin clients" of hardware or software [8].

The examples include online storage options like Dropbox and web-based services like Google Applications.

Software as a Platform: Clients use the tools, including programming languages, provided by the PaaS provider to build applications on the infrastructure. An illustration of one of these programmes is Facebook. To aid in the creation, testing and release of web applications, this kind of platform may offer hosting alternatives and development tools. While the user is in charge of the platform's upgrades, networks, servers and other supporting infrastructure, the supplier oversees and maintains the underpinning infrastructure [9].

Service-based infrastructure: Customers can utilise the basic computing resources provided by IaaS providers anyway they see fit. Customers have the same ability to install, run and manage any operating systems and software they want, just like with desktop personal computers or local servers. The provider maintains the cloud's supporting infrastructure.

For example, Microsoft Azure and Amazon Web Services are two IaaS Examples. They are maintaining their infrastructure (Fig. 1).

5 Cryptography in Cloud Computing

In real life, computer applications are expanding daily. Data security is thus becoming a more and more crucial component of message or data transmission. Information security has been ingrained in our daily lives. Among the different approaches used in information security, hidden information exchange is a worry [10]. This has been accomplished using a variety of techniques, including steganography, coding, etc.

Cloud cryptography secures data that is utilised or stored there by using encryption methods. Like all data that is encrypted by cloud providers, it converts data into an unreadable format for unauthorised users while still enabling users to easily and

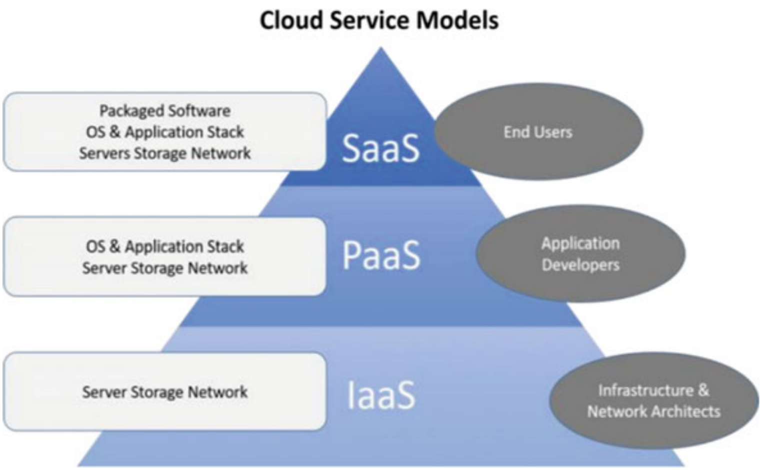
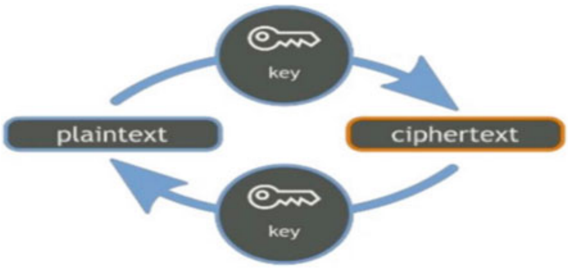


Fig. 1 Cloud computing service model structure

Fig. 2 Structure of cryptography



securely use shared cloud services. Cloud cryptography protects responsive data without necessitating a data transfer.

Greek terms "kryptos" (which means "secret") and "graphein" (which means "writing") combine to get the word "cryptography", which means "hidden writing" [2].

Previously, the terms "cryptography" and "encryption" were interchangeable, but today's cryptography is mostly focused on mathematical theory and computer science methods.

Many different applications, including e-commerce, computer passwords and financial transaction cards, utilise cryptography [11] (Fig. 2).

Suggested Process of Cryptography

On Sender Side:

Step 1: Select the Plaintext and Key from Source Node
Step 2: Convert the message in encoded format using Encryption Algorithm
Step 3: Encoded message is passing from Interceptor/ciphertext

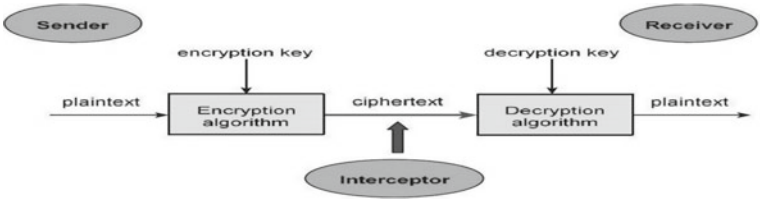


Fig. 3 Process of cryptography

On Receiver Side:

Step 4: Encoded format reached in Destination Location
Step 5: Decoding will be done on encoded format (Plaintext and Key) by using Decryption Algorithm
Step 6: Original Message will be on Receiver Side (Fig. 3).

6 Literature Survey

Author and title	Finding	Remark
Fairosebanu, A. A., and Jebaseeli, A. C. N. (2023). Data security in cloud environment using cryptographic mechanism. Bulletin of Electrical Engineering and Informatics, 12(1), 462–471 [2]	Software, platform and infrastructure components make up the cloud resources	Through a process known as cloud computing, users can access an infinite amount of computing power
Kumari, A. B., and Iyengar, N. C. S. (2023). One Time Pad Encryption Technique in Cryptography. *International Journal of Computational Learning and Intelligence*, *2*(1), 1–7 [12]	In this research, we use cloud-based cryptography, one of the best methods for secure communication	Due to this, cloud computing has evolved into a technological advancement that can manage massive volumes of data that are transported and stored via electronic applications
Alemami, Y., Al-Ghonmein, A. M., Al-Moghrabi, K. G., and Mohamed, M. A. (2023). Cloud data security and various cryptographic algorithms. *International Journal of Electrical and Computer Engineering*, *13*(2), 1867 [13]	Due to its capacity to offer individuals and organisations with access to massive resources, cloud computing (CC) technology has become quite popular	In order to protect data, applications and cloud infrastructure from internal and external threats, cloud service providers (CSPs) must take the necessary precautions

(continued)

(continued)

Author and title	Finding	Remark
Clustering, F. C. M. DESIGN OF ULTRA LOW POWER MULTIPLIERSUSING 49 HYBRID ADDERS [14]	Some businesses encrypt their important data before moving it to the cloud after becoming aware of these security issues	This adds an additional layer of security from the client's perspective for their data in transit
Bhargav, A. J. S., and Manhar, A. (2020). A review on cryptography in cloud computing. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, *6*(6), 225–230 [15]	The low cost, enhanced storage and flexibility of cloud computing are its most significant benefits	Computing services include things like servers, storage, databases, networking and software, to name a few

According to the literature survey, we have find out the failures including key leakage, software bugs, holes in operating systems, side-channel attacks, phishing attacks and social engineering.

7 Research Methodology

The data can be rendered useless using cryptography to prevent reading by unauthorised parties. Clients can access their data or applications anytime, anyplace and on any device with an Internet connection by storing it in the cloud or on the Internet. It offers on demand access and lessens the need for sophisticated technology, making it incredibly practical for users.

By using cryptography is one way to ensure cloud security. Cryptography is one technique for data protection. Unauthorised parties cannot the render data without encryption [16].

Cloud-based services are widely utilised and embraced on a global scale. Utilising the Internet for online applications and data storage has become essential for commercial, industrial and residential organisations in the modern world [17].

Different algorithms can be used to encrypt the data. The data formats, as well as its appearance and feel, are destroyed, and the padding may cause the data's length to grow. The extra padding is removed during decryption, and the encrypted data is decoded to restore it to its original state.

8 Conclusion

This study evaluated and reviewed several different cryptographic methods used in cloud computing, including those employed by some of the industry's biggest firms. It was suggested and explored to utilisation of new algorithms to encrypt data as it is transferred from the cloud user to the cloud provider's platform. Apart from this, we have suggested the four main applications of cryptography that is confidentiality, data integrity, authentication and non-repudiation that are essential for information security services and suggested process of cryptography.

References

1. Kuang R, Barbeau M (2022) Quantum permutation pad for universal quantum-safe cryptography. Quantum Inf Process 21(6):211
2. Fairosebanu AA, Jebaseeli ACN (2023) Data security in cloud environment using cryptographic mechanism. Bull Electr Eng Inform 12(1):462–471
3. Basha SM, Ahmed ST, Naidu SKM (2022) Artificial intelligence: practical approach, vol 1. MileStone Research Publications
4. Prakash GL, Prateek M, Singh I (2014) Data encryption and decryption algorithms using key rotations for data security in cloud system. In: 2014 International conference on signal propagation and computer technology (ICSPCT 2014). IEEE, pp 624–629
5. Albugmi A, Alassafi MO, Walters R, Wills G (2016) Data security in cloud computing. In: 2016 Fifth international conference on future generation communication technologies (FGCT). IEEE, pp 55–59
6. Selvanayagam J, Singh A, Michael J, Jeswani J (2018) Secure file storage on cloud using cryptography. Int Res J Eng Technol 5(3):2044
7. Deepa B, Maheswari V (2022) An enhanced DNA structure for one-time pad together with graph labeling techniques. AIP Conf Proc 2385(1). AIP Publishing
8. Wang X, Wang Y, Zhu X, Luo C (2020) A novel chaotic algorithm for image encryption utilizing one-time pad based on pixel level and DNA level. Opt Lasers Eng 125:105851
9. Ahmed ST, Basha SM (2022) Information and communication theory-source coding techniques-part II. MileStone Research Publications
10. Kumar M, Sharma SC, Goel S, Mishra SK, Husain A (2020) Autonomic cloud resource provisioning and scheduling using meta-heuristic algorithm. Neural Comput Appl 32:18285–18303. Khakim L, Mukhlisin M, Suharjono A (2020) Security system design for cloud computing by using the combination of AES256 and MD5 algorithm. IOP Conf Ser: Mater Sci Eng 732(1). https://doi.org/10.1088/1757-899X/732/1/012044
11. Dubey K, Sharma SC, Kumar M (2022) A secure IoT applications allocation framework for integrated fog-cloud environment. J Grid Comput 20(1):5
12. Kumari AB, Iyengar NCS (2023) One time pad encryption technique in cryptography. Int J Comput Learn Intell 2(1):1–7
13. Alemami Y, Al-Ghonmein AM, Al-Moghrabi KG, Mohamed MA (2023) Cloud data security and various cryptographic algorithms. Int J Electr Comput Eng 13(2):1867
14. Clustering FCM (2015) Design of ultra low power multipliers using 49 hybrid adders
15. Bhargav AJS, Manhar A (2020) A review on cryptography in cloud computing. Int J Sci Res Comput Sci Eng Inf Technol 6(6):225–230
16. Orobosade A, Aderonke T, Boniface A, Gabriel AJ (2020) Cloud application security using hybrid encryption. Communications 7:25–31
17. Zhang L, Xiong H, Huang Q, Li J, Choo KKR, Li J (2019) Cryptographic solutions for cloud storage: challenges and research opportunities. IEEE Trans Serv Comput 15(1):567–587

Secure Data Education: Leveraging Big Data for Enhanced Academic Performance and Student Success in Educational Institutions

A. R. Shylaja, D. A. Shubhashree, M. R. Shrihari, T. N. Manjunath, and N. Ajay

Abstract Big data applications are employed in a variety of industries, including academia, healthcare, marketing, industry, government, and public administration, among others, to address societal issues and progress society (Munshi and Alhindi in Big data platform for educational analytics, 2021). Big data in education has significant benefits since it can be managed at an infinite scale. Big data in education helps students achieve better grades, enhance the grading system, gain attention, and lower dropout rates (Wang and Yang in 2020 International conference on communications, information system and computer engineering (CISCE), 2020). The evaluation of student performance is based on their results on tests, evaluations, quizzes, projects, etc. By analyzing this data, mentors can recognize student behavior and performance and establish the best learning environment. The feedback is highly useful for enhancing the outcomes. Students' heart rates, facial characteristics, and behavioral clues may all be seen visually. Educational despite the number of students in institutions, big data makes it possible to create a specific curriculum for each student. With no compromises, this enables students to enroll in the classes they are interested in and progress at their own speed. All of this data on students' grades, tardiness in submitting tasks, and desire in learning aid mentors and institutions in pinpointing the precise causes of dropouts. The security and privacy of data created in universities, however, are the major drawbacks of big data applications in the educational industry. In this paper, we investigate how Hadoop employs Kerberos for authentication and authorization as well as how big data benefits students' academic performance. A network authentication system called Kerberos offers a safe authentication service based on reliable outside sources.

Keywords Big data · Education · Security · Privacy · Kerberos · Authentication · Authorization

A. R. Shylaja (✉) · D. A. Shubhashree · M. R. Shrihari · N. Ajay
S J C Institute of Technology/Computer Science and Technology, Chickballapur, India
e-mail: nirmalashyalaja@gmail.com

T. N. Manjunath
BMSIT&M/Computer Science and Technology, Bengaluru, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_12

1 Introduction

Data is being produced in massive amounts from a variety of flexible sources. There is no organized framework for the data that originates from sources like the Internet, mobile devices, institutions, etc. We need specialized tools and procedures to analyze this sort of data since conventional computer systems are insufficient for handling and analyzing enormous volumes of data in a variety of forms. We are now processing and analyzing enormous volumes of data using big data analytics. With the use of big data analytics in the educational sector, teachers may make use of learning analytics to identify students' areas of weakness and try to solve the problems with more contemporary teaching methods that can improve students [1] (Fig. 1).

Huge data collections that can't be handled by traditional computer techniques are referred to as "big data". Big data refers to the vast quantities of information that academic institutions produce rapidly and in a variety of formats [2]. The six v's, or characteristics of big data, are volume, veracity, velocity, variety, value, and variability [3].

Volume: A defining characteristic of each dataset is its size. The term "volume" refers to the amount of data that is generated and kept in a big data system. Between petabytes and exabytes of data are what we're talking about. An advanced processing system that is much more powerful than a typical laptop or desktop processor must be used.

Velocity: Depending on how quickly data accumulates, it can either be categorized as huge data or as regular data. Systems need to be able to handle the amount and rate of data creation since a lot of this data needs to be evaluated instantly. Data processing speed implies that there will always be more data available than there was previously, but it also implies that data processing speed needs to be just as quick.

Value: It is also crucial to consider the value of something. The amount of data we preserve or process is important, but it is not the only factor. Saving, analyzing, and evaluating reliable and meaningful data are also essential for gaining insights.

Fig. 1 Big data characteristics

Veracity: Veracity is a phrase used to indicate how trustworthy and high-quality the data is. Even if the data is not reliable and/or trustworthy, the value of big data remains indisputable. When using frequently updated data, this is particularly true. Because of this, there must be checks and balances throughout the entire process of collecting and analyzing big data.

Variability: How equally distributed a set of data is referred to as data dispersion, also known as spread or dispersion. Users can utilize statistics to compare their data to that of other users using a variation to show how much their data differs from other datasets.

Hadoop offers great potential for efficiently and economically identifying patterns and correlations in enormous datasets as compared to using conventional relational databases. Hadoop makes use of a platform for distributed computing [3]. A widely used hardware architecture that is easy to construct and implement is used by Hadoop. Hadoop is an Apache framework that enables massive data to be processed and stored in parallel and over a distributed network. The Hadoop Distributed File System (HDFS) is used to handle and store massive volumes of data in a variety of formats, including structured, semi-structured, and unstructured data [3]. Using Map Reduce techniques, large amounts of data are gathered and reduced to smaller results. Because this method speeds up data processing, aids in parallel execution and can analyze any data type, it is becoming increasingly important to employ big data in the education sector. Making decisions is crucial for every business if it wants to increase its performance. Big data analytics is utilized in educational institutions to examine the data contained in HDFS and make better decisions to enhance institutional performance [4].

Hadoop is made up of two key parts: Hadoop's storage component, Hadoop Distributed File System (HDFS), enables the storing of substantial volumes of data over numerous servers. It is affordable since it is made to function with common hardware. The resource management part of Hadoop, Yet Another Resource Negotiator (YARN), controls the distribution of resources (such as CPU and memory) for processing the data stored in HDFS. Other modules of Hadoop, such as Hive (a query language like SQL), Pig (a high-level platform for developing Map Reduce programmers), and HBase, give more capability (a non-relational, distributed database). Big data applications including data warehousing, business intelligence, and machine learning frequently employ Hadoop. Data processing, data analysis, and data mining are further uses for it. With a straightforward programming model, it enables the distributed processing of massive datasets over clusters of computers.

2 Problem Statement

The educational sector today makes substantial use of big data analytics to gather enormous amounts of data for secure analysis, evaluation, visualization, and decision-making of student performance. The suggested solution will create a web application utilizing Map Reduce for effective performance and security using Kerberos based on the big data Hadoop framework. The major goal of the proposed approach is to raise the institution's overall performance, whether for the benefit of the students or for the institution. Using this solution would help educational institutions expand, encourage data-driven decision-making, and enhance the educational experience for students.

3 Methodology

This section presents the implementation of big data analytics in analyzing the performance of educational institutions. The implementation includes 3 dashboards.

(1) Student dashboard
(2) Teacher dashboard
(3) Institutional dashboard

These three dashboards provide insights into different aspects of the educational institution and help in making data-driven decisions [5].

The illustration of big data analytics implementation in an educational institution can be seen in the Fig. 2.

In the Student dashboard,

- Student results, attendance, and other factors are displayed in the form of a graphical representation

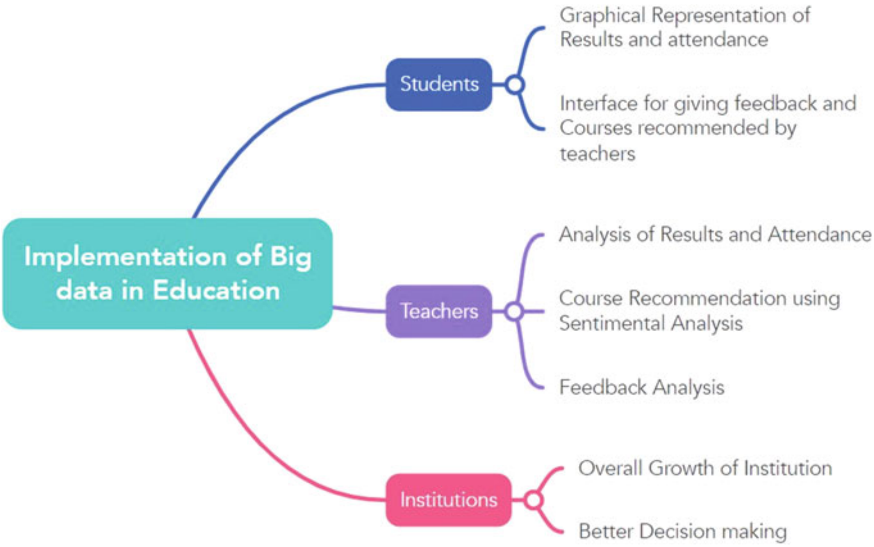


Fig. 2 Implementation of big data analytics in education

- Students can also provide their feedback on the teaching process. Based on the feedback received, teachers can recommend the required courses to the students in whom they are lagging.

This can encourage pupils to take charge of their education and strive for better outcomes. Teachers may provide course recommendations based on student input to better understand the difficulties that students are facing [6].

In the teacher dashboard,

The analysis, evaluation, prediction, and decision-making are done based on the student data, whereas the analysis is done based on the results, attendance, and other factors. Once the evaluation is complete, the outcome is predicted based on the available data so that better decisions can be made to improve the student's performance [6].

- Analysis of results, attendance, and other factors is performed using learning analytics to improve student performance.
- Sentiment analysis is based on student feedback and enables one to understand the educational problems of the students and take necessary actions to solve these problems.

In the Institutional dashboard,

- The overall growth of the institution is displayed in the form of a graphical representation.
- Each institution's growth depends heavily on decision-making, thus better judgments are made based on the information at hand to raise the institution's overall performance [6].

The different algorithms that are implemented are described in the following section.

A. *KERBEROS*

The disadvantage of the traditional education system is the lack of security and privacy. This problem is solved by using Kerberos. Kerberos is a network authentication protocol used in big data management to authenticate and authorize the user.

To offer secure authentication for client/server applications in business settings, Kerberos is frequently utilized. With the endorsement of many software and hardware suppliers, it has emerged as the de facto standard for authentication in many businesses [7]. Kerberos protocol offers a high level of security and has been shown to be a successful method of securing network connections. To make sure that it is offering the desired level of security, it must be properly set and maintained, just like any other security protocol. Users who want to access platforms for big data education can do so securely using Kerberos. This makes it possible for educators to collaborate on projects and safely exchange information with students, instructors, and administrators. This data may be securely shared with administrators and utilized to enhance teacher preparation programs by employing Kerberos authentication.

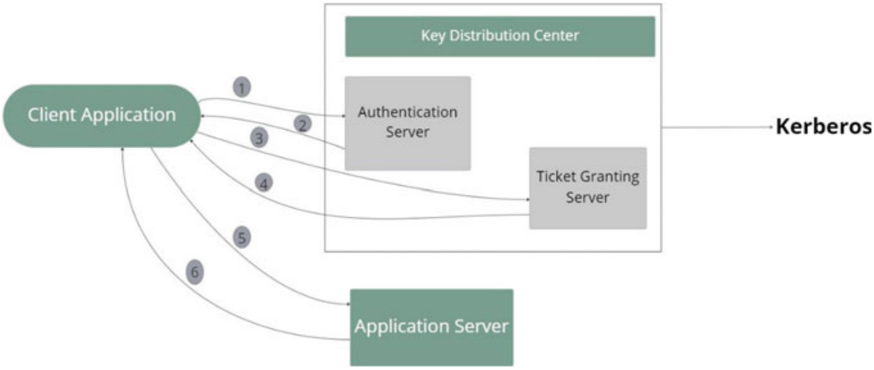


Fig. 3 Implementation of Kerberos protocol

Kerberos uses three servers to authenticate the client (Fig. 3):

- Authentication Server.
- Ticket-granting Server.
- Application Server.

1. Client request to the authentication server for a ticket with a user password.
2. The authentication server verifies the user's credentials and issues tickets by encrypting the ticket with the user's password.
3. An application server ticket is requested by the client from the ticket-granting server.
4. The ticket, which is a secret key, is delivered to an application server after the ticket-issuing server has verified its identity.
5. The client then makes a service request to the application server.
6. The application server grants the client services.

B. *MAP REDUCE TECHNIQUE*

Large educational datasets frequently need to be thoroughly cleaned and pre-processed before being utilized for the study. Cleaning and pre-processing the data may be done more quickly and effectively by dividing these jobs among a cluster of computers using Map Reduce. With vast volumes of text data, such as student comments and course ratings, Map Reduce may be utilized to do sentiment analysis. The amount of time needed to analyze the data may be greatly decreased by using Map Reduce to distribute the sentiment analysis work over a cluster of machines. Map Reduce can make it quicker and more effective to find the key themes and topics in the data by dividing the topic modeling work over a cluster of computers. In organizations, a vast volume of data is produced each day and is stored as data blocks on the cluster's many nodes. As HDFS utilizes standard hardware to store the data, it keeps three copies of the data blocks. When a request is made by a user, simultaneous processing will start across the data nodes. This parallel processing is completed via the Map Reduce method. Analyses of homework completion, exam scores, and attendance may be requested. Map Reduce is a programming paradigm used in big data processing. The two core elements of this technique are the Map and Reduce tasks. As records are inputted into the Map task, they are transformed into new records in the form of key-value pairs. The key-value pairs generated by the Map tasks are sent into the Reduce tasks, which combine them to produce a smaller group of tuples. As implied by the name, the Map chores usually come after the Reduce activities.

Map task's input and output are as follows:

Input: <k1, v1>

Output: list (k2, v2)

The Reduce job has the following input and output:

Input: < k2, list (v2)>

Output: (k3, v3)

For processing huge data in education, Map Reduce is an effective method. Map Reduce may drastically reduce the time and resources needed to analyze and glean insights from massive educational datasets by dividing data processing operations over a cluster of computers (Fig. 4).

Academic Analysis

As seen in the above image, USN, branch, Name, and Phone No. are the inputs provided for the academic study. When a user requests information about a specific student, academic analysis is carried out utilizing Map Reduce tasks on the data stored in the HDFS to produce that student's specific information from a list of student details (Fig. 5).

Result Analysis

According to the image above, the input file contains all the student results, and the Map Reduce approach is used to assess the results rate based on the years. When the rate of student success declines, this analysis aids institutions in making wiser decisions (Fig. 6).

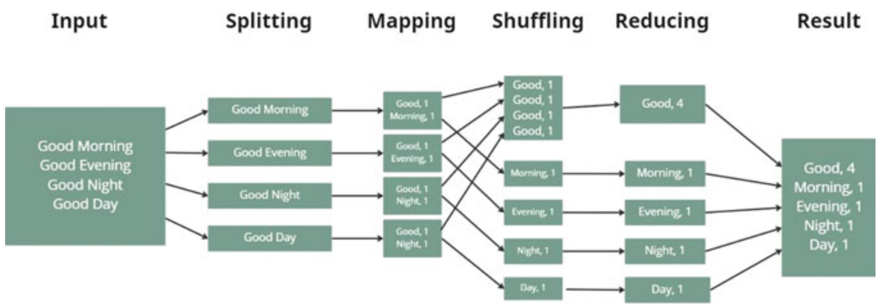


Fig. 4 Map reduce technique

Fig. 5 Academic analysis

Fig. 6 Result analysis

Attendance Analysis

As seen in the picture, the input provides a list of every student's attendance. Using the Map Reduce technique, the average attendance is computed, and the students are then divided into groups according to whether their attendance is below average or above average (Fig. 7).

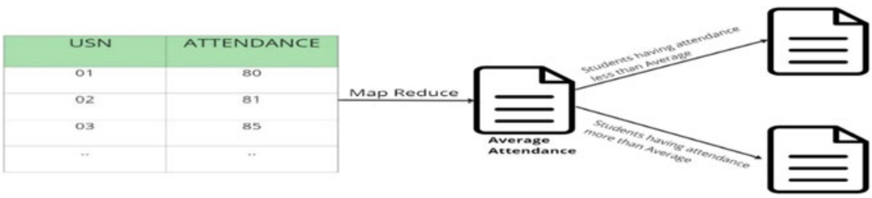


Fig. 7 Attendance analysis

C. *COLLABORATIVE FILTERING*

Big data education may use collaborative filtering to provide students with tailored suggestions for courses, books, and other learning materials. Collaborative filtering algorithms can forecast which courses and materials a student is most likely to find engaging and helpful by examining data on their behavior and interests as well as data on the subject matter and structure of the courses.

The creation of adaptive learning systems is one use case for collaborative filtering in big data education. To tailor the learning experience for each student, these systems make use of information on student performance and behavior [8]. For instance, based on a student's prior performance and interests, an adaptive learning system can suggest various exercises or activities to various pupils. Collaborative filtering uses machine learning algorithms such as random forest and decision tree models to effectively recommend courses by analyzing the similarities between students, classifying them into discrete clusters, and recommending the required courses based on the likes and dislikes of similar students. The implementation of collaborative filtering in education systems leads to remarkable growth in improving the quality of education (Fig. 8).

The development of instructional material is another area where collaborative filtering is used in big data education. Teachers and content producers can find areas where current instructional materials fall short of students' requirements and interests by evaluating data on student behavior and preferences. This allows for the development of new resources that better suit students' interests. The use of collaborative filtering in big data education has the potential to increase the efficacy of educational programs and give students access to more individualized and interesting learning opportunities. To make sure that collaborative filtering is utilized in an ethical and responsible manner, it is crucial to address problems like prejudice and data privacy.

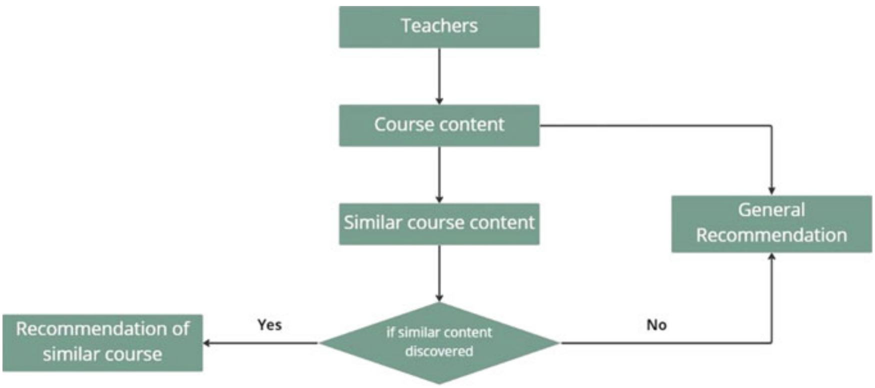


Fig. 8 Collaborative filtering

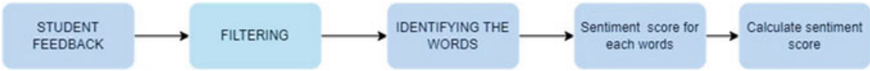


Fig. 9 Working of VADER

Recommendations can be given to groups of students who have similar tastes and interests using collaborative filtering. Collaborative filtering algorithms can forecast which courses and materials will likely be popular among each group by examining data on the behavior and interests of various groups. Collaborative filtering algorithms can identify resources that are expected to be in high demand and allocate them in accordance with assessing data on resource consumption and availability.

D. *VADER*

The Valence Aware Dictionary and sentiment Reasoner (VADER) model can be used to assess the student's emotional polarity and extent (intensity). By the utilization of human resources and the wisdom of crowds, it employs a human-centered methodology that blends qualitative analysis and empirical validation. It determines the sentimental worth of the text that the learner provides as input. The emotion value for each word is determined using the student comments as input. Unwanted words are then filtered out of the phrase by identifying them for further processing. The words in the statement are analyzed using natural language processing. Mentors and educational institutions can identify students' behavior and emotions with the aid of this procedure (Fig. 9).

When VADER analyses a text, it employs both rule-based heuristics and machine learning techniques to produce a sentiment score that runs from −1 to +1 [9]. Together with the emotion score, VADER offers ratings for a few additional sentiment-related characteristics, including positivity, negativity, and neutrality.

The study of political discourse, consumer feedback, and social media monitoring are just a few of the numerous applications in that VADER has been widely employed. VADER can be used in the field of education to assess the opinions and feelings of learners on their experiences with courses, instructors, and other learning materials. Many student feedback datasets, including survey and course assessment results, may be analyzed using VADER. Insights on student views regarding courses, teachers, and other educational resources may be gained by using VADER to this data, which enables educators to discern the sentiment of student input rapidly and effectively. VADER is a tool that may be used to keep an eye out for mentions of educational institutions, programs, and subjects on social media platforms. Educators may learn more about how the public views their institution or programs by studying the sentiment of these references and can also spot areas where they might improve.

VADER may be utilized in e-learning systems to assess sentiment in online discussion forums. Education professionals may learn more about student involvement and pinpoint areas that may require more help by examining the tone of these exchanges. Individual student reactions to individualized learning programs may be analyzed using VADER. Insights about which subjects or instructional strategies are particularly successful or interesting for each student may be gained by examining the sentiment of these replies, which helps instructors further individualize the learning process. In big data education, VADER is a potent sentiment analysis tool that may be applied in several different ways.

4 Results and Discussion

Digital learning is currently being used, although there are several amenities that are lacking. These elements will be introduced to the program to give those facilities, including ideal teacher assistance, organizational support, a solid learning platform, and protection of organizational and student data. Physically keeping student information and organizational records cannot be recovered once they are lost, but the additional capabilities of Hadoop storage will offer backup options if they are, and providing security for each organization will lessen the chance of cyber theft of sensitive data that may result in unlawful activities (Table 1).

Hadoop has the benefit of having quick processing times for massive datasets. When dealing with complicated queries, RDBMS might be sluggish to handle big volumes of data. As it relies on a number of variables, including the quantity of the data, the complexity of the queries, and the hardware design, it is challenging to estimate the processing speed of RDBMS in terabytes per second (TB/s) (Fig. 10).

On the basis of various benchmarks and industry norms, we may nonetheless estimate the processing speed. The maximum amount of data that an RDBMS can handle in one hour is 100 GB or around 0.03 GB/s. It will process more quickly in Hadoop than in RDBMS [10].

Big data settings may safeguard access to critical data and resources, including educational data, by utilizing the well-known authentication system Kerberos. For administrators, instructors, and students who need to access educational data and analytics systems, Kerberos can offer secure authentication. A Hadoop cluster used for studying educational data might have access controlled by Kerberos. Kerberos authentication may be used to verify the identity of instructors, administrators, and students, limiting access to sensitive information and resources to those who are allowed. Access to educational software and resources that are housed on a server

Table 1 Processing speed of RDBMS and big data

Data size	Processing speed in RDMS	Processing speed in big data
1GB	0.1	0.01
100GB	1	0.1
1TB	13	3
100TB	48	11
1PB	90	15

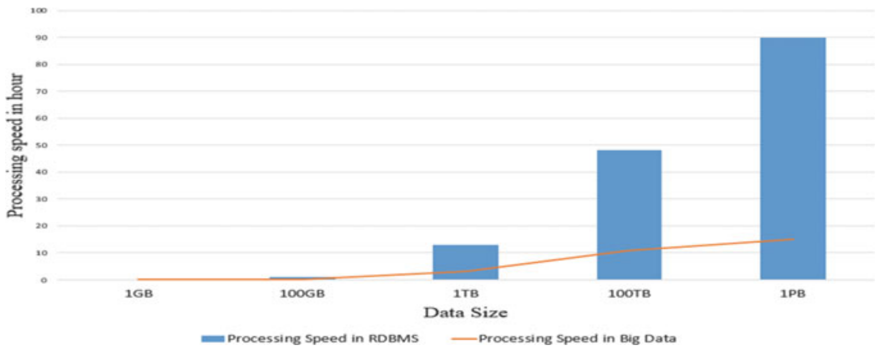


Fig. 10 Comparison of the processing speed of RDBMS and big data

can be secured using Kerberos. Organizations may verify that only authorized users have access to these tools and apps, as well as prevent unauthorized access and data breaches, by requiring users to authenticate using the Kerberos protocol. Kerberos may be helpful in protecting sensitive educational data and safeguarding big data education environments. To make sure that Kerberos is offering the required level of security, however, and to make sure that its use complies with best practices and privacy laws, it must be properly configured and maintained.

Big data education may make use of VADER, a potent sentiment analysis tool, in several different ways. Education professionals may learn more about student attitudes and preferences, spot areas for development, and offer more specialized assistance for students' achievement by evaluating sentiment data.

Collaborative filtering should be used in conjunction with other techniques for assessing educational data; it is crucial to remember that it is not a cure-all. As collaborative filtering depends on the analysis of private student data, it's also critical to make sure that these issues are effectively addressed. The usage of collaborative filtering in big data education is expected to grow as educational institutions continue to implement big data analytics to enhance student outcomes and foster student achievement. The proposed system has the ability to boost educational institutions' general performance by using big data analytics to make data-driven choices, while simultaneously preserving the confidentiality and privacy of sensitive educational data.

5 Conclusion

The strategies that may be used to increase student performance are described, and big data applications and technologies can be used more effectively in education. Collaborative filtering for student course recommendation systems, VADER for student sentiment analysis based on their feedback, Map Reduce for data processing and analysis, and Kerberos for secure authentication processing are some of the approaches mentioned in this article. The major purpose of these suggested big data strategies is to raise students' performance in terms of final results. Big data apps are being used by universities and other educational institutions to manage student data in order to monitor student behavior and performance and make better curricular decisions. Students' input is sought by educational institution administration in an effort to enhance the curriculum. Using the Kerberos authentication protocol helps solve the main security and privacy issue with big data applications. Big data has the potential to transform the area of education by offering insights that can enhance learning, and increase student outcomes, and overall performance. To provide student privacy and guarantee that data is used for the benefit of students and educators, it is crucial to use these insights in a responsible and ethical manner. Big data may be used to assess student performance data and find out what sections of the curriculum and teaching strategies might need to be changed in order to better meet the requirements of the students. Big data analysis may be used to forecast student results by looking at previous data and seeing trends that can hint at future success or failure. By using this data, at-risk pupils may receive targeted help early on in the intervention process. Big data may be used to tailor educational experiences by student data to pinpoint unique strengths and shortcomings. By gathering this data, learning programs that are specifically tailored to each student's requirements may be made. By building a web application based on the big data Hadoop framework using Map Reduce and incorporating Kerberos for secure authentication, educational institutions can leverage big data analytics to improve their overall performance.

References

1. Tasmin R, Muhammad RN, Nor Aziati AH (2020) Big data analytics applicability in higher learning educational system. In: International conference on technology, engineering, and sciences (ICTES)
2. Munshi AA, Alhindi A (2021) Big data platform for educational analytics. Received 28 Feb 2021, Accepted 31 Mar 2021, Date of publication 2 Apr 2021, Date of current version 12 Apr 2021
3. Pratsri S, Nilsook P (2020) Design on big data platform-based in higher education institute. Sajeewan Pratsri, Thepsatri Rajabhat University, Lopburi, Thailand
4. Alghamdi AM, Alghamdi FA (2019) Enhancing performance of educational data using big data and Hadoop. Int J Appl Eng Res 14(19):3814–3819. ISSN: 0973-4562. © Research India Publications. http://www.ripublication.com
5. Vollenbroek W, Jägersberg K, de Vries S (2014) Learning education: an 'educational big data' approach for monitoring, steering, and assessment of the process of continuous improvement of education. In: European conference in the applications of enabling technologies, Glasgow, Scotland. Efthymios Constantinides University of Twente, Enschede, The Netherlands NHL Hogeschool, Leeuwarden, The Netherlands
6. Daniel B (2015) Big data and analytics in higher education: opportunities and challenges. Br J Educ Technol
7. Shrihari MR, Archana RA, Manjunath TN, Hegadi RS (2018) A review on different methods to protect big data sets

8. Rafique A, Khan MS, Jamal MH, Tasadduq M, Rustam F, Lee E, Washington PB, Ashraf I (2021) Integrating learning analytics and collaborative learning for improving student's academic performance
9. Hutto C, Gilbert E (2014) VADER: a parsimonious rule-based model for sentiment analysis of social media text. Georgia Institute of Technology, Atlanta, GA 30032
10. Shrihari MR, Manjunath TN, Archana RA, Hegadi RS (2021) Development of security performance and comparative analyses process for big data in cloud. In: Emerging research in computing, information, communication and applications. Lecture notes in electrical engineering book series, LNEE, vol 789. https://doi.org/10.1007/978-981-16-1338-8_13
11. Wang Y, Yang Y (2020) Research on higher vocational teaching quality improvement based on educational big data. In: 2020 International conference on communications, information system and computer engineering (CISCE)

Drought Monitoring and Assessment Through Remote Sensing Data in Bundelkhand Area of Madhya Pradesh

Dhruvendra Kumar Chourishi, Anil Rajput, and Sanjeev Gour

Abstract Drought, a huge scale natural disaster translating lives of many globally, appears no less often than yearly bringing scarcity of rainfall; dry spells majorly affecting crop fields. It can be characterized based on its intensity, spatial extent, and duration, becoming much crucial in arid and semi-arid zones. Remote sensing is a latest and speedy technology for the acquisition of raster images, converting them into measurable parameters taking into account various software tools. The study uses Linear Image Self Scanner III (LISS III) to provide a multispectral raster data in four bands to acquire the value of Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI). This study aims to assess the vegetational health and water indices in the semi-arid Bundelkhand region of Madhya Pradesh. Using remote sensing data from LISS III, we will analyze NDVI and NDWI values to monitor drought development in the surrounding areas.

Keywords NDVI · NDWI · Raster data and LISS III

Anil Rajput and Sanjeev Gour had contributed equally to this chapter.

D. K. Chourishi (✉)
Department Of Computer Science and Engineering, Government Women's Polytechnic College, Shivaji Nagar, Bhopal, Madhya Pradesh 462016, India
e-mail: dhruve23@yahoo.com

A. Rajput
Higher Education Department, OSD, Government of Madhya Pradesh, Bhopal, Madhya Pradesh 462004, India

S. Gour
Department of Computer Science, Career College, Bhopal, Madhya Pradesh 462023, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_13

1 Introduction

Since ancient times, predicting droughts and monitoring their occurrence has been a significant concern. While their impact on crops, society, and the economy can only be managed to a certain extent once they occur, timely management can help minimize their adverse effects. Remote sensing is a non-contact method of observing and gathering information about various objects on the earth's surface, such as water bodies, vegetation, built-up areas, and bare soil. It involves the use of sensors to collect data without physical contact [1]. Remote sensing technology monitors drought using NDVI and NDWI drought indicators. It is one of the latest tools that provide real-time results. The Normalized Difference Vegetation Index (NDVI) is a prevalent and extensively employed remote sensing index [2]. It has been observed by scientists and researchers that healthy plants from the visible spectrum of light absorb the red (0.62–0.68 μm) and the blue (0.45–0.5 μm) wavelength, reflecting the green (0.52–0.59 μm) and near-infrared (0.77–0.86 μm) wavelengths. NDVI can be used to monitor crop health, estimate crop yield, and detect changes in vegetation cover over time, while NDWI can be used to monitor water resources, detect changes in water levels, and map wetlands. Both NDVI and NDWI are useful indices for environmental monitoring and can provide valuable insights for a variety of applications, including agriculture, forestry, and water management for developing effective strategies for sustainable development in the region.

2 NDVI

By analyzing remote sensing measurements from a distant platform, the NDVI serves as a straightforward numerical indicator to determine the presence or absence of live green vegetation in the observed target or object [3]. The NDVI technique utilizes the disparity in spectral reflectance between near-infrared (NIR) and red wavelengths to quantify vegetation and other characteristics. It measures variations in chlorophyll content by assessing the absorption of visible red radiation within the spongy mesophylls of vegetation canopies and the subsequent reflection of NIR radiation [4]. To derive the NDVI from reflectance images, it is calculated using the average values of the red bands (0.63–0.69 μm) in channel 3 and the near-infrared (NIR) band (0.78–0.90 μm) in channel 4. Alternatively, it can be computed using the average values of the green band (0.52–0.60 μm) in channel 2 and channel 4. The NDVI technique establishes a relationship between the absorption of red radiation by chlorophyll in vegetation and the significant scattering of NIR radiation [2]. Another popular index is Normalized Difference Water Index (NDWI), this parameter serves as an indicator of the moisture level present in vegetation [5]. In other words, it is better used as a supplement to NDVI. NDVI and NDWI are two normalized indices utilized to represent the surface characteristics of the study area through grayscale maps.

3 NDWI

The NDWI is a relatively new index derived from satellite data, which combines the near-infrared (NIR) and short-wave infrared (SWIR) channels. It captures changes in both water content (through SWIR radiation absorption) and the properties of vegetation canopies, such as spongy mesophylls. Therefore, it is important to continuously assess the satellite-derived NDVI and NDWI indices to gain a deeper understanding of their response to vegetation and water bodies. This understanding is crucial for monitoring drought stress on plants [2]. The reflectance ratio of water bodies in the near-infrared (NIR) band is significantly lower compared to the visible band. This characteristic is leveraged by using the NDWI to identify and extract water bodies. A similar methodology has been employed to develop the NDWI index [2].

4 Area of Interest

Our area of interest is the Bundelkhand region situated at (Longitude 78.50′ E to 81.50′ E and Latitude 23.0′ N 26.0′ N which is a part of Madhya Pradesh and Uttar Pradesh, which lies in central India fall under in semi-arid zone. It has been observed that the rainfall in arid and semi-arid zone is less than in other areas. Bundelkhand is a region known for its dry and arid climate, which makes it particularly vulnerable to water scarcity and drought conditions. NDVI and NDWI allows for the monitoring of temporal variations in vegetation and water content, offering valuable insights for the effective management and conservation of water resources in the region. It has also large rural and agricultural region, with many smallholder farmers relying on rainfed agriculture for their livelihoods. By employing NDVI, crop health and productivity can be monitored, supplying farmers and policymakers with crucial data to make informed choices regarding crop management and ensure food security. This region has experienced significant land-use changes and environmental degradation in recent years, including deforestation and soil erosion. NDVI and NDWI enables the monitoring of such changes and facilitates the tracking of progress in restoration endeavors.

5 Data Collection

Data collection for NDVI and NDWI can be done using remote sensing techniques, which involve the use of satellite or aerial imagery [6, 7]. Satellite data acquired from raster images is crucial in computing the normalized difference vegetation index (NDVI) data, which has significantly contributed to monitoring vegetation drought conditions [8, 9]. The following paper has used the LISS III sensor data of NRSC from ISRO, having the 24 days repetition of cycle with 23.5-m resolution and 140-km ground swath. It provides the multispectral remote sensing data in four bands. It is a key source of the most popular Index, NDVI, and NDWI data. The satellite sensors gives the raster image, processed using specialized software to calculate NDVI and NDWI values for each pixel in the image by the reflectance of various wavelengths from the observed vegetation [10]. Remote sensing techniques provide a cost-effective and efficient way to collect data for large areas.

6 Methods

To analyze the vegetation stress with the help of satellite data through NDVI and NDWI in Sagar, Charlatanry, Panna, Tikamgarh, Datia, Damoh which are more drought-stressed districts than other parts of Madhya Pradesh. We used QGIS to analyze the various raster images of the research area, it is an open source geographical information system software which converts raster images into numerical values. It is based on the geographic information system (GIS), which allows the users to analyze spatial information. It also supports both raster and vector image data. NDVI uses $B3$ and $B4$ bands while NDWI measures the water index using $B4$ and $B5$ bands. These bands of certain reflected wavelengths are used to measure the indices. NDVI estimates the health of vegetation, while NDWI value indicates the water content [8]. Healthy and green vegetation absorbs more electromagnetic spectrum of visual range, mainly blue (0.4–0.5 μm) and red (0.6–0.7 μm) while the reflectance of the green spectrum (0.5–0.6 μm) and near-infrared (0.7–1.3 μm) is highest. It is due to the internal structure of the leaves (Table 1).

To analyze, process and calculate the NDVI and NDWI data obtained from the raster image (Linear imaging self-scanning system). Resource I/Resource II (LISS III) from Bhuvan's NRSC (government of India portal). A toposheet represents a specific area of a station which is repeatedly observed by a satellite. With a toposheet number assigned to each area under study, the data of a toposheet in five years of all the district has been used, three sets of values in a year is been taken of each toposheet and bounding box numbers indicates the latitude and longitude of that area. LISS III is a multispectral camera that operates in four bands $B2$ (0.52–0.59 μm), $B3$ (0.62–0.68 μm), $B4$ (0.77–0.86 μm), $B5$ (1.55–1.70 μm). The sensor provides a 23.5-m

Table 1 Range of NDVI values

S. No.	Range of value	Objects
1	− 1 to 0	Water bodies
2	− 0.1 to 0.1	Barren rock, sand, and snow
3	0.2–0.4	Sparse vegetation
4	0.4–0.6	Moderate vegetation
5	0.6–1.0	Dense vegetation

resolution with a swath area of 140 km. It repeats scanning of the same toposheet after every 24 days and gives the GeoTIFF image file format. We obtain four different bands of certain reflected wavelengths. The QGIS software is employed to extract the NDVI and NDWI values from the raster image acquired from the LISS III sensor.

The Normalized Difference Vegetation Index (NDVI) is calculated using the following formula:

$$\mathrm{NDVI} = \frac{\text{Band 4 (NIR)} - \text{Band 3 (RED)}}{\text{Band 4 (NIR)} + \text{Band 3 (RED)}}$$

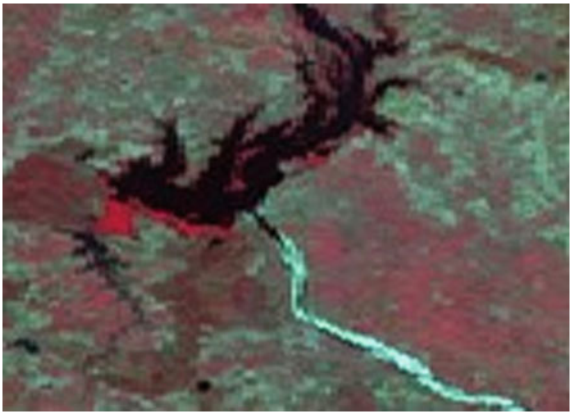
The NDVI values can range from − 1.0 to 1.0, with higher values indicating the presence of green vegetation and lower values representing other typical surface materials [2]. It is used for drought monitoring and famine early warning. Higher NDVI value also indicates healthier vegetation. While the Normalized Difference Water Index (NDWI) is calculated using the following formula:

$$\mathrm{NDWI} = \frac{\mathrm{NIR} - \mathrm{SWIR}}{\mathrm{NIR} + \mathrm{SWIR}}$$

NDWI serves as an effective indicator of vegetation water content. The intensity of both NDVI and NDWI can be classified into three categories (low, moderate, and high) based on mean, standard deviation (SD), and value ranges derived from the NDVI and NDWI outcomes [9]. Using QGIS we obtain minimum and maximum values of NDVI and NDWI which provide us a range rather than a precise value (Figs. 1, 2, 3, 4, 5 and 6; Table 2).

The output of the raster calculator will be a new raster layer showing the NDVI or NDWI values for each pixel in the image. QGIS is a widely used open-source geographic information system (GIS) software known for its versatility in various applications, including remote sensing analysis. Here only a minimum and maximum value of NDVI and NDWI is considered.

Fig. 1 Chhatarpur



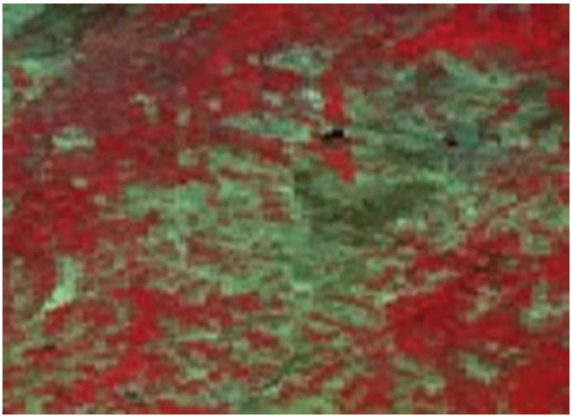


Fig. 2 Datia

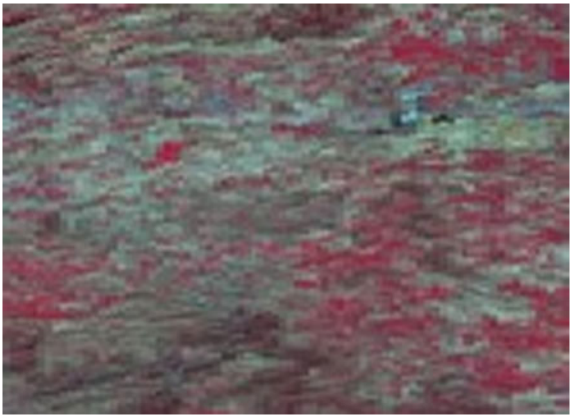


Fig. 3 Damoh

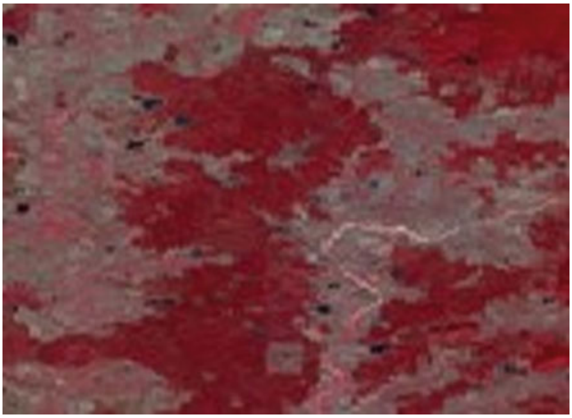


Fig. 4 Panna

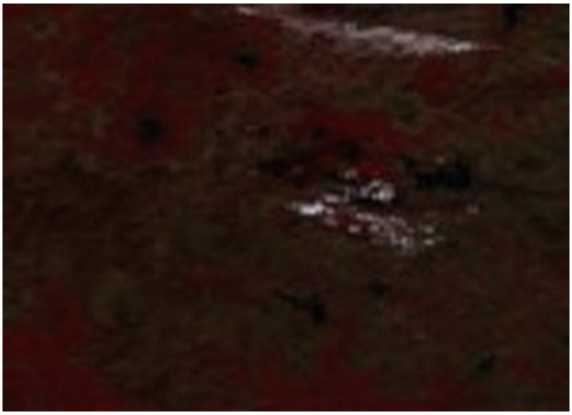


Fig. 5 Sagar

Fig. 6 Tikamgarh

Table 2 Geographical location of Chhatarpur, Datia, Damoh, Panna, Sagar, and Tikamgarh

S. No.	Area	Topo sheet number	Longitude	Latitude
1	Chhatarpur	G44T09	79.5910° E	24.9168° N
2	Datia	G44MO6	78.4609° E	25.6653° N
3	Damoh	F44G15	79.4387° E	23.8323° N
4	Panna	G44U02	80.1875° E	24.4892° N
5	Sagar	F44A11	78.7378° E	23.8388° N
6	Tikamgarh	G44S12	78.8321° E	24.7456° N

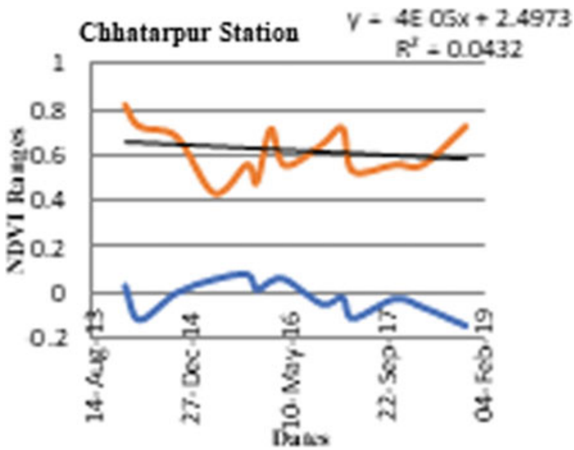
7 Results and Discussion

The remote sensing data of six years from 2014 to 2019 has been used to calculate the vegetation health in research areas for different seasons by taking the raster images of Charlatanry, Panna, Damoh, Datia, Sagar, and Tikamgarh stations for different seasons. NDVI is a measure of vegetation greenness and health, and low NDVI values can indicate less vegetation cover and/or less healthy vegetation, which can be a result of drought conditions. From the dataset, we observed that there are several instances where the NDVI values are low such as 2014, 2015, 2016, 2017, 2018, and 2019 shows significant decreasing trend. While Normalized Difference Water Index (NDWI) is an index commonly used to monitor the water content or water stress in vegetation. In particular, 2014, 2015, 2016, 2017, and 2019 the lowest mean NDWI values, which could indicate drought conditions during those years. NDWI values range from − 1 to 1, where values closer to − 1 indicate a higher probability of drought and values closer to 1 indicate a higher probability of wet. However, we should note that the standard deviation is also high for these years, indicating variability in the data. It has been observed that the maximum NDVI values lie under the range of − 0.9767 to 0.98649 in different years, while on the other hand NDWI for the proposed area having the range of − 0.986 to 0.98649. It is also observed that the maximum range of NDVI from 2014 to 2019 varies for different areas from 0.76101 to 0.98649, while the minimum range of NDVI varies from − 0.1447 to − 0.9767. The analysis of NDWI values for the same areas and years showed a range of maximum values from 0.72414 to 0.98649 and minimum values ranging from − 0.9859 to − 0.3122. These results suggest that there were variations in the water content and distribution in the vegetation for different seasons and areas. The areas with higher maximum NDWI values could indicate the presence of water bodies or moist soil, while areas with lower values could indicate dry soil conditions. Overall, the results of both NDVI and NDWI analyses provide valuable information about the vegetation health and water content in the study areas and can be used for further research and management purposes.

8 Conclusion

Result from this study indicate that there is a strong relationship between remote sensing data derived vegetation index and precipitation data. We can conclude from the calculation that during the winter season the vegetation health is good in all the districts, after while it varies from 0.35 to 0.98 and indicate moderate vegetation and sparse most of the time. Every year, April and May find vegetation stress while October to January there is a comparatively dense vegetation. Therefore, we may predict that the likelihood of severe drought conditions is low in the near future. It’s important to continue monitoring the NDVI values and other relevant factors to assess the current and future drought conditions. We can also consider human activities such as deforestation or agriculture, or other environmental stressors that affected vegetation growth. Which may impact vegetation health. To predict future drought conditions, we can analyze the trend of NDVI values over time. If we notice a trend of decreasing NDVI values over time, this may be indicative of worsening drought conditions. The interpretation of NDVI values should be done cautiously due to potential influences from factors like cloud cover and sensor calibration, which can impact the accuracy and reliability of the data. We also observed that the low NDWI values for most of the years are negative, indicating a lack of water content in the vegetation. To predict drought conditions, we can look at the trend of Low NDWI values over the years. If there is a decreasing trend, it could indicate a worsening of drought conditions. Similarly, an increasing trend in High NDWI values could indicate improved water content in vegetation. It is important to monitor NDWI values over time to identify and predict drought conditions, which can help in managing water resources and minimizing the impact on ecosystems and agriculture. This can help inform decisions related to water management and agriculture. This result improve our understanding of how satellite based index respond to vegetation health. We recommend continue evolution in NDVI and NDWI as a mandatory tool for all the geographic location. The model can take into account various factors that affect NDWI, such as weather, land cover, and soil properties. We can take proactive measures to mitigate the impact of drought on agriculture and other sectors that rely on water resources. A sudden decrease in NDWI values could indicate a decrease in vegetation water content, which could be an early warning sign of drought. This section showcases the experimental results using satellite images captured over different durations and years, providing visual representations of the findings. This provide the visual comparison as well as the quantitative validation. This remote sensing data is very effective in in drought assessment. Future research scope can be done through the higher spatial and temporal resolution satellite and sensors with improved technology (Figs. 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 and 18).

Fig. 7 Chhatarpur NDVI



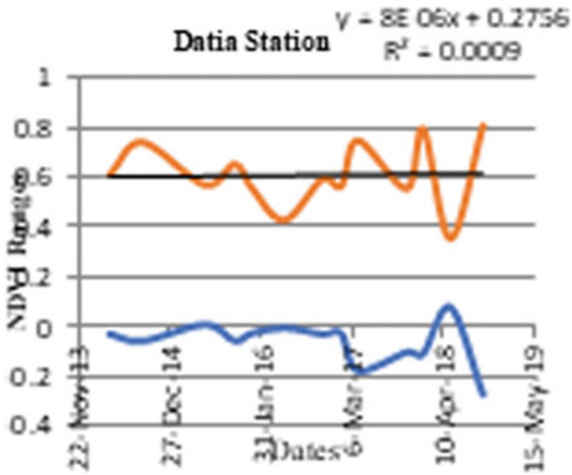


Fig. 8 Datia NDVI

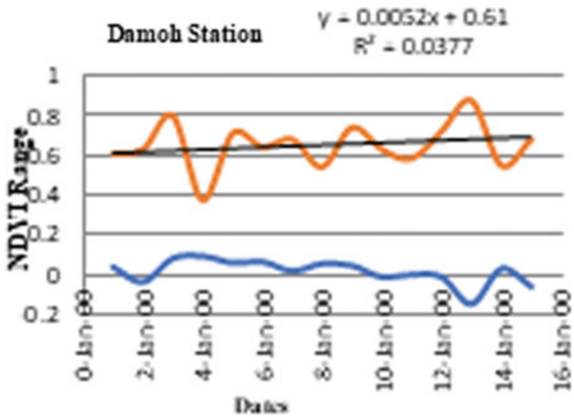


Fig. 9 Damoh NDVI

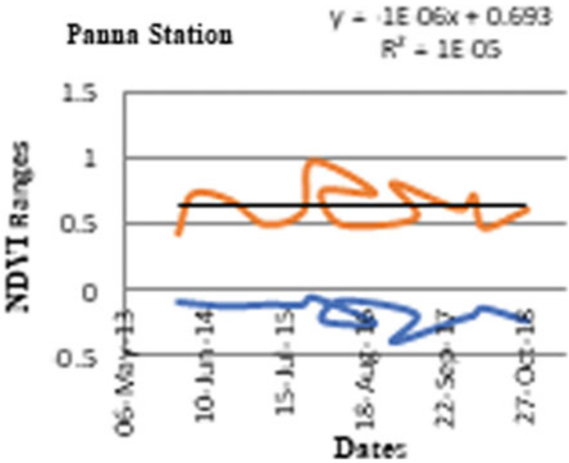


Fig. 10 Panna NDVI

Fig. 11 Sagar NDVI

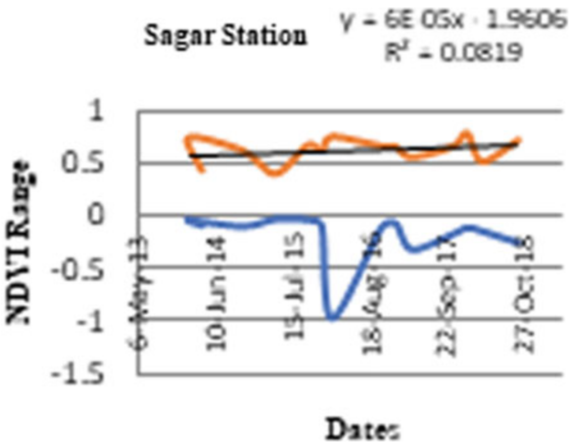


Fig. 12 Tikamgarh NDVI

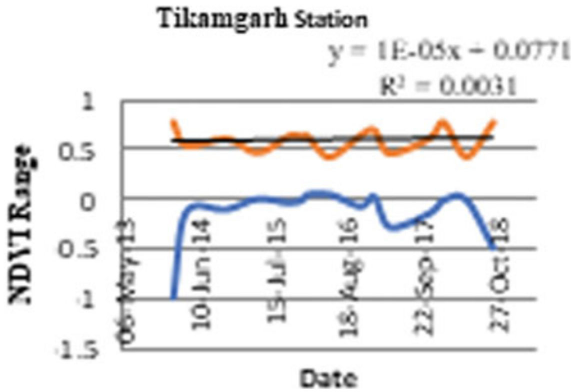
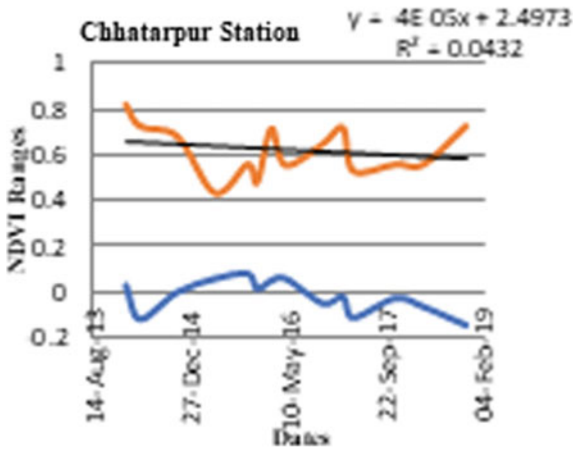


Fig. 13 Chhatarpur NDWI



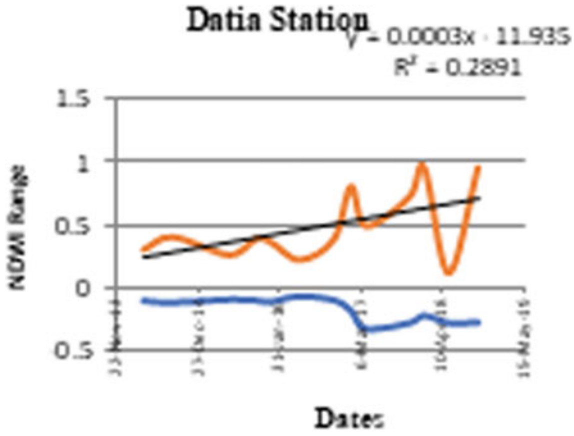


Fig. 14 Datia NDWI

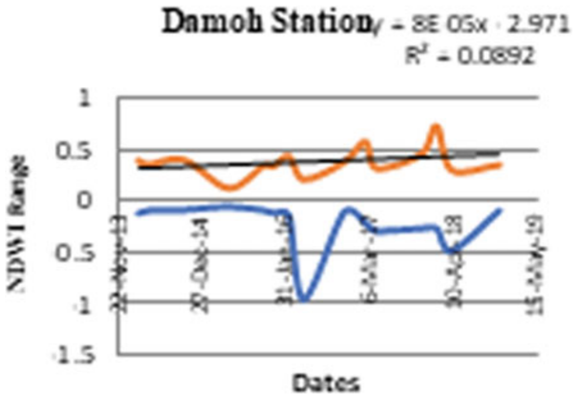


Fig. 15 Damoh NDWI

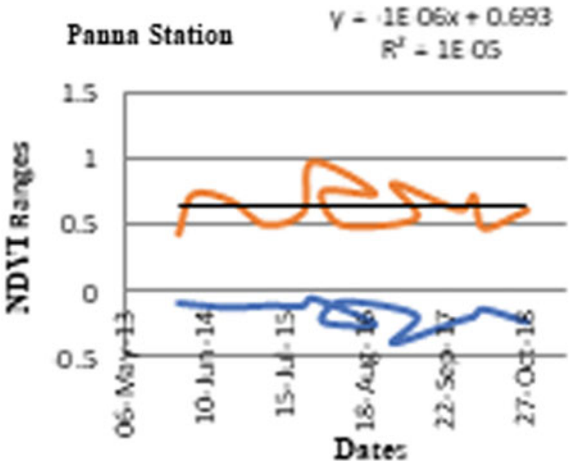


Fig. 16 Panna NDWI

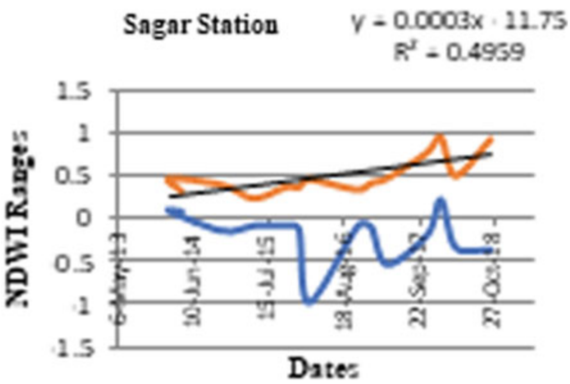


Fig. 17 Sagar NDWI

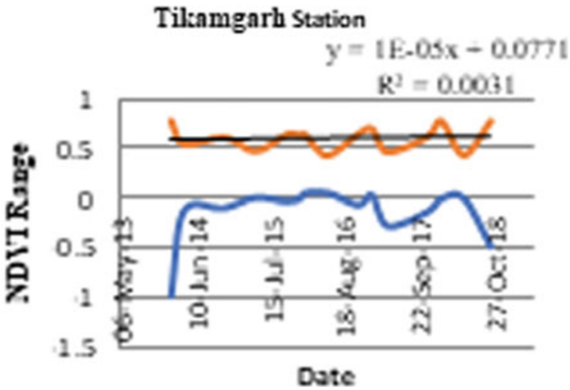


Fig. 18 Tikamgarh NDWI

References

1. Zhong Y, Ma A, Ong Y, Zhu Z, Zhang L (2018) Computational intelligence in optical remote sensing image processing. Appl Soft Comput 64:75–93
2. Bhandari A, Kumar A, Singh GK (2015) Improved feature extraction scheme for satellite images using NDVI and NDWI technique based on DWT and SVD. Arab J Geosci 8:6949–6966
3. Santos P, Negri AJ (1997) A comparison of the normalized difference vegetation index and rainfall for the Amazon and Northeastern Brazil. J Appl Meteorol Climatol 36(7):958–965
4. Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8(2):127–150
5. Gao B-C (1996) NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens Environ 58(3):257–266
6. Goward SN, Markham B, Dye DG, Dulaney W, Yang J (1991) Normalized difference vegetation index measurements from the advanced very high resolution radiometer. Remote Sens Environ 35(2–3):257–277
7. Somvanshi S, Kunwar P, Singh N, Kachhwaha T (2011) Water turbidity assessment in part of Gomti River using high resolution google earth's Quickbird satellite data. Geospatial World Forum 18–21

8. Gu Y, Hunt E, Wardlow B, Basara JB, Brown JF, Verdin JP (2008) Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data. Geophys Res Lett 35(22)
9. Kogan FN (1995) Application of vegetation index and brightness temperature for drought detection. Adv Space Res 15(11):91–100
10. Gu Y, Brown JF, Verdin JP, Wardlow B (2007) A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central great plains of the united states. Geophys Res Lett 34(6)

Smart Glasses Designed Using ESP32-Cam Coupled with Google Lens

G. Jai Sreedhar, Jainam S. Mardia, T. Jagan Balaji, and S. Sudharsan

Abstract Smart glass, one of the most recent developments, is a wearable gadget known as a switchable glass that can handle a wide variety of computer tasks that a smartphone can. The fast development of this wearable technology has resulted in various research projects, including smart glasses applications in the healthcare sector and societal challenges. In this regard, the proposed work provides an audio assistive technology for visually impaired people and those who struggle to read in order to improve their socializing capacity and independence. The suggested work is intended for visually impaired people who own a smartphone. The proposed work is a camera-mounted smart glass that is meant to be connected to a smartphone via Wi-Fi module to provide an easy and advanced solution for vision-challenged people to detect text and some items in front of them via auditory output.

Keywords MacroDroid · ESP32-cam · ESP8266 · TTS · Google Lens · OCR

G. J. Sreedhar (✉) · J. S. Mardia · T. J. Balaji
Department of EEE, Rajalakshmi College of Engineering, Chennai, Tamil Nadu, India
e-mail: 200901032@rajalakshmi.edu.in

J. S. Mardia
e-mail: 200901033@rajalakshmi.edu.in

T. J. Balaji
e-mail: 200901031@rajalakshmi.edu.in

S. Sudharsan
Department of ECE, Rajalakshmi College of Engineering, Chennai, Tamil Nadu, India
e-mail: sudharsan.s@rajalakshmi.edu.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_14

1 Introduction

According to the World Health Organization (WHO), around 39 million people worldwide are blind, and 284 million are visually impaired, including the elderly, who are visually impaired due to age [1]. These people are dependent on visually normal people to interpret the text or to identify the object for them. This group of people were constrained to depending on someone for their daily tasks. And since many subjects or distinguished books were not accessible in Braille format or in audio format, blind people had a restricted number of books to read.

Many useful devices are available in this digital era that are associated with smartphones, and these gadgets are mostly designed for ordinary people. Some technologies can even assist people with impairments, although they are not widely available and are sometimes extremely expensive. One of them is a reading device for the visually impaired. There are several text-to-speech converters accessible with smartphones, but they cannot be utilized by the visually impaired for real-time text recognition. The difficulties in using those gadgets by visually impaired people are numerous [2]. A function is added to the smartphone through this proposed work to enable them to evolve toward an independent existence and to assist them. This system uses the Wi-Fi module coupled with a camera module to receive real-time photos from the camera in the spectacles, and the images are sent to smartphones using the Wi-Fi module and the images are transformed to text using Optical Character Recognition (OCR). Despite the fact that there are several technologies for picture to text conversion, Google Lens always has an advantage over others since it can detect all languages and numeric. It can also recognize items. Google Lens's capability, reliability, and accuracy in text and object recognition is a widely researched issue [3]. This simplifies the complexity of the device's construction to a larger extent by integrating it with the smartphone. This is because it is necessary to perform complex operations that require a high level of processing power and a huge volume of prerecorded data to perform OCR. This reduces the price of the microcontroller used as well as the size and weight of the whole module, making the device more affordable. Bulky modules can generate a lot of heat and radio frequency, which can be harmful to humans over a long period of usage.

2 Literature Survey

The electronic visual aids are mostly used in the following ways. Refreshable braille displays, electronic travel aids (ETA), and location trackers are examples. Each alphanumeric element on a braille display is represented by a group of six binary dots. The six-segment display comprising these dots is designed in such a manner that the user's hands can recognize the letters via the inclination of the dots. In other scenarios, real-time messages are recognized and converted into braille code [4, 5].

Electro active polymers (EAP) [6], piezoelectric linear motors [7], and soft actuators [8] are used to manufacture braille displays. The various ways of building an effective refreshable braille display have been thoroughly investigated [9]. Li et al. [10] developed and built a navigation device that navigates the visually impaired using Google Tango and mobile computing. It functions by utilizing Google Tango and a smartphone that enables interior navigation. Unlike the works outlined above and similar works [5, 11–18], this proposed work focuses on developing a prototype which is both smartphone compatible and cost effective.

3 Proposed System

The proposed work is a smartphone-controlled assistive glass that allows visually impaired users to recognize real-time messages. Texts displayed on boards, notes, books, instruction manuals, product descriptions on the covers, and other similar practical contexts are examples of such texts. Texts that are intended to be recognized by users are digitally pronounced. This is achieved via the text-to-speech converter, which is included in the majority of smartphones as a tool for the visually impaired. OCR is used to feed the texts and to digitally pronounce them. Photos of real-time texts should be obtained before employing OCR. Because of its low cost, high performance, and ease of installation, a Wi-Fi controlled camera module (ESP32 camera module) is utilized for capturing real-time photos in this proposed work. The ESP32's ability to collect photos for OCR has been extensively investigated [19]. Figure 1 illustrates the device's overall operation. The picture is captured when the trigger is pushed and then the image is transferred to the smartphone via the ESP32 image server. The recorded image is then transferred to Google Lens via a pre-automated procedure performed with MacroDroid. The acquired image is analyzed by Google Lens. The MacroDroid's software capability to help impaired people with their daily activities has been thoroughly established [20]. Studies on the effect of employing mobile automation on a group of impaired people show that MacroDroid can be effective for disabled people [21, 22]. The text on the picture is turned into audio, which the user receives. In the case of object identification, the picture is converted to text and then into audio whenever the user presses the trigger button.

3.1 Capturing the Image

When the user wishes to hear the texts that come their way, they must press the trigger button. This method begins by defining the size of the pictures that will be captured by the camera. The picture extension is binary structure (BIN), since it allows for quicker loading without sacrificing image quality. This extension was chosen because it uses less memory and has a faster access speed. This is done because detecting messages in real time is more crucial than collecting a high-quality image. In this

Fig. 1 Process flow diagram of the proposed work

work, the pixel rate has been set at (852 × 480), which is 480p resolution, so that the module does not get overheated. The ESP32 cam module can capture images at a resolution of up to (2048 × 1536) pixels. When these operations are performed, the picture is captured by the ESP32 cam module and delivered to the ESP32 cam module's image server. And all of these operations are automated with a single trigger button.

3.2 Processing of Image to Text

Once the image reaches the image server of the ESP32 cam module, it is automatically transferred to the Google Lens application, which would be a pre-installed application in all Android smartphones and an installable software in all other devices. When an image is sent to Google Lens, it is processed to extract text from the image or to identify the image that was captured and sending the image to the Google Lens cannot be directly done via ESP server so the process of transferring the data from the server to the Google Lens application is assigned to the separate functional application MacroDroid, as shown in Fig. 2 as a flowchart.

4 Design and Implementation

4.1 Schematic Representation of the Circuit

An ESP8266 is used to upload the code to an ESP32 camera module instead of an Future Technology Devices International (FTDI) programmer. This is done so that, in addition to being beneficial for camera programming, it can also be utilized for future work in enhancement of this proposed work. The problem of the ESP32 camera's limited number of pins may be solved by utilizing ESP8266. The ESP32 camera's GPIO0 pin is linked to the ESP32's ground (GND) pin. This ensures that the hardware is prepared for programming, i.e., that it enters the writing mode. To keep the ESP8266 powered on, the enable pin is linked to the ground pin of the same device. The transmission and receiving pins of both hardware modules are linked to ensure that the software is correctly uploaded. To leave programming mode, disconnect the connection between the GPIO0 and GND pins. Figure 3 depicts these linkages.

4.2 Design of the Glass

Figure 4 illustrates the hardware implementation of the proposed work, which includes the installation of the modules over the glasses. In this design, the ESP32 camera module is placed over the glass component, which functions as an eye and can capture pictures in the same way as the human eye functions. To avoid interfering with or complicating the connections between the modules, the EPS8266 module is placed over the adjacent frame near the ESP32 cam module.

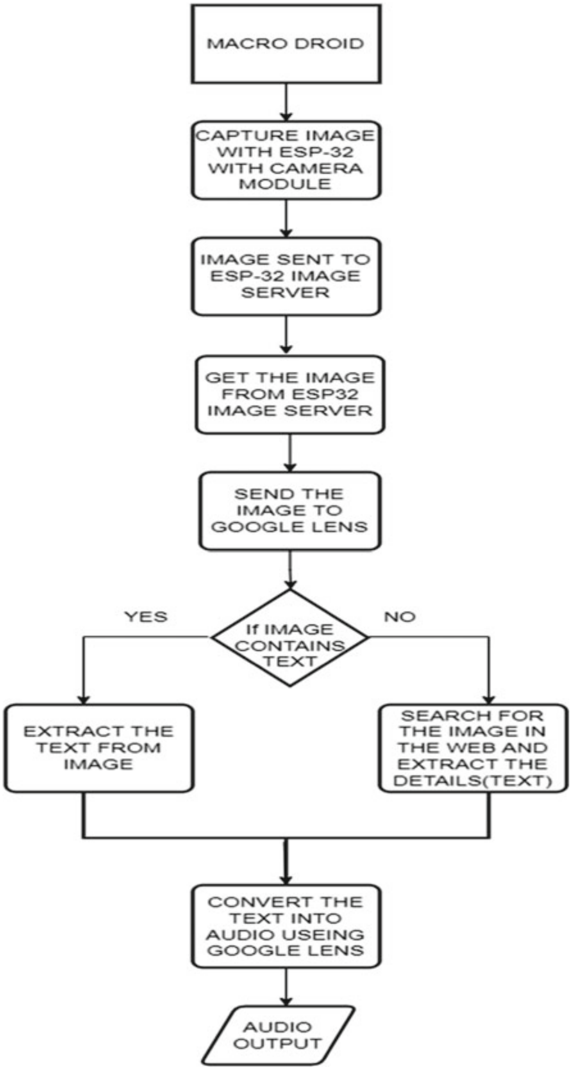


Fig. 2 Automation process involved

5 Results and Discussion

When the user presses on the trigger button, the module serves as a server. When the server is initialized, its IP address is copied. As shown in Fig. 5, this duplicated IP address is instantly open. It is launched anytime a user wishes to detect a real-time text message, via a trigger button on their smartphone as shown in Fig. 6. That webpage displays the image collected from the server (ESP32 cam). This picture is now available for OCR.

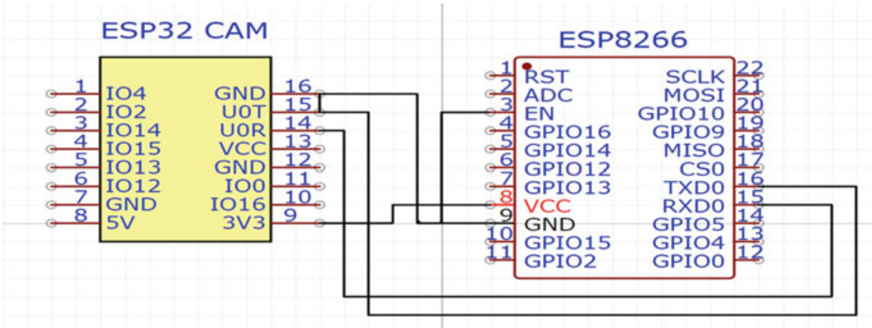


Fig. 3 Schematic circuit diagram

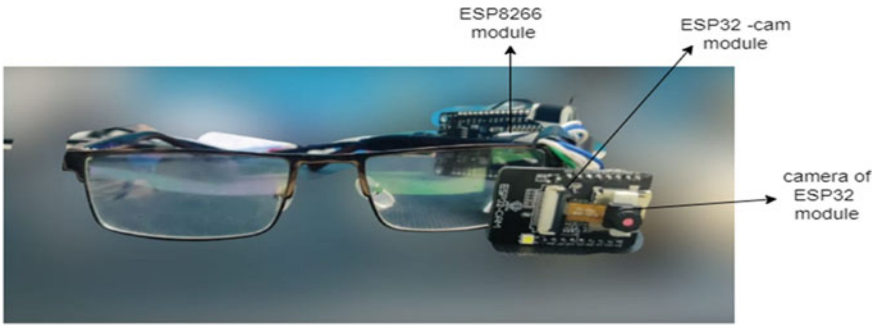


Fig. 4 Design of glass

As illustrated in Fig. 7, this captured image is automatically searched with Google Lens, and the process of character recognition begins. The MacroDroid automation is configured to search for the button labeled "text." Once it locates the button, the text will be automatically clicked. The lens begins character recognition. The phrase "Welcome to Rajalakshmi Engineering College" has been chosen. The non-editable texts are turned into editable ones at this stage. Figure 8 illustrates this option. When the text is selected, the MacroDroid automation looks for a button labeled "listen." When the listen button is discovered, it is automatically clicked. TTS is applied to it (text-to-speech conversion). This could be heard by the user through the phone's speakers. As a result, the user recognizes the phrase "Welcome to Rajalakshmi Engineering College." The square encircling the phrase "Rajalakshmi," as shown in Fig. 9, indicates that this name is spoken. The pause icon at the bottom shows that the material is being read aloud which is the major difference between Figs. 8 and 9. The remaining phrases are pronounced in the same manner. This proposed work, which is now supported by Google Lens, can make the user identify over 100 languages and special characters and numeric and even helps in identifying objects.

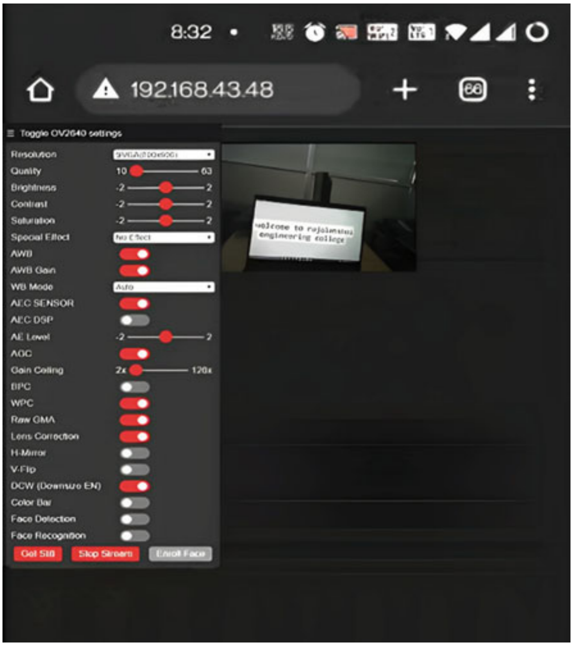


Fig. 5 Webpage showing the captured image

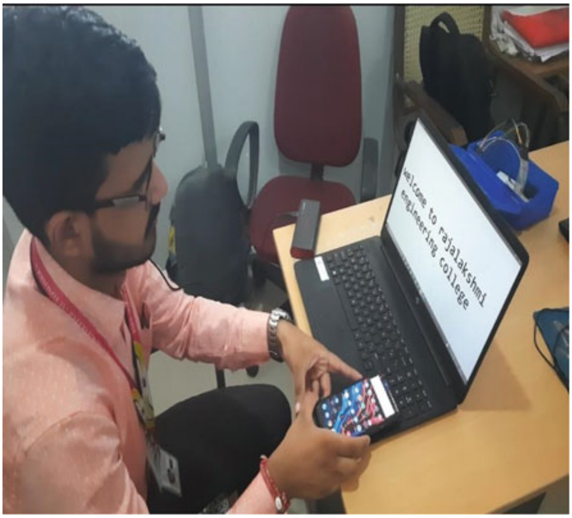
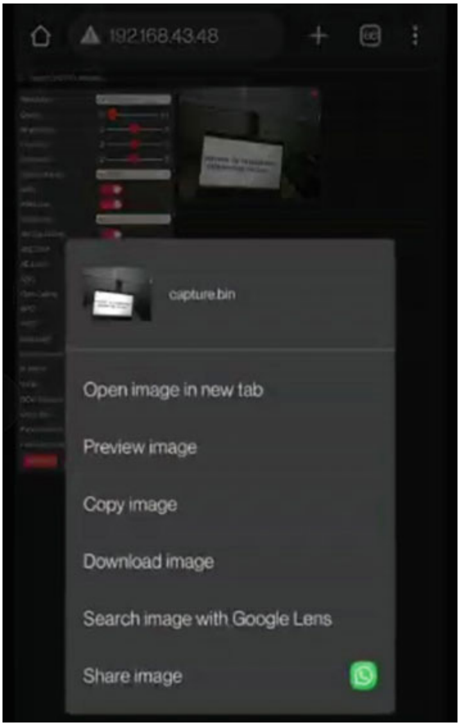


Fig. 6 User pressing the trigger button on phone

Fig. 7 Searching the image



6 Conclusions and Future Work

As a result, this work was developed and implemented in a smartphone-controllable manner. This work is a lightweight, low-cost camera integrated microcontroller ESP32 camera device that can be driven by smartphones. When the same process is tested with different languages, the gadget effectively recognizes the texts. This element may enable this work to fulfill global demand. Aside from that, the suggested system may be calibrated and taught to detect faces, lead users to their destinations, warn users of potential physical disruptions, and so on. The main objective for considering Google Lens instead of the pre-queried object detection in the ESP sever is because Google has pretrained all the possible objects, texts and languages. Google has achieved this using the data collected from millions of users and enormous research. This makes it easier to implement Google Lens for any purpose which needs image recognition. Google also has the function to enable a voice assistant in the output for text or image input which is an important outcome of the proposed system.

Fig. 8 Selecting the texts on image

To automate all these tasks and ensure convenience, the pre-built application MacroDroid has been configured and deployed. A special function is designed such that the entire process takes place at a single click or touch. This configuration is based on predefined time sequence ensuring a prompt response.

And on future work, it is currently being test to be fully automated using the cloud server and to bring it to a single module which will work without the assist of the smartphone, and further addition of obstacle detection on 360° angle using the ESP8266 will increase in more assistive for the needful people. This approach is undeniably compact while still effective.

Fig. 9 Pronunciation of text on Image

References

1. Information on http://www.who.int/mediacentre/factsheets/fs282/en/
2. Brady E, Morris MR, Zhong Y, White S, Bigham JP (2013) Visual challenges in the everyday lives of blind people. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 2117–2126
3. Lucia B, Vetter MA, Moroz O (2021) The rhetoric of Google Lens: a postsymbolic look at locative media. Rhetor Rev 40(1):75–89
4. Kumari S, Akole A, Angnani P, Bhamare Y, Naikwadi Z (2020) Enhanced Braille display use of OCR and Solenoid to improve text to braille conversion. In: 2020 International conference for emerging technology (INCET). IEEE, pp 1–5
5. Holanda GB, Souza JWM, Lima DA, Marinho LB, Girão AM, Frota JBB, Rebouças Filho PP (2018) Development of OCR system on android platforms to aid reading with a refreshable braille display in real time. Measurement 120:150–168
6. Ren K, Liu S, Lin M, Wang Y, Zhang QM (2008) A compact electroactive polymer actuator suitable for refreshable Braille display. Sens Actuators, A 143(2):335–342
7. Cho HC, Kim BS, Park JJ, Song JB (2006) Development of a Braille display using piezoelectric linear motors. In: 2006 SICE-ICASE international joint conference. IEEE, pp 1917–1921
8. Lee S, Jung K, Koo J, Lee S, Choi H, Jeon J, Nam J, Choi H (2004) Braille display device using soft actuator. In: Smart structures and materials 2004: electroactive polymer actuators and devices (EAPAD), vol 5385. SPIE, pp 368–379
9. Leonardis D, Claudio L, Frisoli A (2018) A survey on innovative refreshable braille display technologies. In: Advances in design for inclusion: proceedings of the AHFE 2017 international conference on design for inclusion, The Westin Bonaventure Hotel, Los Angeles, California, USA 8. Springer International Publishing, pp 488–498

10. Li B, Munoz JP, Rong X, Chen Q, Xiao J, Tian Y, Arditi A, Yousuf M (2018) Vision-based mobile indoor assistive navigation aid for blind people. IEEE Trans Mob Comput 18(3):702–714
11. Selvaraj C, Bhalaji N (2018) Enhanced portable text to speech converter for visually impaired. Int J Intell Syst Technol Appl 17(1–2):42–54
12. Neto R, Fonseca N (2014) Camera reading for blind people. Procedia Technol 16:1200–1209
13. Rizzo JR, Conti K, Thomas T, Hudson TE, Wall Emerson R, Kim DS (2018) A new primary mobility tool for the visually impaired: a white cane—adaptive mobility device hybrid. Assist Technol 30(5):219–225
14. Jeong GY, Yu KH (2016) Multi-section sensing and vibrotactile perception for walking guide of visually impaired person. Sensors 16(7):1070
15. Everding L, Walger L, Ghaderi VS, Conradt J (2016) A mobility device for the blind with improved vertical resolution using dynamic vision sensors. In: 2016 IEEE 18th international conference on e-health networking, applications and services (Healthcom). IEEE, pp 1–5
16. Gupta S, Sharma I, Tiwari A, Chitranshi G (2015) Advanced guide cane for the visually impaired people. In: 2015 1st International conference on next generation computing technologies (NGCT). IEEE, pp 452–455
17. Pattanshetti AC, Bhat SI, Choudhari HG (2016) Advanced bat hat for the visually impaired. Imperial J Interdiscip Res 2(6)
18. Rizzo JR, Pan Y, Hudson T, Wong EK, Fang Y (2017) Sensor fusion for ecologically valid obstacle identification: building a comprehensive assistive technology platform for the visually impaired. In: 2017 7th International conference on modeling, simulation, and applied optimization (ICMSAO). IEEE, pp 1–5
19. Adi PDP (2022) Performance evaluation of ESP32 camera face recognition for various projects
20. Lancioni GE, Singh NN, O'Reilly MF, Sigafoos J, Alberti G, Chiariello V, Carrella L (2020) Everyday technology to support leisure and daily activities in people with intellectual and other disabilities. Dev Neurorehabil 23(7):431–438
21. Lancioni GE, O'Reilly MF, Sigafoos J, Desideri L, Alberti G, Campodonico F, Chiariello V (2019) Smartphone-based technology to support functional occupation and mobility in people with intellectual disability and visual impairment. Adv Neurodev Disord 3:334–342
22. Griffin-Shirley N, Banda DR, Ajuwon PM, Cheon J, Lee J, Park HR, Lyngdoh SN (2017) A survey on the use of mobile applications for people who are visually impaired. J Vis Impairment Blindness 111(4):307–323

Hardware Architecture of Reinforcement Learning for Edge Devices

H. Sai Vardhan Reddy, Saroja V. Siddamal, and Suneeta V. Budihal

Abstract *Q* learning algorithm is a part of classic reinforcement learning algorithm. The *Q* leaning algorithm has a wide range of applications in the fields of robotics, traffic control, industry automation, etc. In this paper, a *Q* learning algorithm is implemented for an environment having nine states and four actions. This algorithm is used to determine best path that agent can take in an environment by maximizing reward value for each iteration of state and action pair. The process of maximizing reward function makes use of Markov decision process equation with Bellman equation. Using this equation *Q* table is updated for each iteration, rewards are awarded for each state action pair. Based on the updated *Q* table, agent determine the best path that the agent can take.

Keywords Reinforcement learning · *Q* table · Parallel architecture · Markov decision process

1 Introduction

Reinforcement learning is a feedback-based machine learning in which an agent learns to behave in an environment by performing the selected actions and gaining rewards as per selected actions. For each action selected, the agent might get positive or negative reward. The agent gets positive reward for each good action selected and gets negative reward for each bad action selected. The agent tries to interact with the environment and explores it by itself. The primary goal of an agent in a reinforcement learning is to maximize rewards. It is a trial and error iterative process, i.e., agent

H. S. V. Reddy · S. V. Siddamal (✉) · S. V. Budihal
KLE Technological University, Hubballi, India
e-mail: sarojavs@kletech.ac.in

H. S. V. Reddy
e-mail: harshavvardhan1412@gmail.com

S. V. Budihal
e-mail: suneeta_vb@kletech.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_15

always continues to select action and either change state or remain in same state and gets feedback.

By doing these iteratively, agent learns which state action pair leads to positive feedback and which state action pair leads to negative feedback. As this is a trial and error iterative proves, concept of reinforcement learning is very similar to human learning. Q table-based reinforcement learning algorithm is a model free off policy reinforcement learning where the best action is selected for a given state depending on where the agent is in the environment and tries to maximize reward. The quality function for Q learning algorithm is updated based on Eq. 1.

$$\mathrm{Qt} + 1(\mathrm{St}, \mathrm{At}) = Q(\mathrm{St}, \mathrm{At}) + \alpha[\mathrm{Rt} + 1 + \Gamma * \max[Q(\mathrm{St} + 1, \mathrm{At})] - (\mathrm{St}, \mathrm{At})] \quad (1)$$

Upon simplifying Eq. (1)

$$\mathrm{Qt} + 1(\mathrm{St}, \mathrm{At}) = (1 - \alpha) * Q(\mathrm{St}, \mathrm{At}) + \alpha * \mathrm{Rt} + 1 + \alpha * \Gamma * \max(Q(\mathrm{St} + 1, \mathrm{At})) \quad (2)$$

where,

$Q(\mathrm{St}, \mathrm{At})$ is quality value for current state action pair, α is learning factor, Γ is discount factor, $\mathrm{Rt} + 1$ is reward value which will be awarded for selecting particular action in a state, $\mathrm{Max}(Q(\mathrm{St} + 1, \mathrm{At}))$ is the Q value of next state action pair having maximum Q value in next state.

RL is interaction of directed agent with environment that is uncertain [1]. To maximize the rewards, the RL agents decide the current state sensing the environment to perform the next state dependent action [2]. To learn the behavior of the agent Q table-based reinforcement learning (QRL) algorithms are used [3]. It determines the Q value for state action pairs, where Q value is the quality to improve the agent's evaluation. All these state actions are stored in Q table. There has been an exhaustive research in the area of deep reinforcement learning (DRL). This has a capability to learn above large state spaces [4].

Hardware implementations of these algorithms are gaining importance due to parallelism and energy efficiency. FPGAs are the platform of choice [5]. By taking advantage of parallelism, high throughput can be achieved. These FPGA devices provide on chip memory to support Q table [6]. Due to increased computational power for more state and action pairs, reinforcement learning algorithms have to be implemented in hardware. By implementing Q learning algorithm in hardware throughput for different state action pairs has been increased when compared with CPU and SARSA algorithm [1]. Path planning using Q table-based reinforcement learning has been significantly used in the field of robotics research. Q learning is a field of reinforcement learning having absolute advantages for path planning in an unknown environment. For its high computation speed requirements, it is necessary for accelerating Q learning algorithm which is used in path planning [2].

The authors in [3] have designed a fast, parallel, and pipelined Q learning accelerator for a scalable architecture of RL accelerator and the proposed architecture is implemented on AvnetUltra96-V2 platform and real-time evaluation is performed in realistic use. In this work, the authors have built a parallel architecture to implement Q learning.

Rest of the paper is organized as follows. Section 2 discusses the proposed architecture for Q learning. In Sect. 3 results are discussed, followed by conclusion.

2 Proposed Methodology

Figure 1 shows the proposed architecture to achieve Q table-based reinforcement learning based on Markov decision process with Bellman equation. In this architecture, Q table is updated by trial and error process by selecting state action pair randomly. From above architecture, each value in Q table is known as quality value or Q value. Initially Q values for all state action pairs are initialized to value 1.

These numbers are then converted into 32-bit floating point numbers. A reward function is defined initially which gives a reward value based on present state and selected action by agent. In this architecture, three types of reward values are defined as reward0, reward1, and reward2. Reward0 is a positive reward which will be awarded for each transition from one state to another state based on action performed. Reward1 is a negative reward and will be awarded for transition from a state by performing particular action that leads agent out of the boundary. Reward2 is also a negative reward and will be awarded for transition from any state by performing particular action that leads to obstacles. The above architecture is proposed to

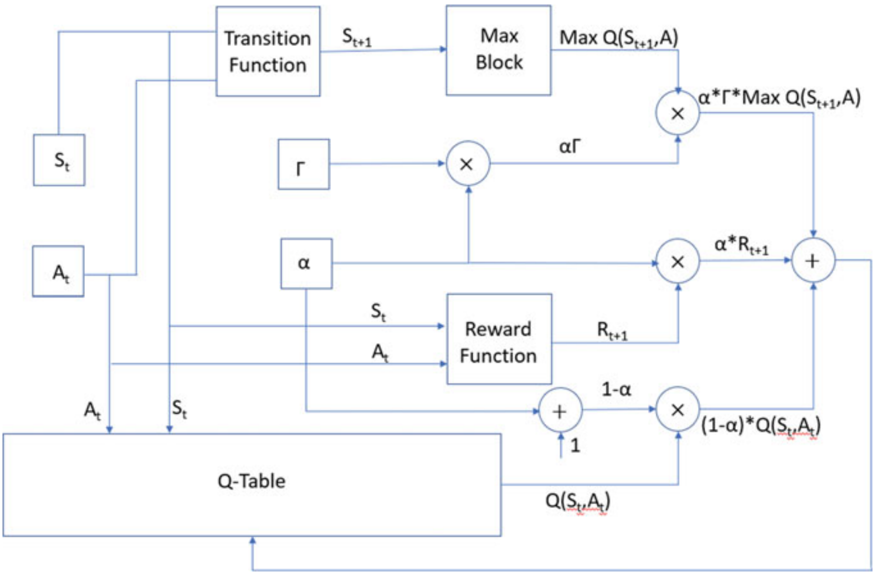


Fig. 1 Proposed architecture to achieve Q table-based reinforcement learning

achieve Q table-based reinforcement learning based on Markov decision process with Bellman equation. In this architecture, Q table is updated by trial and error process by selecting state action pair randomly. Value of alpha (α) lies in the range of 0–1. Value of alpha set near to 0 means Q value for a particular state action pair will never get updated and agent can't learn anything and Q table will never get updated. So, keeping value of α near to 1 is important so that agent learns quickly. In this architecture, value of α is considered as 0.9. Value of Gamma (Γ) lies in the range of 0–1 and quantifies importance given to future rewards. The agent will consider intermediate rewards if value of Γ is set near to 0 and will consider rewards with greater weight if value of Γ set near to 1. Here, value of Γ is considered as 0.9. The values of alpha and gamma are converted into 32-bit floating point numbers.

A random state St and random action At are selected for first iteration and for other iterations value of St is given transition function which takes present action and present state as input. The next state generated by transition function acts as input to max block which in turn provides maximum Q value of all the possible state action pairs.

Max block internally has comparators. As shown in the Fig. 2 inputs for comparators are Q values of all possible state action pairs for next state generated by transition block. As per Q value, updating equation value of $(1-\alpha)$ is calculated using a 32-bit floating point adder and value of $\alpha * \Gamma$ is also calculated using 32-bit floating point multiplier. The Q table which holds all the Q values of each state action pair takes present state and present action as inputs and provides Q value of corresponding state action pair known as $Q(\mathrm{St}, \mathrm{At})$. This corresponding Q value is multiplied with $(1-\alpha)$ using 32-bit floating multiplier. The value of α and the reward value $\mathrm{Rt}+1$ which is generated by reward function are multiplied using 32-bit floating point multiplier. The value of α and Γ and output from max block that is action having highest Q value, i.e., $\max(Q(\mathrm{St}+1, A)$ among all possible actions for next state-based on transition function are multiplied using 32-bit floating point number. The value of $\alpha * \Gamma * (\max(Q(\mathrm{St}+1, A)))$, $\alpha * \mathrm{Rt}+1$ and $(1-\alpha) * Q(\mathrm{St}, \mathrm{At})$ are added using 32-bit floating point adder and this Q value is updated back into Q table for present state and present action. This entire process is repeated for each iteration and after much iteration, the Q table is updated so that agent can learn optimal action selection policy in real time.

3 Experimental Results

The proposed architecture is designed in Xilinx ISE, the simulation is done using Vivado Design Suite 2019.1. The same is implemented on FPGA (Sparten 6). The performance of the designed accelerator is studied with grid example. In the use case example, the environment is considered as cells, the agent traverses from one cell to next avoiding the walls. It chooses the actions, collects rewards and updates Q table. Figure 3 shows an environment with nine states, four actions—left, up, right, and down.

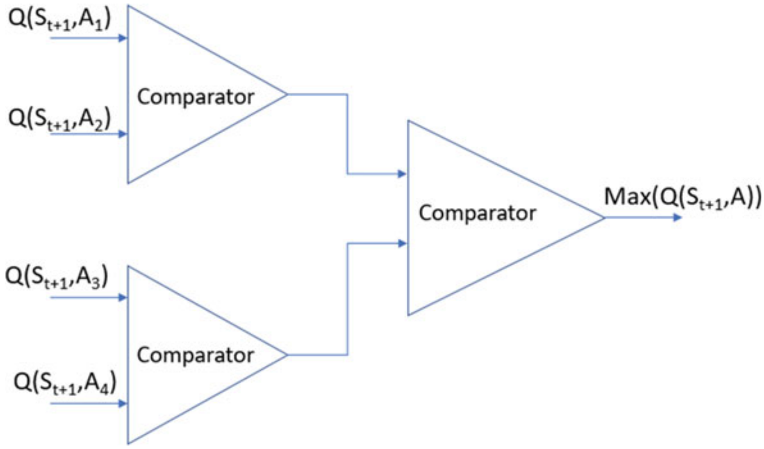


Fig. 2 Proposed architecture for max calculation

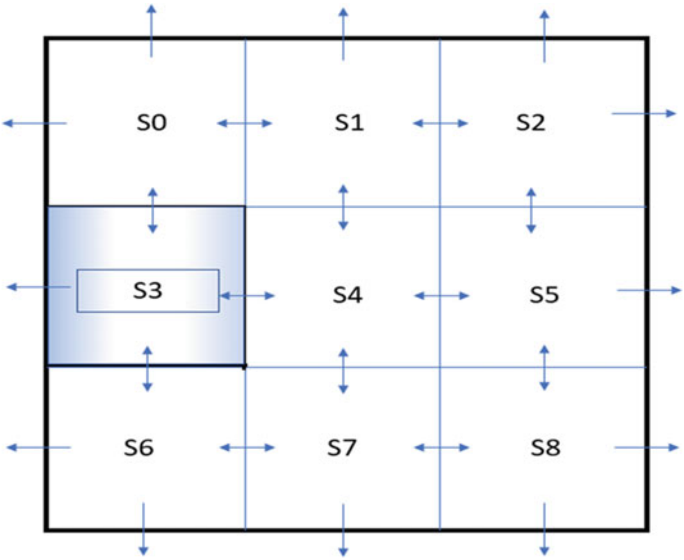


Fig. 3 Example of grid

The black border line shows the boundary of the environment. The environment is divided into nine states: S0–S8. A reward value is awarded to every state action pair. The value of α and Γ are considered to be 0.9. A positive reward is given for transition from one state to another state and a negative reward is given if state after transition is either an obstacle or a boundary. The maximum reward is awarded upon reaching the goal. When a wall is hit, negative rewards are awarded. In this environment state, S3 is considered as obstacle. Initially the Q table is initiated with

States	UP	DOWN	RIGHT	LEFT
S0	1	1	1	1
S1	1	1	1	1
S2	1	1	1	1
S3	1	1	1	1
S4	1	1	1	1
S5	1	1	1	1
S6	1	1	1	1
S7	1	1	1	1
S8	1	1	1	1

a

States	UP	DOWN	RIGHT	LEFT
S0	1	1	45.90	1
S1	1	187	1	1
S2	1	245	204	196
S3	45.90	1	1	1
S4	157	206	254	1
S5	116	243	204	271
S6	-44	1	16	-48
S7	230	1	176	22
S8	266	92	97	206

b

Fig. 4 **a** Initial *Q* table, **b** Final *Q* table

value 1 and in each iteration of this training phase this table is updated as per *Q* table-based reinforcement learning algorithm. Initially the *Q* table is initiated with value 1, and in each iteration of this training phase this table is updated as per *Q* table-based reinforcement learning algorithm (Fig. 4).

Figure shows the updated *Q* table as per *Q* table learning-based reinforcement learning algorithm. Based on this updated *Q* table, the agent learns to take optimal action policy in real-time applications.

The performance of the design is analyzed with various state action pairs. The states are represented as ‘S’ and actions as ‘A’. The actions are encoded as binary bits depending on the actions. If four actions are considered, they are represented as 00-left, 01-up, 10-right, 11-right.

4 Conclusion

In this paper, a *Q* table-based reinforcement learning algorithm architecture is proposed and the performance is evaluated for an environment having nine states and four actions, and even an obstacle is also considered in this environment. Based on the Markov decision process with Bellman equation, *Q* table is updated iteratively and path to final destination is determined by edge devices.

References

1. Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT press, New York
2. Li Y (2017) Deep reinforcement learning: an overview. arXiv preprint arXiv:1701.07274
3. Watkins CJ, Dayan P (1992) Q-learning. Mach Learn 8(3–4):279–292
4. Nair A, Srinivasan P, Blackwell S, Alcicek C, Fearon R, De Maria A, Panneershelvam V, Suleyman M, Beattie C, Petersen S et al (2015) Massively parallel methods for deep reinforcement learning. arXiv preprint arXiv:1507.04296
5. Liu X, Diao J, Li N (2022) A FPGA-based accelerator implementation for path planning using Q_learning algorithm. J Phys Conf Ser 2245(1):5475
6. Sutisna N, Ilmy AMR, Syafalni I, Mulyawan R, Adiono T (2023) FARANE-Q: fast parallel and pipeline Q-learning accelerator for configurable reinforcement learning SoC. IEEE Access 11:144–161. https://doi.org/10.1109/ACCESS.2022.3232853
7. Kalmath M, Kulkarni A, Siddamal SV, Mallidu J (2022) Implementation of 32-bit ISA five-stage pipeline RISC-V processor core. In: Pandit M, Gaur MK, Rana PS, Tiwari A (eds) Artificial intelligence and sustainable computing. Algorithms for intelligent systems. Springer, Singapore. https://doi.org/10.1007/978-981-19-1653-3_19
8. Mallidu J, Siddamal SV (2023) A survey on in-order 5-stage pipeline RISC-V processor implementation. In: Kaiser MS, Xie J, Rathore VS (eds) Information and communication technology for competitive strategies (ICTCS 2021). Lecture notes in networks and systems, vol 401. Springer, Singapore. https://doi.org/10.1007/978-981-19-0098-3_73

Strengthening the Privacy of Blockchain with Zero Knowledge Proof Case Study: Online Exam Student Verification

Ramya Rajamanickam and Saumya Chaturvedi

Abstract Blockchains provides a decentralized, distributed and immutable ledger which is verifiable and digitally available over a network, bringing in a technological revolution in this fast-forwarding digital world which helps in building smart supplychains, eGovernance, eHospital, etc. Nevertheless, blockchains are bound by divergent issues such as scalability, security and privacy. Privacy issues are of great concern these days, when every Nation is fighting against various potential privacy hackers and safe guarding the data from various types of heists. Despite numerous endeavors to concoct and merge new privacy techniques, present day blockchain methods are quite farfetched to survive the privacy issues. This situation undermines rights of the user, similar as the option to become anonymous in specific cases, the option to edit or delete data, thus decreasing the consummation of a rightful privacy preserving system on the framework of blockchain. Retaliatory action to overcome multitudes of crime causing havoc to the blockchain privacy and to work on the existing privacy mechanisms is the main focus for blockchain developers and users. The Zero Knowledge Proof (ZKP) is a method of cryptography which has the prospects to change how data is processed. It provisions to prove the users' validity without sharing any other private information, thereby opening possibilities in preserving the privacy of blockchain applications. Further we discuss a use-case which is studied and evaluated with ZKP model to facilitate user verification without gathering any other private information. The Hyperledger Fabric platform is used and Hyperledger Caliper to analyze the use-case and to calibrate the results.

Keywords Blockchain · Privacy issues · BCT · Zero knowledge proof · ZKP in blockchain

---

R. Rajamanickam (✉)
National Informatics Centre, Lodhi Road, New Delhi, India
e-mail: ramya.rajamanickam@nic.in

S. Chaturvedi
School of Engineering and Technology, Sharda University, Greater Noida, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_16

1 Introduction

Blockchain has seen great development over the past few years, and today its usage applications are well beyond the bitcoin and the cryptocurrencies. The digital transactions in a decentralized environment through a ledger barring any third party, at the same time allowing verification publicly and provisioning the transactions of the data digitally is blockchain. In recent times, blockchain is subjected to various different issues, namely legal compliance with regulations, scalability, security, privacy issues, response time which sabotage the confidentiality, anonymity and privacy of the user in their ledger transactions. The recent years have seen numerous heists in cryptocurrency, bitcoin, and blockchain. The heist from 2016 to 2021 are recorded in this chart which had caused loss in huge millions (Fig. 1).

Despite all the unfazed endeavors to concoct and embed new privacy techniques, present day blockchain methods are quite farfetched to survive the privacy issues. The challenges are involved in analyzing and evaluating the open issues which are 51% attack, Byzantine generals' problem, Double Spending attack, wallet privacy leakage, selfish mining attack, etc. Several crucial open adversities that need to be acknowledged before integration of blockchain with other technologies, includes asset management, transaction data management, scalability and cost factor. The privacy matters that contribute to blockchain are Encryption Time, Memory Utilization, Throughput, Time and Space, Key Length, Cipher Algorithm, Group Size, Percentage of Nodes accessing transaction traffic which are to be evaluated and predicted for confirming the efficiency of the security and privacy.

Here not only the transaction data but the data related to each transaction are termed prized and privacy related to them is called transaction privacy. Unlike conventional methodology of defense against preventing the unsanctioned access of such potential information, it is vital to make sure that transaction data is not likely to be

Fig. 1 Heist in blockchain in recent years, https://arstechnica.com, 2021

misused or hacked by any other unverified nodes. It becomes vital to check genuineness of each transaction such that it does not give away any vulnerable information that are not fully encrypted in the transaction. These inconsistencies also sum up to adversities and complications in privacy preservation of the blockchain. The current setup is unable to give utter privacy and at loss to protect the user. It becomes a mandate to add some methodologies, embed some privacy protection algorithms, introduce protocols or other secure approaches in achieving the blockchain privacy. This paper, focuses on blockchain privacy security issues and proposes a plausible solution using Zero Knowledge Proof (ZKP).

In terms of sensitive user information, the public blockchain systems cause privacy concerns. As per the transparent design properties of blockchain, any participant from the network can view all the contents of the ledger as it is a permission less blockchain system ex: Bitcoin. The zero knowledge proof is a method of cryptography which enables first party to prove to second party that they know the secrecy without revealing any data other than possession of the secret. This makes ZKP capable of protecting the privacy of the users in the network and using its verification process to make it a part of blockchain network.

2 Zero Knowledge Proof

In 1980 the Zero Knowledge Proof was proposed by MIT researchers Shafi Goldwasser, Silvio Micali and Charles Rackoff. Generally, the message is exchanged between two parties the prover and the verifier. The prover has to convince the verifier that some mathematical statement is true. Even though the prover has infinite resources for computation ends up not being trusted. Though the verifier has less computation power is supposed to verify the prover and expected to be honest. They both exchange messages until the verifier is convinced with the answers provided by the prover. This theory had some difficulties as the prover can trick the verifier in the course of message exchanges. Zero Knowledge Proof was proposed as a remedial to the information leakage issue where the prover can prove statements, without revealing any information to the verifier (Table 1).

Table 1 The properties of zero knowledge proof

Properties	Statement	Prover	Verifier
Completeness	True	Convinces the verifier	Follows protocol and is convinced
Soundness	False	Cannot convince the verifier	Secure against hostile prover
Zero knowledge	True	No sensitive information is shared	Only that prover is authentic

The Distributed Ledger Technology and the Blockchain Technology has enlarged its area into numerous applications. It generates consensus among all stakeholders removing the requirement of a trusted third party. In the public blockchain network all the participants can view all the transactions. When the transactions contain sensitive information, it becomes an issue of privacy. A solution to this is provided by ZKP. It only shows that a valid transaction has been carried out and withholding the other details such as sender details, recipient details, type of assets, quantity of assets. Among the interactive and non-interactive ZKP (NZKP), in this paper, NZKP is used in the proposed system.

When using blockchain with ZKP, the problem of data privacy and security in transactions can be solved. However, the current scenarios of implementations need a significant computational power to produce these zero knowledge proofs and make them work. With the increase in problem complexity the executing time of proof generation also increases. Thus an, efficient ZKP model implementations is the need of the hour. This paper proposes a Zero Knowledge proof and a permissioned blockchain strategy for achieving the privacy preservation in verification of various existing models such as Food Delivery, Home Delivery, Parcel Delivery, User Verification applications.

This paper proposes ZKP with permissioned blockchain network for verification and protection for safety and privacy of all the stakeholders by withholding any sensitive information and by benchmarking the performance with a case study.

3 Case Study: Online Exam with ZKP Model in Blockchain Network

An application caters in conducting online examination for the students. Say, the university exam center and students of the university are the main stakeholders in the system. Both parties get registered in the blockchain network. The application generates key of proof and key of verify to the students and the peer nodes in the blockchain, respectively. Like most application, the registration is a onetime activity. Before start of every examination, the online application requests for verification from their students. The student uses the key of proof and generates zero knowledge proof for identity verification. This ZKP gets send to peer nodes in blockchain network. Then, the peer node uses the key of verifier to verify the ZKP generated by the student. The student and the examination center get notification after verification is complete. Now the usual process of blockchain like execution of smart contract and insertion in ledger takes place. And access to online exam is granted to the student. The ZKP model is constructed in the blockchain network as per below Fig. 2.

The three steps of initiation, proof and verification of ZKP are:

1. Initializes a student to act as the prover. The student ID card is used and the online application generates the key pair, key of proof for student and the key of verify for the peer nodes in Fabric.

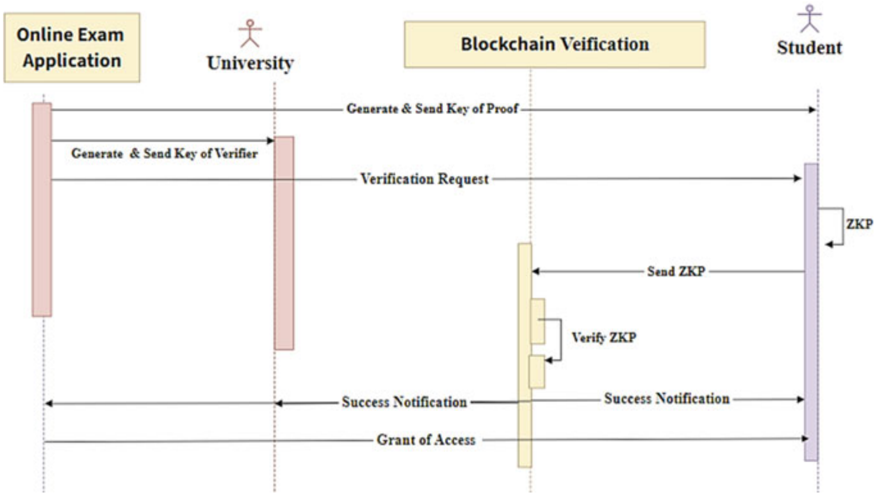


Fig. 2 Workflow of the online student verification

2. The student uses the key of proof to generate a ZKP for the secret salt and Hash-Based Message Authentication ode (HMAC) of the ID card
3. A peer node from the network verifies the ZKP from the student. The verification function takes the proof, the secret salt and HMAC, the key of verify and passes the parameters as inputs to verify the proof.

4 Framework: Hyperledger Fabric

Hyperledger Fabric is a consortium blockchain technology which is used in the implementation. Before participation the nodes must be certified to enter and be part of the network. These nodes need not be owned by one entity necessarily. To define all allowable interactions smart contracts/chaincodes can be written in any programming language in the network. The access control functionality is available for each chaincode function so that only certain peers can call it. Here, Organizations are termed to associate peers and users in a connected flow that reduces access control overhead.

1. Through the client application, a request for transaction can be initiated by a participating node of the member organization
2. The endorser peers receives the call for request through the application broadcast.
3. To endorse any transaction, the endorser verifies the certificate attributes and then runs the chaincodes by returning approved/rejected response.
4. Once approval of transaction is received, ordered peer receives instructions to order and add the accepted transaction in the all the blocks

5. The other member organizations receive these transactions through their respective anchor nodes which in turn broadcast them to the peers nodes of their organization.
6. The latest block is then updated in the local ledger. Thus the entire network has synchronized ledger.

The Hyperledger Fabric Node SDK3 is used here. It is wide known for building modular application for wide range of use-cases. Its modular and versatility satisfies the requirement of our casestudy. To connect to the network, the configuration is set up in Common Connection Profile of the SDK node. Whenever a request has to be sent to the network, any client who initiates the request has to be priorly registered with the fabric. This happens at the SDK level where their identity is created and stored. By creating these identities, it is used in confined access and correct authentication.

The users are supposedly, not to know the logic of any application program, the logic and behind the code has to be encapsulated and hidden. Express, JS framework is used here, implementing the Representational State Transfer (REST). In this fabric network, composers are used as interface for interacting with the functions.

Inherently blockchain technology is a distributed network of nodes preserving the ledger and its state. The heavier the data gets, because of the number of copies made on all the nodes, longer the transaction becomes, ending up in slowing the network. Alternatively, instead of shifting the whole database on the blockchain, it is better to find a better methodology to have access and to maintain the data probity.

Between the blockchain network and layer of database, a dataless framework is Verity that makes the assertion that it can eliminate illicit data manipulation while maintaining data privacy. It is fairly inferred that the centralized databases are classified in SQL DBMS category. Establishment of a connection between the data stored on the ledger and the existing tuple on the database is done in order to use the existing database with the blockchain network. Every tuple of data in the DBMS generates metadata about it. The primary key's value and the generated metadata are both put into the ledger. As a result, whenever a data tuple is in dispute, the subsequent procedures are carried out: (1) The data tuple's metadata are formed, (2) Using the primary key's value, Ledger is accessed to find the metadata of the tuple in dispute, (3) The retrieved metadata is verified.

5 Evaluation

In a decentralized system it gets difficult to test due to a number of reasons: establishing and instantiating components on multiple machines takes time and effort; in the course of implementation and operation, connections can get disturbed; some parts may stop responding when a specific set of operation is performed; and when a component fails, it is frequently necessary to examine the entire flow and interrelations of all the sequences. Testing a blockchain network's performance presents unique difficulties because the appropriate metrics and analysis tools are still in the early stages of development. In contrast, measuring performance poses new difficulties because it's crucial to develop test suits that cover every component and replicate an actual condition in which each one is put under duress. The further work offers a choice of methods, techniques and metrics to evaluate the effectiveness for the implemented models.

Hyperledger Caliper is a tool that provisions ways to test blockchain use-cases. This is a performance benchmark framework. Users can visualize the performance of the various case though the transaction rate, throughput, etc. The transaction throughput is valid transactions that are committed against the total time taken for the transaction. The throughput latency is calculated by confirmation time—submission time. The transaction policies define the number of peers that need to reach agreement on a transaction which can even further be categorized as normal and extreme case.

The endorsement peers were shuffled between 2, 4, 6, 8, organization between 1 and 10 and deployment type involved single and multi-hosts. The throughput of the network sees growth linearly as the arrival rate keeps increasing. This increase happens only till the optimal value (Fig. 3).

In the network, the average latency is seen to be remarkably low. This is due to quick processing of transactions which sees appreciably almost no to less waiting time. Also, a significant rise in the average latency is seen as the rate of arrival goes over the optimal throughput. With additional increase in the rate of arrival rise in the latency obviously gets larger. At one point in the network, the processing time of transaction comes nearer to the transaction's arrival time. This is when rate of arrival falls below the optimal threshold. When single host is compared with multi-host architecture, the system executes better efficiency in the later (Fig. 4).

During the evaluation, for all the different send rates and different endorsement policies, the network can always achieve 100% success rates. For every cycle, the success in transaction rate is measured by maximum number of transactions

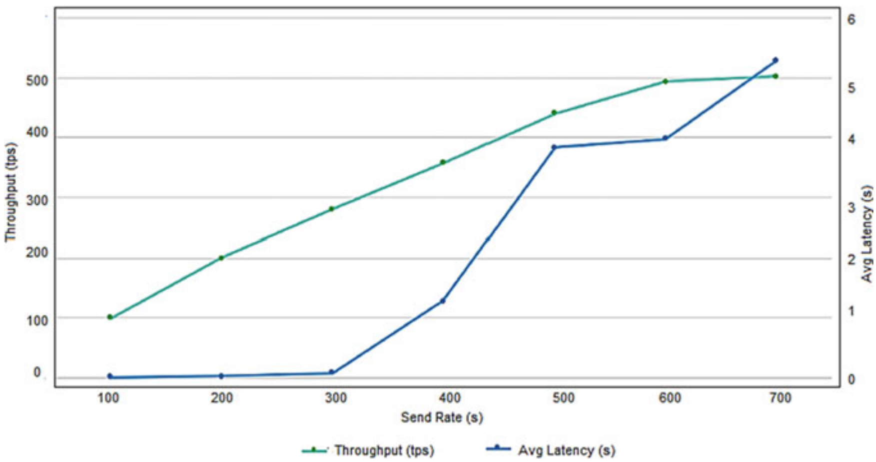


Fig. 3 HL caliper report on read transaction throughput tps

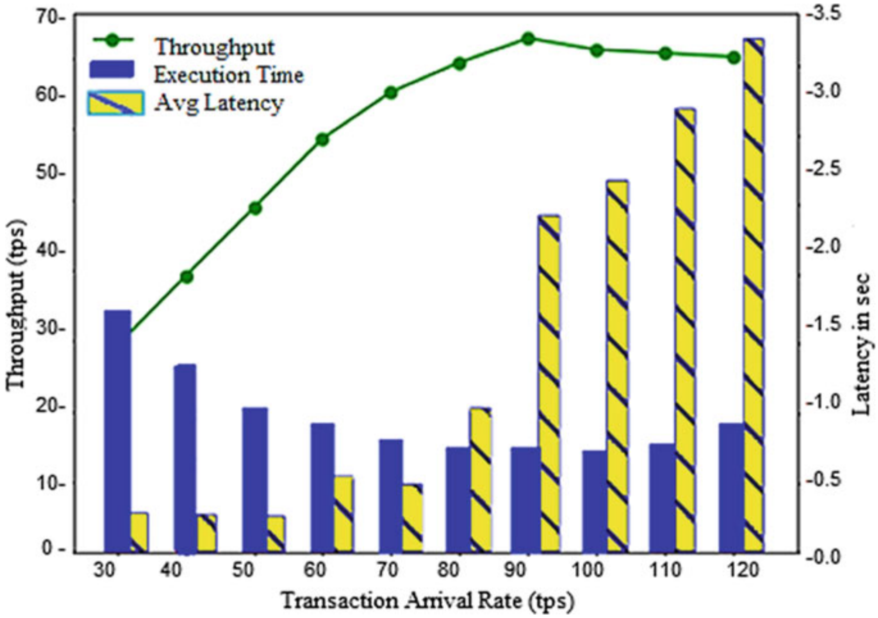


Fig. 4 Transaction throughput and multi-host servers

committed, from all the transactions that were submitted, have been successfully processed registered in the blockchain. An unsuccessful transaction could be due to number of other reasons like network misconfiguration, changes in smart contract or a network connection error.

6 Conclusion

Blockchain has the possible future to revolutionize the world of economics, business, health, supplychain, IoT and other industries. The last few years marked the growth in applications contributed in blockchain on the basis of the immutability, traceability, trust and the data privacy. Blockchain rendered possibility for the users to behave anonymous and make unassailable transactions. However, the prevailing concerns with blockchain needs to be addressed mainly technical, regulatory and privacy issues. The problems with blockchain have a significant impact on privacy. New ingenuity is the need of the moment with the insight of improving privacy and security by making the blockchains acquiescent with the standards and mandate of privacy. Though Blockchain Technology include chaincodes and valid endorsement plan for the privacy techniques and handling open security, but is it enough now? To ensure that blockchain is used in a secure manner, businesses, organizations and government needs to take steps to protect privacy. This includes creating regulations, embedding encryption, other security measures, usage of ZKP being one among them. This paper has discussed the preservation concern of privacy in blockchain and had given a casestudy with implementation of ZKP model.

Bibliography

1. Konkin A, Zapechnikov S (2021) Privacy methods and zero-knowledge poof for corporate blockchain. Proced Comput Sci 190:471–478. https://doi.org/10.1016/j.procs.2021.06.055
2. Bansod S, Ragha L (2022) Challenges in making blockchain privacy compliant for the digital world: some measures. Sādhanā 47(3):31. https://doi.org/10.1007/s12046-022-01931-1
3. Balasubramanian S, Shukla V, Sethi JS, Islam N, Saloum R (2021) A readiness assessment framework for blockchain adoption: a healthcare case study. Technol Forecast Soc Change 165:120536. https://doi.org/10.1016/j.techfore.2020.120536
4. Bernabe JB, Canovas JL, Hernandez-Ramos JL, Moreno RT, Skarmeta A (2019) Privacy-preserving solutions for blockchain: review and challenges. IEEE Access 7:1. https://doi.org/10.1109/access.2019.2950872
5. Bocek T, Rodrigues BB, Strasser T, Stiller B (2017) Blockchains everywhere: a use-case of blockchains in the pharma supply-chain. In: IFIP/IEEE symposium on integrated network and service management (IM), pp 772–777
6. Bryatov S, Borodinov A (2019) Blockchain technology in the pharmaceutical supply chain: researching a business model based on hyperledger fabric. In: International conference on information technology and nanotechnology (ITNT), Samara, Russia
7. Chukleang T, Jandaeng C (2022) Security enhancement in smart logistics with blockchain technology: a home delivery use case. Informatics 9(3):70. https://doi.org/10.3390/informatics9030070
8. Dagher GG, Mohler J, Milojkovic M, Marella PB (2018) Ancile: privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. Sustain Cities Soc 39:283–297
9. Dasgupta D, Shrein JM, Gupta KD (2019) A survey of blockchain from security perspective. J Bank Finan Technol 3(1):1–17. https://doi.org/10.1007/s42786-018-00002-6
10. https://ethereum.org/en/zero-knowledge-proofs/
11. https://en.wikipedia.org/wiki/Zero-knowledge_proof
12. Ghayvat H, Pandya SN, Bhattacharya P, Zuhair M, Rashid M, Hakak S, Dev K (2021) CP-BDHCA: blockchain-based confidentiality-privacy preserving big data scheme for healthcare clouds and applications. IEEE J Biomed Health Inform 14:1. https://doi.org/10.1109/jbhi.2021.3097237
13. Hasan J (2019) Overview and applications of zero knowledge proof (ZKP). IJCSN Int J Comput Sci Netw 8(5):12364
14. Homoliak I, Venugopalan S, Reijsbergen D, Hum Q, Schumi R, Szalachowski P (2021) The security reference architecture for blockchains: toward a standardized model for studying vulnerabilities, threats, and defences. IEEE Commun Surv Tutor 23(1):341–390. https://doi.org/10.1109/comst.2020.3033665
15. https://www.hyperledger.org/use/fabric
16. Medical Sensor Networks. IEEE Internet Things J, p 1. https://doi.org/10.1109/jiot.2021.3117762
17. Morais E, Koens T, van Wijk C, Koren A (2019) A survey on zero knowledge range proofs and applications. SN Appl Sci 1(8). https://doi.org/10.1007/s42452-019-0989-z
18. Goldwasser MSS (1986) Private coins versus public coins in interactive proof systems. In: Proceedings of the eighteenth annual ACM symposium on Theory of computing
19. Peng L, Feng W, Yan Z, Li Y, Zhou X, Shimizu S (2020) Privacy preservation in permissionless blockchain: a survey. Dig Commun Netw. https://doi.org/10.1016/j.dcan.2020.05.008

20. Sahai S, Atre M, Sharma S, Gupta R, Shukla SK (2020) Verity: blockchain based framework to detect insider attacks in DBMS. IEEE Xplore. https://doi.org/10.1109/Blockchain50366.2020.00012
21. Goldwasser SCS (1985) The knowledge complexity of interactive proof systems. In: Proceedings of the 17th ACM symposium on theory of computing
22. Shivers R, Rahman MA, Shahriar H (2019) Toward a secure and decentralized blockchain-based ride-hailing platform for autonomous vehicles. arXiv:1910.00715
23. Singh SK, Park JH, Sharma PK, Pan Y (2022) BIIoVT: blockchain-based secure storage architecture for intelligent internet of vehicular things. IEEE Consum Electr Mag 11(6):75–82. https://doi.org/10.1109/MCE.2021.3089992
24. Wang W, Qiu C, Yin Z, Srivastava G, Gadekallu TR, Alsolami F, Su C (2021) Blockchain and PUF-based lightweight authentication protocol for wireless
25. Singh R, Dwivedi AD, Mukkamala RR, Alnumay WS (2022) Privacy-preserving ledger for blockchain and internet of things-enabled cyber-physical systems. Comput Electr Eng 103:108290. https://doi.org/10.1016/j.compeleceng.2022.108290

Meta-Analysis of Popular Encryption and Hashing Algorithms

Parth Rathod, Riya Sakhiya, Riya Shah, and Shivangi Mehta

Abstract Data is any information that is useful, and we have a variety of resources to protect it. In order to stop illegal access to computers, personal databases, and websites, data security entails upholding digital privacy protections. Confidentiality, authenticity, integrity, and nonrepudiation are all elements of the modern information security system scenario. Cryptography, which offers the capability to safeguard users using various kinds of encryption and decryption methods, is the answer to this specific challenge. We employ compression to protect the data so that it uses fewer bits and takes up less disc space, ultimately costing less money. One of the most essential procedures is encryption, which enables us to conceal the true nature of data so that only the decryption procedure can reveal the original data. Encryption helps to protect data from unauthorized access. The data is encrypted and decrypted using a variety of symmetric and asymmetric key-based techniques. There is another way that has been popularized in recent times which consists of the use of hashing. Hashing offers better security and reliability than encryption-decryption algorithms.

Keywords Cryptography · Hashing · Encryption and decryption · SHA-3 · AES · 3DES · Blowfish

1 Introduction

One of the key concerns for any IT firm is information security. We are utilizing cryptography, which will enable effective data security, to address this type of issue. Organizations need to make sure that there is enough storage capacity to satisfy their needs since in addition to safety, they are also dealing with the problem of rising disc storage costs. With data compression, we can reduce the size of the original message to a smaller format known as a code word, so resolving the problem of rising costs. In cryptography, there are two operations: encryption and decryption, which operate on

P. Rathod (✉) · R. Sakhiya · R. Shah · S. Mehta
Devang Patel Institute of Advance Technology and Research, Changa, Anand, Gujarat 388421, India
e-mail: parth.rathod12@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_17

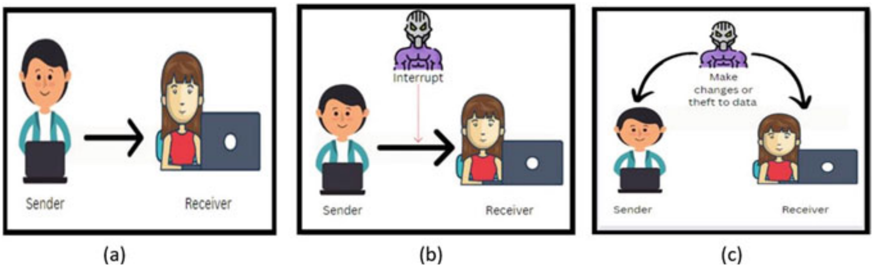


Fig. 1 **a** Encrypted file being transferred to the receiver, **b** Attacker interrupts the transfer process, and **c** Attacker manipulates the message

plain text and cipher text, respectively. By encoding the message or information, we can increase security by limiting access to it to approved entities. This technique is known as encryption. While decryption is taking place, the message or information that has been encoded as a code word is being decoded. Depending on the key values or type, our diverse range of cryptographic techniques can be loosely divided into two categories: symmetric key cryptography and asymmetric key cryptography (Fig. 1).

The other broad category of algorithms consists of hashing techniques. In our analysis of these papers, we found various insights about MD5, SHA-1, SHA-2, and SHA-3 along with its different architectures as listed by NIST. Various Research papers included in this review paper include using hash along with data encryption techniques and finding out which ones are most suitable in terms of time analysis and CPU utilization.

2 Various Algorithm

2.1 3DES Algorithm

Due to the enormous volume of transactions that occur daily, cryptography is becoming increasingly important. One such symmetric block cipher technique is Data Encryption Standard (DES) wherein a 64-bit key is used to encrypt plain text of 64-bit into cipher text of 64-bit (Fig. 2).

2.2 AES Algorithm

The Advanced Encryption Standard algorithm, previously named as Rijndel algorithm, is a block cipher-based encryption method that was introduced by the National Institute of Standards and Technology in 2000. A new architecture method has been introduced to decrease the complexity of the AES algorithm when applied to the

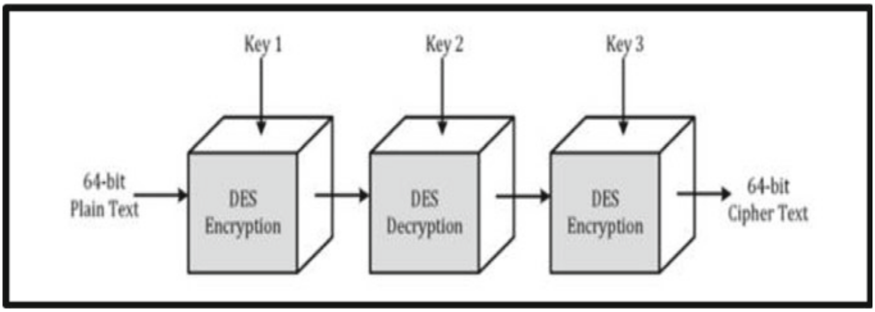


Fig. 2 Block diagram of 3DES

Table 1 AES key length versus rounds

Key length (bits)	Number of rounds
128	10
192	12
256	14

hardware devices such as mobile phones, smart cards. This method focuses on some important features such as SubBytes and Mixcloumn module to generate cipher text (Table 1).

2.3 Blowfish

Blowfish was created by Bruce Schneier is a 64-bit which uses a range of key lengths between 32 and 448 bits and has a predefined value of 128 bits (Fig. 3).

2.4 Hash Functions

A mathematical function that goes one-way by taking input data of arbitrary size and outputs a fixed-size value typically used for authentication, verification, data retrieval, and security purposes.

SHA-0 and SHA-1

SHA-0 and SHA-1 were introduced in 1993 and 1995, respectively, by the National Institute of Standards and Technology which comes under the National Security Agency in the United States. SHA-0 and SHA-1 were created to generate a 160-bit

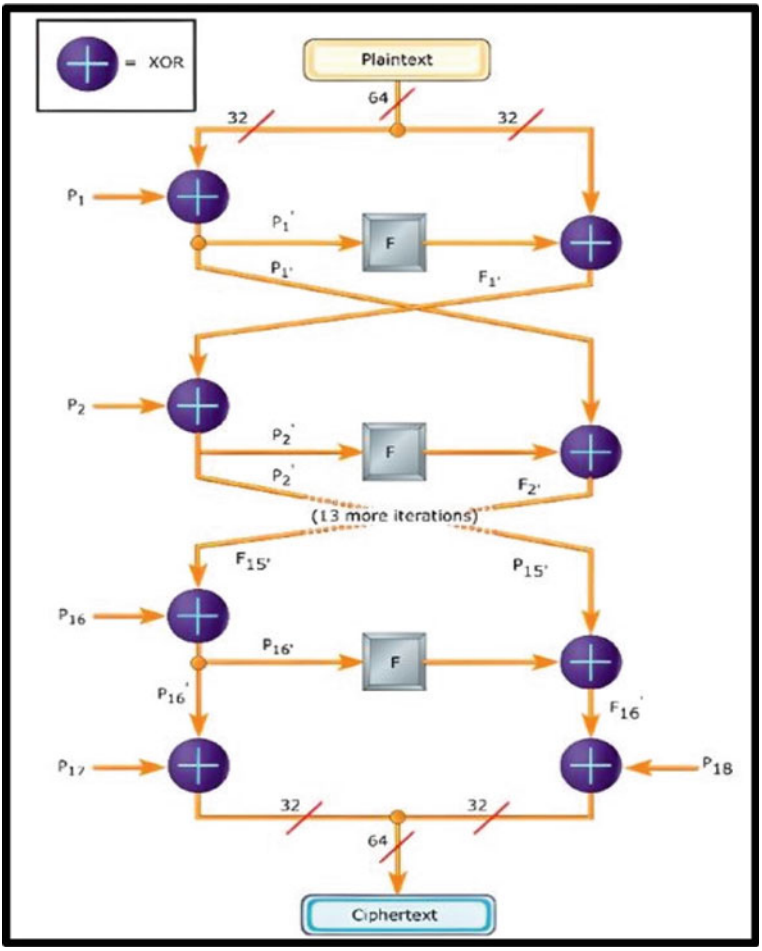


Fig. 3 Blowfish architecture

message digest. But in 1995, flaws in SHA-0 that made it vulnerable to attack were found, and NIST subsequently withdrew support for it which led to the creation of SHA-1.

SHA-2

The National Institute of Standards and Technology (NIST) announced the Secure Hash Algorithm-2 (SHA-2) family of cryptographic hash functions in 2001 which were developed by the National Security Agency. The six hash algorithms that together make up SHA-2 each provide message digests of different size: SHA-224, SHA-256, SHA-384, SHA-512, SHA-512/224, and SHA-512/256. The functions construct the fixed-size message digest by accepting input of any size and processing it using a sequence of mathematical procedures.

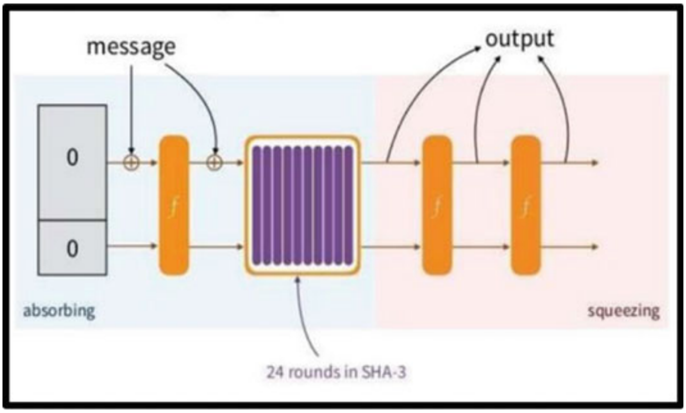


Fig. 4 Keccak sponge construction [1]

SHA-3

The National Institute of Standards and Technology (NIST) published SHA-3 in 2015 after choosing it as the 2012 hash function competition winner of the 2012 hash function competition. As explained by Sharma and Khanum [1] it consists of four hash functions SHA3 has variants 224, 256, 384, and 512. Each of these techniques produces a different size of the message digest. The architecture of SHA-3 is based on the Keccak Family which is based on random permutation which accepts any amount of data also called absorbing and outputs any amount of data also called squeezing and acts as a pseudorandom function allowing it to have great flexibility (Fig. 4) [1].

3 Performance Analysis

The metrics used for evaluation of different algorithm to compare its performance with algorithms such as Blowfish, 3DES, DES, and AES were battery power usage, data block sizes, varying encryption-decryption rate, and various key sizes. It was discovered that there would be no difference in the output of the AES algorithm for the audio, video, and text-related documents when the data format included hexadecimal or 64 bit encodings. In the study conducted by Alanazi et al. [2] summarized the important properties in the Table 2 [2].

Research conducted by Karthik and Muruganandam [3] showed the performance standards for several of the most well-known cryptographic algorithms as shown in Table below. These outcomes are useful for providing context for the comparison results that were presented. It has been demonstrated that Blowfish and AES perform

Table 2 Comparison between AES, 3DES, and DES [2]

Factors	AES	3DES	DES
Possible keys	2^{128}, 2^{192}, or 2^{256}	2^{112} or 2^{168}	2^{256}
Cryptanalysis resistance	Strong against differential, truncated differential, linear, interpolation, and square attacks	Vulnerable to differential, Brute-force attacker could be analyze plain text using differentia; cryptanalysis	Vulnerable to differential and linear cryptanalysis; weak substitution tables
Security	Considered secure	One only weak which exist in DES	Proven inadequate
Time required to check all possible keys at 50 billion keys per second	For a 128-bit key: 5×10^{21} years	For a 112-bit key: 800 Days	For a 56-bit key: 400 Days
Block size	128,192 or 256 bits	64 bits	64 bits

the best among the competition. Also, it is well-known that neither of them has inferior encryption than the other two, making them stronger against data attacks (Figs. 5 and 6) [3].

Hundred and twenty-eight bits key is used by 3DES, which means that there are 2^{128} possible keys. Brute-force attacks, which involve trying every possible key, are not practical with such a large key space, as it would take an infeasible amount of time to test all keys. The time taken to decrypt 3DES encryption can be reduced by using various forms of other attacks. For example, a meet-in-the-middle attack

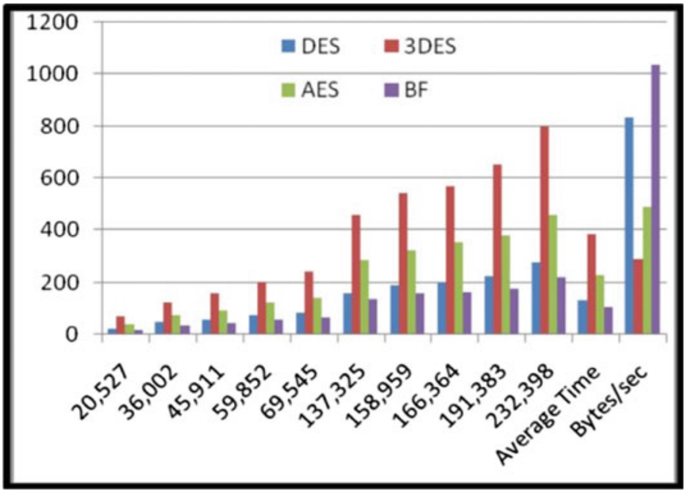


Fig. 5 Comparative execution time in encryption algorithm system 1 [3]

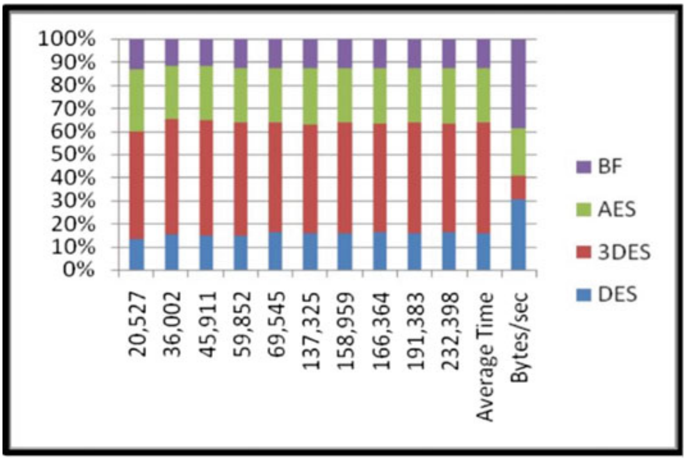


Fig. 6 Comparative execution time in encryption algorithm system 2 [3]

can be used to break 3DES with an effective key length of 112 bits in about 2^{112} operations, which is still infeasible to perform with current technology.

3.1 ECB and CBC Modes of Application

In Electronic Code Book (ECB) the input plain text is broken into numerous blocks. It is a simple method of processing a string of message blocks that have been listed in order, and it mostly employs symmetric key encryption. In ECB each block has been encrypted individually using the encryption key and the same goes for the decryption process. Cipher block chaining (CBC) differs from ECB mode as it hides away patterns in the plaintext. To achieve this, it uses fixed-size initialization vectors which are XOR-ed with the first plaintext block (B1) to introduce randomization before encryption. By making use of block chaining, it performs XOR of each subsequent plaintext block with the ciphertext of the previous block.

Experiments conducted by Karthik and Muruganandam [3] compared the processing times of the Blowfish, AES, and DES algorithms, and it was discovered that Blowfish outperformed the others. However, it demonstrates that AES uses more resources when the size of the data block is relatively large. Worm flaws in the security mechanisms of DES and 3DES are well-known; Blowfish and AES, on the other hand, have none yet. These results have little to do with the other loads on the computer, even though each experiment was conducted multiple times and produced results that were substantially identical to those anticipated (Fig. 7) [3].

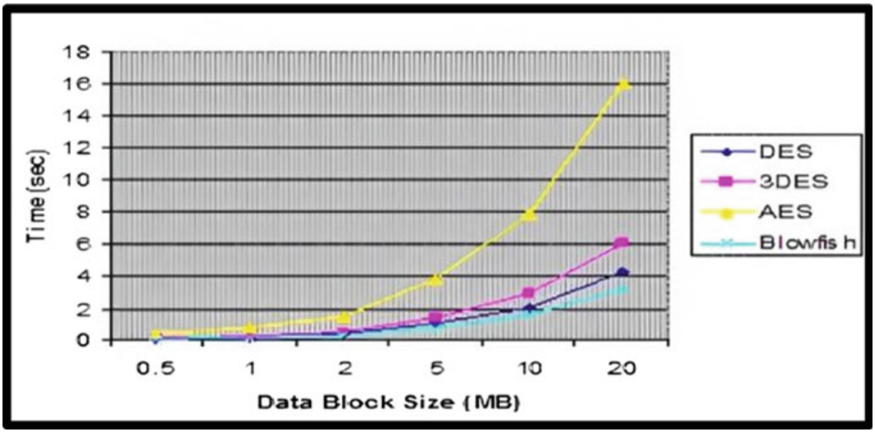


Fig. 7 Performance results with ECB mode [3]

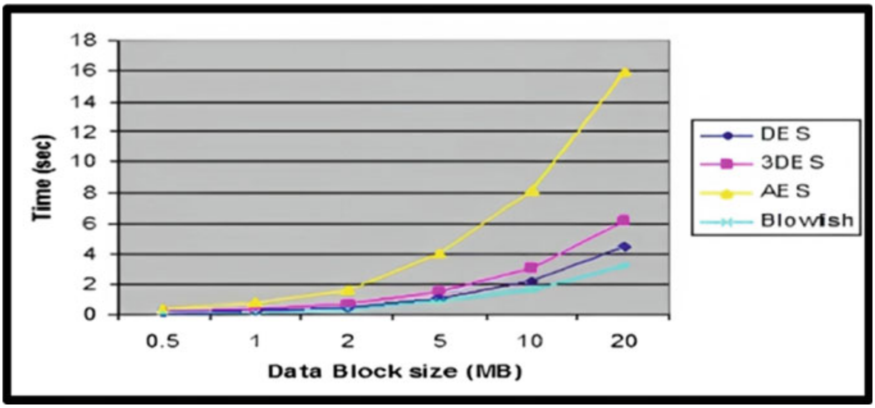


Fig. 8 Performance results with CBC mode [3]

In terms of protection, CBC is superior to ECB. A comparatively little 0.059896 s separate ECB and CBC on average. As a result, we can draw the conclusion that the two operating modes are virtually identical and that discrepancies can be overlooked in practice (Fig. 8) [3].

3.2 Performance of Hashing Algorithms

A study by Vadhera and Lall [4] on SHA and its Variants, went into the attacks that SHA and its family can be exposed to. Due to the algorithm having a block and iterative structure that lacks extra final stages, the function becomes susceptible to partial-message collision and length-extension attacks. There were four security issues in SHA-1, including the possibility of a mathematical vulnerability, which suggests the need for a new and more robust hash function. Even though SHA-2 and SHA-1 have certain similarities, SHA-2 has not yet shown any vulnerabilities to these attacks. To obtain the optimal hash algorithm, they calculated the time needed for each of these algorithms and found the one that would take the least amount of time to compute the hash value. Important insights are given in the following Table 3 [4].

In a study published by the National Institute of Standards and Technology [5] which is responsible for publishing an official series of standards and guidelines. They analyzed permutation-based SHA-3 algorithms of the Keccak family. They summarized different variants of SHA-3 with the most common attack schemes such as collision attack, preimage, and 2nd preimage attacks. We can see that compared to SHA-2, SHA-3 is at par or even superior in resisting the attacks found in SHA-2 (Table 4) [5].

Further performance analysis was conducted by Yoshida and Biryukov [6] where they performed a differential attack on SHA-256 XOR and SHACAL-2-XOR and tried to break the hashes. Pseudo-collision resistance is a vulnerability of SHA-2-XOR with up to 34 cycles. To perform an attack, a 31-round distinguisher was used for up to 32 rounds on SHACAL-2-XOR. They further went on to demonstrate that the probability is $> 2^{-16}$ in the absence of 2-round iterative patterns. Resistance to discovering a collision resulting from more flexible conditions that allow for selecting several beginning vectors is known as pseudo-collision resistance. An important conclusion of the study was that for a scenario where a hash function is created using the MD construction, pseudo-collision resistance becomes significant. This is because collision resistance can be converted into pseudo-collision resistance for the compression function. Their study showed SHACAL-2 with different attack types with their time and memory requirements (Table 5) [6].

The next study was conducted to improve data encryption algorithms such as AES, since the practical implementations of hashing algorithms are limited. The research was carried out which went into the details of combining data encryption methods like AES along with hashing methods like SHA-256 for securing the transfer of a multimedia file over chat application such as WhatsApp. This study was performed by Fauziah et al. [7]. They used AES to encrypt, and the key was generated with SHA-256. They found that the encrypted and decrypted files had expanded in terms of memory, was due to the increase in blocking length found in the header cipher text file and was caused by the usage of extension and encryption operations. The increase in file size of the original file after decryption was on average about 22% (Tables 6 and 7) [7].

Table 3 Performance of various SHA algorithms [4]

Algorithm and variant		Output size (bits)	Internal state size (bits)	Block size (bits)	Max message size (bits)	Word size (bits)	Rounds	Bitwise operations	Collisions found	Performance (MiB/s)
Md5 (reference)		128	128	2^{64}–1	32	64	64	And, or, xor, rot	Yes	335
SHA-0		160	160	2^{64}–1	32	80	80	And, or, xor, rot	Yes	–
SHA-1		160	160	2^{64}–1	32	80	80	And, or, xor, rot	In theory	192
SHA-2	SHA-224	224	256	2^{64}–1	32	64	64	And, or, xor, rot, shr	None	139
	SHA-256	256								
	SHA-384	384	512	2^{128}–1	64	80	80	And, or, xor, rot, shr	None	154
	SHA-512	512								
	SHA-512/224	224								
	SHA-512/256	256								

Table 4 NIST published strengths of various hashing algorithms [5]

Function	Output size	Security strengths in bits		
		Collision	Preimage	2nd Preimage
SHA-1	160	< 80	160	160-L(M)
SHA-224	224	112	224	Min(224, 256-L(M))
SHA-512/224	224	112	224	224
SHA-256	256	128	256	256-L(M)
SHA-512/256	256	128	256	256
SHA-384	384	192	384	384
SHA-512	512	256	512	512-L(M)
SHA3-224	224	112	224	224
SHA3-256	256	128	256	256
SHA3-384	384	192	384	384
SHA3-512	512	256	512	512
SHAKE128	d	Min(d/2,128)	$\geq$ min(d,128)	min(d,128)
SHAKE256	d	Min(d/2,256)	$\geq$ min(d,256)	min(d,256)

Table 5 Time and memory consumption of various attacks on different Hashing Architectures like SHACAL-2 [6]

Attack type	#R	Data	Time	Memory
Impossible differential attack on SHACAL-2	30	744CP	$2^{495.1}$	$2^{14.5}$
Differential-linear attack on SHACAL-2	32	$2^{43.4}$CP	$2^{504.2}$	$2^{48.4}$
Related-KEY RECTAngle attack on SHACAl-2	37	$2^{43.}$2RK-CP	$2^{484.95}$	$2^{238.16}$
Distinguisher attack on SHACAL-2-XOR in this paper	31	2^{248}CP	2^{248}	2^{22}
Differential attack on SHACAL-2-XOR in this paper	32	$2^{243.3}$CP	$2^{246.3}$	

#R, # of rounds; CP, chosen plaintexts; RK-CP, related-key chosen plaintexts; Time, encryption units; Memory, bytes of memory

Table 6 Comparison of file sizes before and after using AES [7]

File name	Original file size	Encrypted file size	Decrypted file size
Kalimba.mp3	8,415,232	9,898,557	9,437,184
Sleep Away.mp3	4,843,520	5,494,538	5,242,880
Tulips.jpg	621,568	1,075,200	1,048,576
Chrysantemum.jpg	879,616	1,075,200	1,048,576
Wildlife.wmv	26,246,144	27,674,624	27,262,976

Table 7 Time (in seconds) to encrypt and decrypt files using Hashing [7]

File name	Encrypted time	Decrypted time
Kalimba.mp3	7.3	9.1
Sleep Away.mp3	5.4	6.4
Tulips.jpg	2.6	2.9
Chrysantemum.jpg	2.3	2.7
Wildlife.wmv	21.0	23.9

4 Conclusion

From all the studies which we included in our review paper, we found out that the most effective algorithm is SHA-3. Its speed along with the ability to provide security against attacks like collision attacks, the preimage, and 2nd preimage attacks which hashing algorithms are most susceptible to, but its lack of practical applications is what holds it back. Practically speaking, AES when compared with DES and 3DES, increases the time taken for the encryption process but offers far greater security and reliability. It is also used in other cryptographic protocols, including the Transport Security Layer protocol and the Socket Security Layer (SSL), which greatly enhances the security of communications over the Internet. There are some innovative approaches to increase the security and credibility of AES by generating the key and encrypting it with the SHA algorithm. Hence, we can say both algorithmic methods would coexist, and creative solutions are to be envisaged.

References

1. Sharma S, Khanum S (2022) Performance analysis of SHA 2 and SHA 3. In: Proceedings of the international conference on innovative computing and communication (ICICC)
2. Alanazi HO, Zaidan BB, Zaidan AA, Jalab HA, Shabbir M, Al-Nabhani T (2010) New comparative study between DES, 3DES and AES within nine factors. J Comput 2(3):4085
3. Karthik S, Muruganandam A (2014) Data encryption and decryption by using triple DES and performance analysis of crypto system. IJSER 2(11):24–31
4. Vadhera P, Lall B (2014) Review paper on secure hashing algorithm and its variants. IJSR 3(6):629–632
5. Dworkin MJ (2015) SHA-3 standard: permutation-based hash and extendable-output functions. NIST FIPS
6. Yoshida H, Biryukov A (2006) Analysis of a SHA-256 variant. In: Preneel B, Tavares S (eds) Selected areas in cryptography. SAC 2005. Lecture notes in computer science, vol 3897. Springer, Berlin, Heidelberg, pp 245–260
7. Fauziah NA, Rachmawanto EH, Moses Setiadi DRI, Sari CA (2018) Design and implementation of AES and SHA-256 cryptography for securing multimedia file over android chat application. In: International seminar on research of information technology and intelligent systems (ISRITI). Yogyakarta, Indonesia, pp 146–151

Cloud-Based Internet of Things Architecture for Hydroponics Farm Automation

Kunal Kulkarni, Vinaya Gohokar, Pradnya Kulkarni, and Trushita Chaware

Abstract Due to rising food prices, labor costs, unfavorable environmental conditions, and dwindling agricultural land, indoor growing techniques like hydroponics and aeroponics are becoming more and more popular. It is carried out in a controlled environment and requires precise control of various parameters to improve yield and quality. This paper proposes a cloud-based Internet of things architecture for controlled environment agriculture. Farmers can control the nutritional requirements, pH values, water level, and humidity to assure that the plants are getting the exact amount of nutrients they need. The plants are cultivated in an indoor setup; thus, it becomes easy to control temperature and light exposure to improve plant production and have better quality yield. A detailed discussion on the nutrient requirement and measurement is also presented in the paper.

Keywords IoT · Hydroponics · Vertical farm · Sensors · Mobile apps

1 Introduction

Section 1

The world population is expected to reach 9.6 billion by 2050. This represents a serious problem for the agricultural sector [1]. Despite common obstacles such as bad weather, adverse climate change and impacts on agriculture, food demand is getting increasingly out of control [2]. Researchers are starting to look for smart IoT

K. Kulkarni (✉) · V. Gohokar · P. Kulkarni · T. Chaware
School of ECE, Dr. Vishwanath Karad MIT-WPU, Pune, India
e-mail: kunalkulkarni2599@gmail.com

V. Gohokar
e-mail: vinaya.gohokar@mitwpu.edu.in

P. Kulkarni
e-mail: pradnya.raykar@gmail.com

T. Chaware
e-mail: trushita.chaware@mitwpu.edu.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_18

technologies for agriculture to meet this growing need. Traditional farming uses huge amounts of resources. The practice of growing plants indoors is known as controlled environment agriculture (CEA), also referred to as climate and weather-independent farming or, more frequently, indoor farming. While greenhouses have been used for centuries, indoor farming is a more modern development that further simplifies the growing process by getting rid of all the extraneous parts of conventional farming. We can remember the essential components for plant development as energy, nutrients, water, and CO_2 as per the photosynthesis process. Following this fundamental formula, CEA eliminates all extraneous inputs that have come to be required for our existing agricultural system, such as soil and insecticides [3]. Seeds are sown instead of soil, using a soilless growing medium such as coconut husks to provide a surface for the seedlings to attach their roots to. This soilless method provides a much cleaner and easier process by reducing the risk of insects and weeds entering the growing environment.

Hydroponics can be utilized to grow a range of flowers, vegetables, and herbs. Farmers are now able to manage hydroponic systems easily because of the development of IoT. IoT can control and monitor pH of water, water level, flow, temperature, and light intensity [4–6]. For instance, in some regions during the winter, the water tends to freeze, which can completely impede the cultivation process. The hydroponics farm's water temperature sensors can detect the temperature loss and notify the grower as needed. Like this, the pH sensor can determine when nutrient levels have changed and can then pump in the necessary quantity of minerals [3].

In the study by Rahimi et al. [7] a highly secure communication and control system in the sensors and the IoT platform is presented. The MFA security system developed for this IoT platform has multiple steps that required to be done before data can be uploaded to the cloud database. Things Sentral IoT platform is used for implementation. Work published by Ludwig and Fernandes details of plant grown in a hydroponic setup was monitored for pH, conductivity, and luminosity, and information was sent to a microcontroller. Here, a microcontroller showed the status of the parameters being watched and used a relay switch to regulate the lighting for the plant [8]. The automation of irrigation system using cost effective sensors is presented in [9]. An automatic calibration pH sensor is used to monitor and adjust pH values of the nutrient solution in hydroponic system. The work published in [10] presents IoT network which improves reliability and allow remote monitoring and control. Variety of sensors, actuators, and microcontrollers are used in the development of IoT-based architectures. The microcontrollers include Arduino Uno, Raspberry pi, ESP32, NodeMcu. There are systems using cloud for saving data and take control action based on the requirements.

The research presented in [11] focuses on Edge-Based Autonomous Management of Vertical Farms, which automates vertical farming and improves crop yields using LPWAN technology. The research presented in [12] describes an Automated Hydroponic System that uses ThingSpeak and Arduino microcontrollers for automation, pH, and EC sensors to measure nutrient levels in a deep water culture hydroponic system, and an Automated Hydroponic System. The study published in [10] describes a Fully Automated Hydroponic System for Indoor Plant Growth, which uses Arduino microcontrollers, Raspberry Pi, and various sensors to monitor and control the environment for optimum plant growth.

The structural flow of this paper is as follows:

- Basic building blocks of hydroponic-based system, Nutrient requirements as per the crop and its analysis is presented in Section 2.
- In Section 3, discussion is given on Sensors used and their characteristics, interfacing of the sensor with microcontroller is also presented.
- Section 4 details the proposed architecture, circuit diagrams, interfacing details, and flowchart of the system.
- Section 5 presents observations, and concludes the paper with future scope of the work.

2 Basic Building Blocks of Automated Hydroponic System

Section 2

There are various structures used for growing hydroponic systems including Deep water culture (DWC), Drip, Nutrient film Technique (NFT). Every technique has its own pros and cons. NFT is a highly accepted technique of all and is used in this research for the implementation of automation in hydroponics. Dr. Alain Cooper discovered this technique to overcome the flaws of the flow and the ebb systems. Nutrient solutions are continuously pumped across channels in which the plants are kept. When the solution reaches the terminal of the channel, it is reverted. This recirculates the system. The plant's roots are not completely submerged in water [13]. In hydroponic systems, substrates like Rockwool, Grow Rock, Coco Fiber are used as a growing medium to support plant roots and provide them with the necessary nutrients and oxygen. A substrate is selected depending upon the plant. Coco Fiber is a commonly used substrate. It decomposes very slowly, retains moisture. It is neutral to pH.17 vital elements are required for plant development and reproduction, out of which carbon, oxygen, and hydrogen are the primary elements. The remaining 14 elements are classified into two groups:

1. Macro-Elements: These are required by plants in relatively large amounts and consist of Nitrogen (N), Phosphorus (P), Potassium (K), Calcium (Ca), Sulfur (S), and Magnesium (Mg).
2. Micro-Elements: These are required by plants in relatively small amounts and consist of Manganese (Mn), Iron (Fe), Zinc (Zn), Boron (B), Copper (Cu), Chlorine (Cl), Molybdenum (Mo), and Nickel (Ni).

In hydroponics, the nutrient solution can be customized based on the plant's growth stage and specific needs by adjusting the levels of the primary and secondary elements. For example, during the vegetative stage, plants require more nitrogen, so a high nitrogen formula, such as the Grow formula, is used. During the flowering stage, plants require more phosphorus and potassium, so a high phosphorus

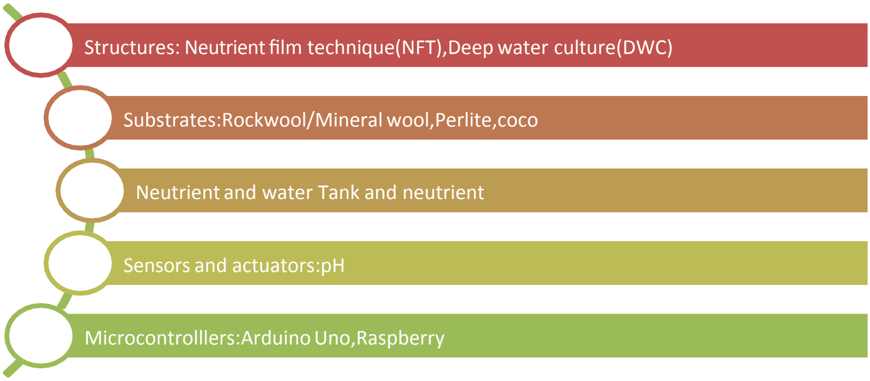


Fig. 1 Basic building blocks of automated hydroponic system

and potassium formula, such as the Bloom formula, is used. By controlling the nutrient solution, hydroponic growers can ensure optimal plant growth and yield. Other important parameters include-pH Level, TDS, Dissolved Oxygen, Lighting, Temperature, Substrate, Electric conductivity (Fig. 1).

3 pH Level

The pH of any solution indicates hydrogen ion concentration (H) in a solvent. Higher the concentration of hydrogen ions, the weaker the solvent (more acidic). At 25 °C, a pH value 7 is considered to be neutral, < 7 is acidic, alkaline above 7. The availability of nutrients for plants is regulated by proper pH value. An unmonitored pH hampers absorption of nutrients and can also lead to symptoms indicating deficiency in plants. Plants having pale and yellow leaves is a sign of iron deficiency. Calcium deficiency leads to burnt tips of leaves or cup-shaped leaves. The pH value requirements are different for plants cultivated in hydroponic techniques as compared to soil cultivation. In hydroponic systems, one must constantly monitor the pH value. 5.5 and 6.0 is an optimal pH value range for growing a majority of plants in hydroponics. An incorrect pH can be either a very high or very low quantity of a particular nutrient. Low pH (< 5) can cause calcium and magnesium deficiencies or iron and copper poisoning. However, a pH more than 6.0 or 6.5 will cause deficiency of iron [14] (Fig. 2).

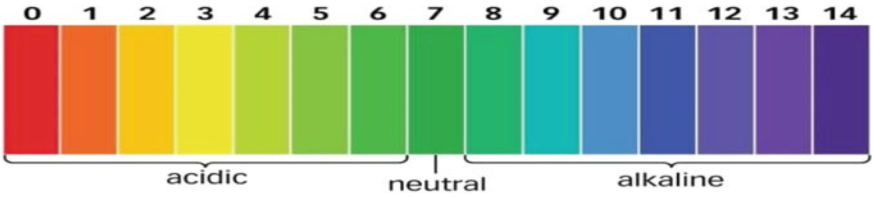


Fig. 2 pH level

4 TDS

Total Dissolved Solids (TDS) quantifies the concentration of a nutrient solution. It is denoted in values like in parts per million (PPM) or conductivity (EC), where the value is in Milli Siemens (mS) or micro-Siemens (μS). It is a measure of how many salts and minerals are dissolved in a solution. Hard water contains more minerals and soft water has less. For reference, clean drinking water is below 200 ppm, and above 400 ppm is not safe for human drinking. A plant needs < 20 elements to endure. However, the absorption of said elements is possible from the roots within a certain pH range. This range varies as it is 5.5–6.5 for hydroponics. It is also significant to consider the strength of the elements. Even at a correct pH level, the elements will not be absorbed if they are too strong. Both the pH and TDS have to be monitored as an imbalance may cause damage to the plants [15]. As with pH levels, PPM should be monitored throughout the growth cycle. As your PPM goes down, add more nutrients to your system. A high PPM level is harmful for plants. One must add the nutrients one time in a week. PPM meters and EC meters both measure nutrient concentrations in a similar manner and accuracy, however, they display different readings. Both EC and PPM devices measure the concentration by measuring the current that flows in two metal probes. This is based on the principle that a nutrient solution with a high concentration of nutrients will have higher electrical conductivity [15, 16].

5 Dissolved Oxygen (DO)

Oxygen is crucial for plants. Root systems require oxygen for aerobic respiration, which is an essential process in plants that releases energy for growth of roots. The oxygen demand of plants in the flowering state is generally higher than in the vegetative state. This is not related to a specific growth stage, but to the root system size, temperature, and rate of nutrient uptake. Oxygen insufficiency leads to wilting of the plant roots even in presence of light and appropriate weather conditions. Oxygen decreases the permeability of water in the roots. Due to this toxin proliferate causing insufficient absorption of water and minerals which support plant survival. The ability of an aqueous solvent to retain DO decreases with an increase in the temperature [17].

6 Lighting

Lighting is an essential element in hydroponics, as it provides the necessary energy for plant growth and development. The intensity, duration, and spectrum of light influence plant growth, flowering, and fruiting. In hydroponics, artificial lighting is usually used to supplement natural light or replace it entirely. LED grow lights are a popular choice for hydroponic lighting because they are energy-efficient, long-lasting, and emit the right spectrums of light that plants need for optimal growth. Proper lighting in hydroponics can help decrease the time required for plants to mature, increase yields, and improve overall crop quality.

Short-day plants: Need long periods of darkness for photosynthesis and flowering. It will not bloom if exposed to light for > 10–12 h a day. Examples for short-day plants cauliflowers, strawberries, and chrysanthemums. Long-Day Plant: Requires up to 18 h of sunlight per day. These include wheat, potatoes, lettuce, spinach, and beets. Day Neutral Plant: The most flexible. It bears fruit regardless of the amount of light. Examples are rice, eggplant, corn, and roses. Lighting systems: bulbs, reflector hoods, ballasts, and timer unit.

Section 3

Figure 3 illustrates the proposed cloud-based architecture for hydroponic system automation. The main building blocks of the architecture are physical layer consisting of sensor subsystem, processing unit and actuators. Data acquisition, preprocessing is done in this layer. Network Layer: It consists of communication design between microcontroller and internal server as well as cloud. Application Layer: Consists of mobile application/web page for visualization and control of parameters. Following part of the section details about the sensor selection, testing, and interfacing with the microcontroller used in the design (Fig. 4).

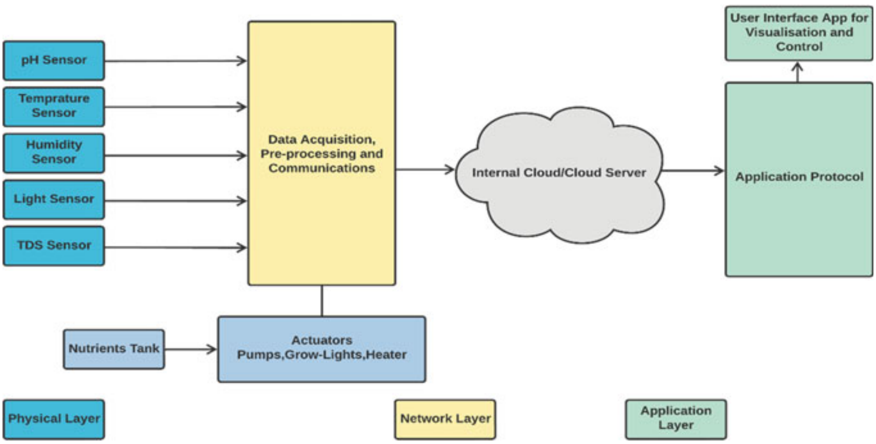


Fig. 3 Is an IoT-enabled hydroponics system

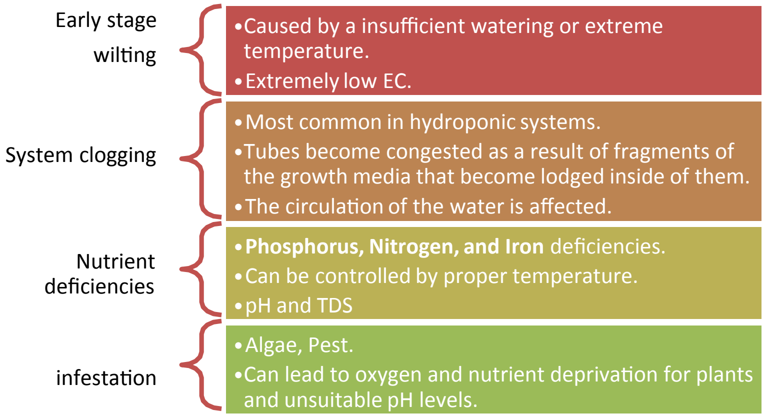


Fig. 4 Challenges in automation of hydroponics

7 Sensor Subsystem

In order to closely monitor parameters to prevent or address any issues that may arise as presented in Fig. 5, sensors play a crucial role. This section discusses about various sensors used in the implementation of the project, the low-power system-on-a-chip (SoC) microcontroller is used ESP32, which is a microcontroller designed by Espressif systems. It is a successor to the popular ESP8266 microcontroller and features a dual-core Tensilica LX6 processor, Wi-Fi and Bluetooth connectivity, and a range of peripheral interfaces. Development Environment: Espressif IoT Development Framework Supports: Arduino and Micro Python programming as well.

The sensor subsystem consists of 5 sensors which includes pH Sensor, Temperature Sensor Probe, Humidity and Temperature Sensor, Light Sensor, TDS Sensor

Fig. 5 Interfacing diagram of TDS sensor

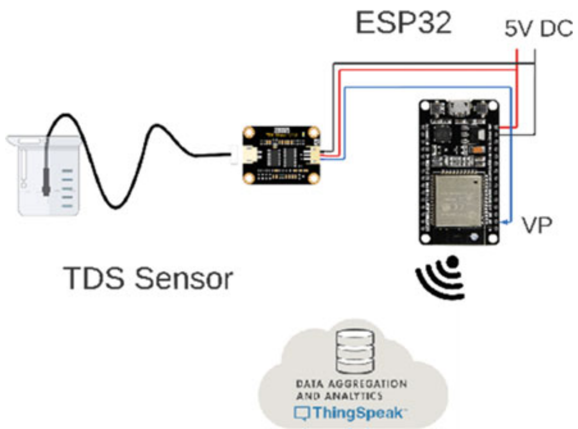


Table 1 Sensors used in the project

Parameters	pH-Probe PE-03	TDS-TDS V1.0	DS18B20	DHT22
Operational Voltage	5 V-DC	5 V-DC	3–5 V-DC	5 V-DC
Current	5 ~ 10 mA	3 ~ 6 mA	1 mA	2.5 mA
Range	0–14 PH	0–1000 PPM	− 55 to 125 °C	Temp-40–80 °C Humidity-0–100%RH
Accuracy	0.01 pH	2%	± 0.5 °C	± 0.5 °C
Manufactured by	Erma INC.Tokyo Japan	Keyesstudio	Bombay electronics	Bomb ay electronics
Purchase link	pH Sensor module	TDSMeter	Ds1820	DHT22

which provide the data to the MCU which further processes the data and takes required actions and updates the cloud/server. The actuators that are Pumps, Heater, Lights are controlled according to the analyzed sensor data by the MCU. Table 1 lists and gives brief about the sensors used in the project.

8 Testing and Interfacing

TDS Sensor

Total Dissolved Solids (TDS) refers to the concentration of inorganic salts and other dissolved substances in water. A TDS sensor can measure the amount of dissolved nutrients in the water, which is important for maintaining the proper nutrient balance in a hydroponic system. The following Fig. 6 shows the Interfacing of TDS Sensor with ESP-32 Microcontroller. As it is an analog sensor the output pin of TDS sensor is connected to A0 pin of ESP32.

9 pH Sensor

This type of sensor measures the alkalinity or acidity of a solution, such as the water in a hydroponic or aquaponic system. It works by measuring the concentration of (H^+) ions that are hydrogen ions in the solution, which determines the pH value. The pH sensor can help ensure that the water pH level is within the optimal range for plant growth. Figure 7 represents the interfacing of pH sensor with ESP-32 microcontroller. Testing and calibration of pH sensor probe using tap water with known pH value of 6.5 and lemon juice with known pH value of 2 is shown Fig. 7.

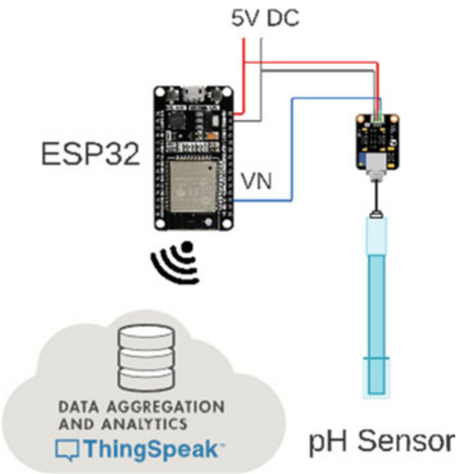


Fig. 6 Interfacing diagram of pH sensor

Fig. 7 **a**, **b** Readings before and after adding nutrients and pH balancing solution to the reservoir water tank which contains tap water with pH-8 and TDS 120 ppm

DS18B20 Temperature Probe: This is a digital temperature sensor that can be used to measure the water temperature in a hydroponic or aquaponic system. It is waterproof and can be submersed in the water. The DS18B20 temperature probe is a popular choice for temperature monitoring in these types of systems because of its accuracy and reliability. DHT22 Sensor: This type of sensor can measure temperature and humidity in the air surrounding the hydroponic or aquaponic system. It is useful for monitoring the environmental conditions in the growing area, as both temperatures (Figs. 8 and 9).

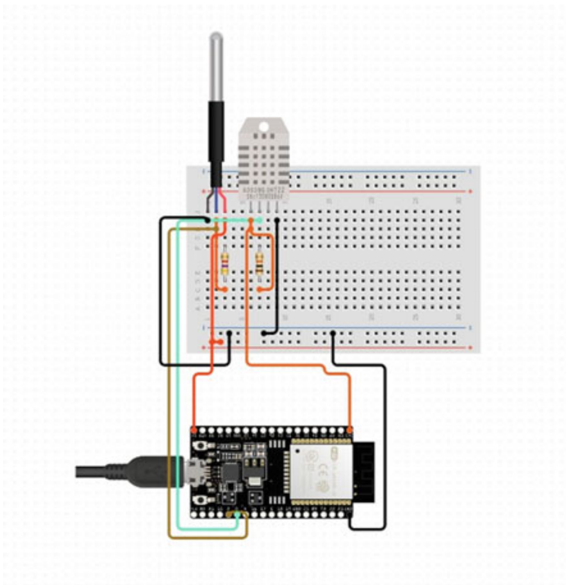


Fig. 8 Interfacing dia. of DHT22 and DS18B20 sensor

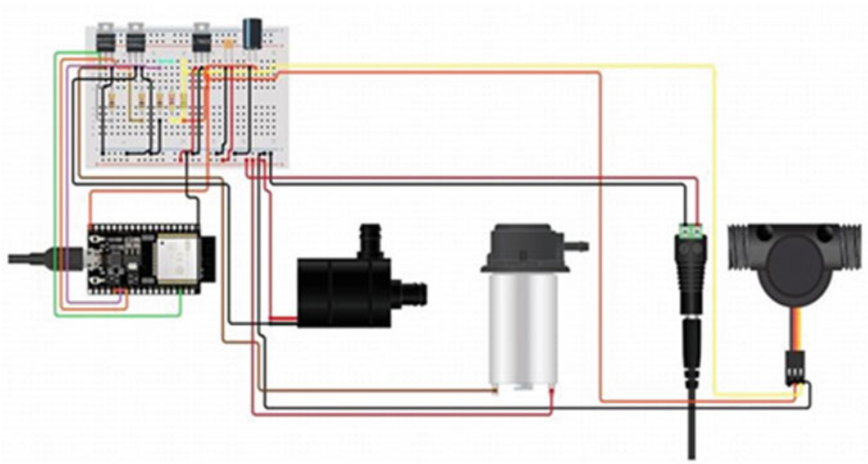


Fig. 9 Interfacing diagram of flow sensor and water pump

The maximum flow rate for water pump in this setup is 650 L/min and pump head are ~ 3 ft.A 1/2 inch. Hall sensor monitors the water flow of the pump. Dosing pumps/peristaltic pumps are used to add nutrients to the reservoir tank, to balance pH of the water. Kamoer NKP-DC-S04 12 V 10 ml/min pump is used in this setup.

10 Implementation of the System

Section 4

This section gives information regarding the actual implementation. The Nutrient Film Method (NFT) cultivation method is used for experimentation (Fig. 10) where plant roots are permanently immersed in a nutrient solution. The solution is kept in a thin film that allows the plant to extract the nutrients through the roots that are in constant contact with the solution. The air pump saturates the solution with oxygen, preventing the plant from suffocating.

The system is developed for spinach. The setup for NFT structure is shown in Fig. 10, the NFT structure consists of 5 PVC pipes each of 45 inches in length and 3 inches in diameter placed 6–7 inches parallel to each other. Holes of diameter 2 inches for net cups have been drilled at a distance of 5 inches from each other. Each pipe consist of 5 net cups and 5 such pipes have been placed so total 25 net cups will be planted with spinach. The net cups are filled with coco peat as a substrate. The spinach growing in this system can be ready to harvest within 30–45 days (Table 2).

The system contains four sensors: pH sensor, TDS sensor, a temperature sensor probe, and a temperature and moisture sensor. TDS sensors are used to estimate the salinity or nutrient content of the water, and pH probes must be calibrated using a calibration solution to regulate the pH of the water before introducing plant saplings

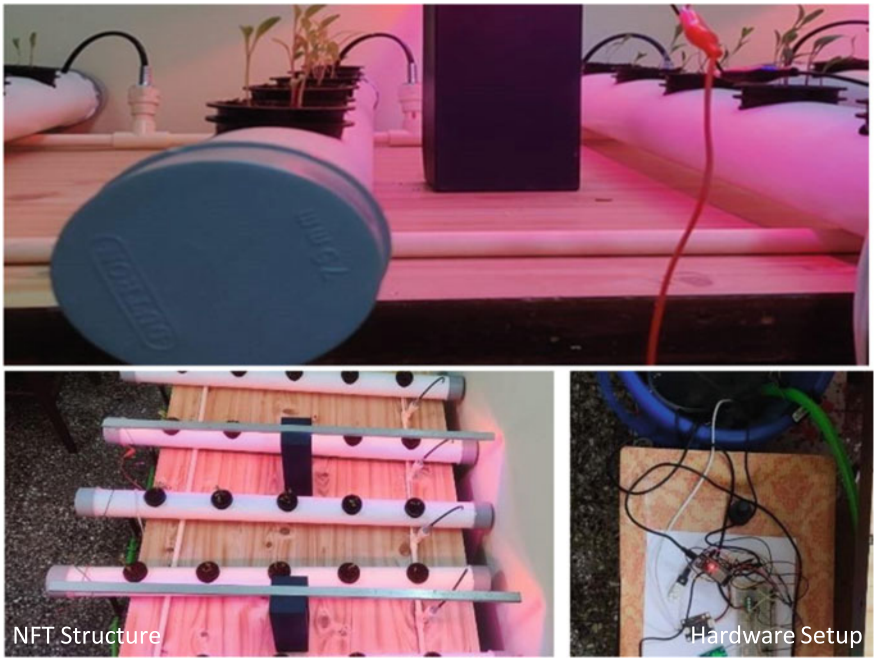


Fig. 10 Nutrient film technique setup

Table 2 Nutrient requirements for spinach plant

Nutrient	Essential for	Range
Nitrogen (N)	Leaf growth, chlorophyll production, and overall growth	120–140 ppm
Phosphorus (P)	Root development and energy transfer within the plant	40–60 ppm
Potassium (K)	Water regulation, disease resistance, and overall growth	180–220 ppm
Calcium (Ca)	Proper cell wall development and root growth	100–200 ppm
Magnesium (Mg)	Chlorophyll production and overall plant health	40–80 ppm
Sulfur (S)	Amino acid production, protein synthesis, and overall growth	20–30 ppm
Iron (Fe)	Chlorophyll synthesis and overall plant growth	2–5 ppm
Manganese (Mn)	Photosynthesis and enzyme activation	1–2 ppm
Boron (B)	Cell division, pollen germination, and over all—plant growth	0.1–0.2 ppm
Zinc (Zn)	Enzyme activation and protein synthesis	0.5–1 ppm
Copper (Cu)	Chlorophyll synthesis and enzyme activation	0.1–0.2 ppm
Molybdenum (Mo)	Nitrogen fixation and enzyme activation	0.01–0.05 ppm
Chlorine (Cl)	Osmotic regulation and photosynthesis	50–100 ppm
pH level	Affects nutrient availability and uptake	5.5–6.5

in the system. Because pH is temperature-dependent, it is important to measure the pH of water at a constant temperature. The system uses an TDS sensor to measure the nutrient level of the solvent. RO water is used, and the pH is adjusted to 6–6.5 as per requirement of spinach, after adding nutrients. The water heater maintains a temperature around 25 °C, which is required for healthy growth of root system (Fig. 11).

The flow chart in Fig. 12 gives us the idea of IoT-enabled hydroponics system. The MCU used here is ESP32 once it is connected to network it initiates the sensors to acquire the sensor data. Based on this sensor data and cloud database the ESP32 controls various functions in the system. The level of water in the reservoir is set with the database if the water is lower than threshold it will alert the system through level sensor and water will be added to the reservoir using the pumps this action will continuously sense till the water reaches its threshold after that the sensor will check after 30 min of delay.

The water temperature in the reservoir is monitored by DS18B20 probe with threshold of 25 °C if water temperature drops below threshold MCU will switch on the water heater. The flow sensor will continuously monitor the water flow and if the water is not circulating the MCU will alert the owner and switch on the back up pump.

The TDS sensor will monitor the total dissolved solvents and as the plant will grow according to its growth stages it will increase the TDS value by adding nutrients using Peristaltic pump/Dosing pump at a precision amount of 10 ml/min. Also, as

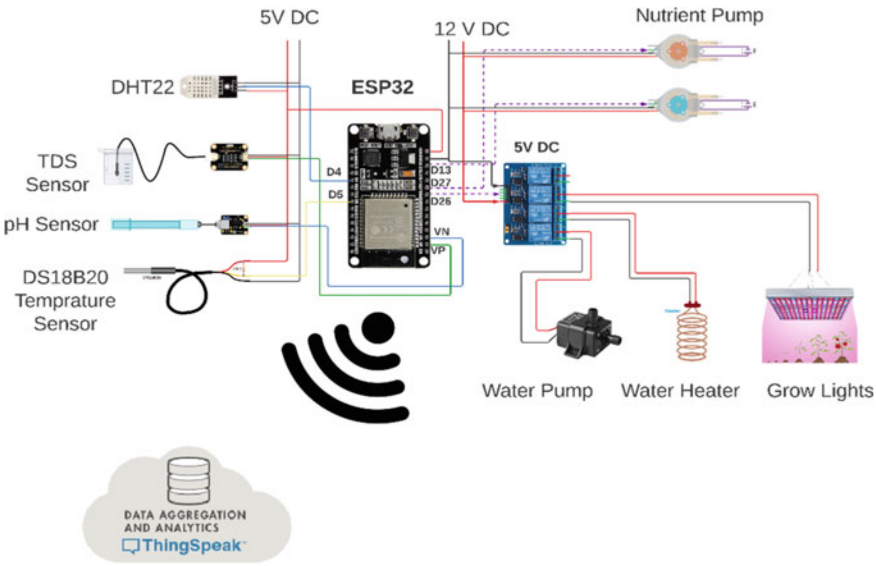


Fig. 11 Interfacing diagram of Wi-Fi enabled ESP32 with sensors and actuators

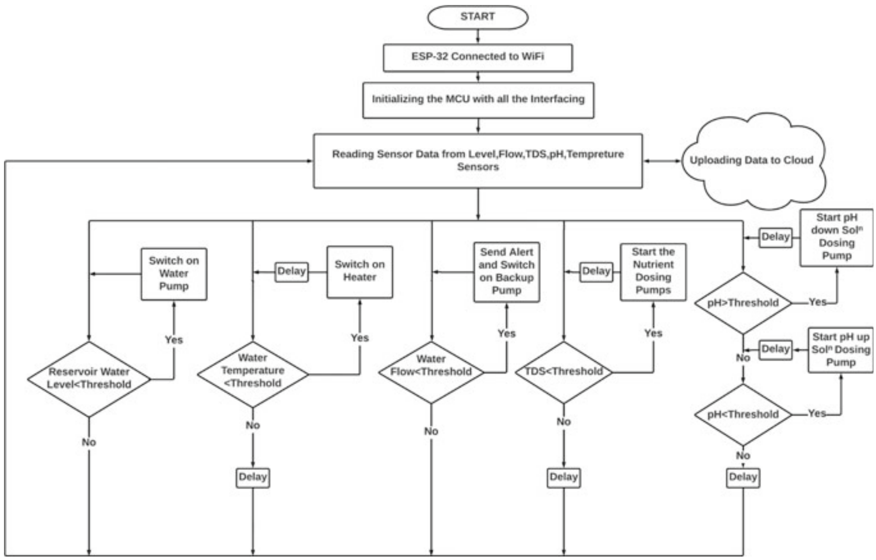


Fig. 12 IoT-enabled hydroponic system flow chart

the plants will absorb the nutrients it will drop the TDS level in the water during its growth cycle the TDS sensor will monitor the TDS at the interval of 24 h.

The threshold for pH is set at 5.5–6.5 for spinach plant and if the pH value is greater than threshold, Peristaltic pump/Dosing pump will add 1–2 ml of pH down solution at delays of 30 min to circulate the water and test the pH before adding next dose. Similarly, the pH will be balanced if it is less than threshold. Simultaneously every 1 h the data will be uploaded to the cloud. The sensor data is uploaded to ThingSpeak platform for analysis and visualization.

11 Observations

Figure 13 shows the time series data received in the cloud for all the physical parameters. There is slight variation in temperature and humidity as per the environmental conditions, and time of the day. The value of pH and TDS is kept well within the required range.

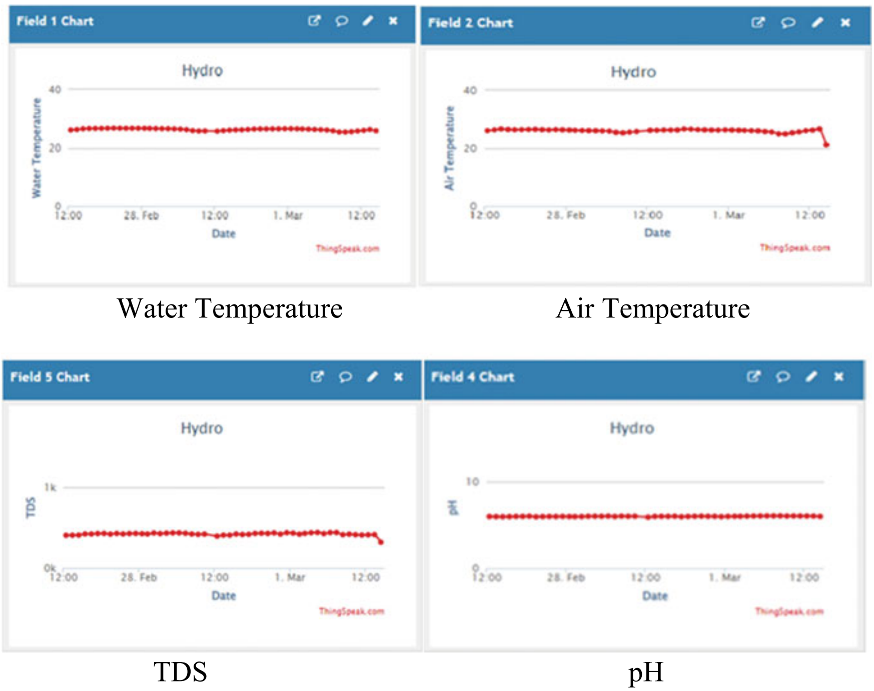


Fig. 13 Time series data received in the cloud for the setup

12 Conclusion and Future Direction

Section 5

Hydroponics has a lot of potential to revolutionize agriculture and food production, especially in areas where traditional farming methods may not be feasible due to factors like lack of arable land, water scarcity, or extreme weather conditions. Hydroponic systems can be designed to optimize plant growth, reduce water consumption, and eliminate the need for harmful pesticides or fertilizers. Additionally, hydroponic farming can be done indoors or in urban areas, allowing for year-round production and a reliable supply of fresh, healthy produce. With increasing demand for unstainable and locally grown food, it's likely that hydroponic farming will continue to gain popularity and be adopted in more regions. This project is an attempt to design a low-cost solution to improve the yield of spinach. The outcome is encouraging, the quality and growth of the produce is better as compared to the crop in soil. The time required to get the harvest is also less as compared to the soil. However, as with any new technology or method, there may be challenges to overcome in terms of cost, scalability, and accessibility. It will be important for researchers and farmers to continue developing and refining hydroponic systems to make them more efficient, affordable, and accessible to people of all socioeconomic backgrounds.

References

1. United Nations Department of Economic and Social Affairs, Population Division (2022) World population prospects 2022: summary of results. UN DESA/POP/2022/TR/NO. 3
2. Calicioglu O, Flammini A, Bracco S, Bellù L, Sims R (2019) The future challenges of food and agriculture: an integrated analysis of trends and solutions. Sustainability 11(1):222
3. Folnovic T (2021) Loss of arable land threaten world food supplies. Agrivi, London, U.K., Technical report. https://blog.agrivi.com. Accessed 1 May 2021
4. Almadani B, Mostafa SM (2021) IIoT based multimodal communication model for agriculture and agro-industries. IEEE Access 9:10070–10088. https://doi.org/10.1109/access.2021.3050391
5. Hamdalla J, Haque F, Reaz MBI, Al Shafei A, Al-Emadi N (2020) Design, construction and testing of IoT based automated indoor vertical hydroponics farming test-bed in Qatar. Sensors 20(19):5637
6. Bardi S, Palazzi CE (2022) Smart hydroponic greenhouse: internet of things and soilless farming. In: Conference on information technology for social good (GoodIT'22), September 7–9, 2022, Limassol, Cyprus. ACM, New York, p 6
7. Rahimi M, Saad M, Hanif M, Hussain A, Hamdan M (2022) Secure cloud connected indoor hydroponic system via multi-factor authentication. Int J Adv Comput Sci Appl 13:925. https://doi.org/10.14569/IJACSA.2022.0130925
8. Ludwig F, Fernandes DM (2013) Electrical conductivity and pH of the substrate solution in gerbera cultivars under fertigation. Hortic Bras 31(3):356–360
9. Cambra C, Sendra S, Lloret J, Lacuesta R (2018) Smart system for bicarbonate control in irrigation for hydroponic precision farming. Sensors 18:1333. https://doi.org/10.3390/s18051333
10. Palande V, Zaheer A, George K (2018) Fully automated hydroponic system for indoor plant growth. Proced Comput Sci 129:482–488

11. Jandl A, Frangoudis PA, Dustdar S (2022) Edge-based autonomous management of vertical farms. IEEE Internet Comput 26(1):68–75. https://doi.org/10.1109/mic.2021.3129271
12. Peuchpanngarm C, Srinitiworawong P, Samerjai W, Sunetnanta T (2016) DIY sensorbased automatic control mobile application for hydroponics. In: Proceedings of the 2016 5th ICT international student project conference ICT-ISPC 2016, pp 57–60
13. Farooq MS, Riaz S, Helou MA, Khan FS, Abid A, Alvi A (2022) Internet of things in greenhouse agriculture: a survey on enabling technologies, applications, and protocols. IEEE Access 10:53374–53397
14. Nugroho ED, Putrada AG, Rakhmatsyah A (2021) Predictive control on lettuce nft-based hydroponic IoT using deep neural network. In: Proceedings of the 2021 international symposium on electronics and smart devices (ISESD). IEEE, pp 1–6
15. Chowdhury ME, Khandakar A, Ahmed S, Al-Khuzaei F, Hamdalla J, Haque F et al (2020) Design, construction and testing of IoT based automated indoor vertical hydroponics farming test-bed in Qatar. Sensors 20(19):5637
16. Jan S, Rashid Z, Ahngar T, Iqbal S, Naikoo A, Majeed S, Bhat T, Gul R, Nazir I (2020) Hydroponics: a review. Int J Curr Microbiol Appl Sci 9:1779–1787. https://doi.org/10.20546/ijcmas.2020.908.206
17. Judith S (2019) Ph in hydroponics: how to maintain the ph levels of hydroponic systems. Jenco Water Quality Blog. https://blog.jencoi.com/ph-in-hydroponics-how-to-maintain-the-ph-levels-of-hydroponic-systems. Accessed 6 March 2023

AMAX-AR: A Way to Maximize Augmented Reality

Sakshi Modi, Siddhi Bafna, Sneha Belapurkar, Trapti Mishra, and And Prashant Panse

Abstract Augmented reality intention will rise exponentially because it amplifies its authority to have valued and remold challenges. AR will advance as the cover interface between engines and humans, bridging the digital and right humanities. While difficulties in planting dwell, presenting institutions, analogous as Amazon, Google, Facebook, Mayo Clinic, the U.S. Navy, and General Electric formerly dealt with AR and observed a considerable strike on caliber. AR operations are formerly guided and posted in shops and across the valuation train, and their composition and reach will exclusively rise. AR operations give a kind of X-ray delusion, discovering interior features that would exist subtle to experience. AR gets aimed at a wide consumer cult that has generally lived aimed for smartphones, holding the upper hand over their plainness and ubiquity. For farther smart guests, pots use tablets, which give huge defenses, better plates, and lower processing authority. In 2013, Bujak et al. (Comput Educ 68:536–544, 2013) refocused on 'augmented reality (AR) is just starting to scrape the look in informative operations.' This paper deals with our operation AMAX-AR which will help in the business and education sphere by giving a three-dimensional view of the arrangements or plates which can be more understood in three-dimensional rather than two-dimensional. Also tapping the generated model will give the accurate labels of the model. This operation will be accessible on both Android-grounded biases as well as Windows-grounded biases that have a camera or web camera access.

S. Modi (✉) · S. Bafna · S. Belapurkar · T. Mishra · A. P. Panse
Medi-Caps University, Indore, India
e-mail: EN19IT201072@medicaps.ac.in

S. Bafna
e-mail: EN19IT201085@medicaps.ac.in

S. Belapurkar
e-mail: EN19IT201087@medicaps.ac.in

T. Mishra
e-mail: trapti.mishra@medicaps.ac.in

A. P. Panse
e-mail: prashant.panse@medicaps.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_19

Keywords Augmented reality (AR) · Unity · Vuforia

1 Introduction

Augmented reality (AR) is an evolved explanation of the material humankind that is accessible through the exercise of digital illustrations, tones, or distinct technological instruments. It's a rising trend among organizations affected by the operation of laptops and business agendas. Augmented reality maintains to broaden and introduce redundant patients amongst a huge variety of operations. Since its inception, entrepreneurs and technology pots have had to war with the notion that it is little further than an advertising and marketing tool. But there is evidence that guests are starting to decide on the palpable advantages of this functionality and count on it as a part of their buying system.

Augmented reality can be a great help in the education sector for scholars to grasp the diagram and maintain their interest in the content. Amongst the epidemic situation, the education sector is one of the most affected areas. So our design AMAX-AR aims to help scholars in understanding the diagram without actually seeing them in the real world. This design also aims to act as a ground between a person and the real-world academy using technological aids. Using AMAX-AR, the user can see the two-dimensional image into a three-dimensional image.

1.1 Definition

'Augmented reality is a technology that superimposes computer-generated two-dimensional illustrations into three-dimensional real-world things exercising digital optical aids, detectors, etc.'

1.2 Brief History

Over the half century, augmented reality has changed the route we've devoured and adhered to content in the real world.

In 1968, Ivan Sutherland, a professor and scientist, built the initial block-mounted display named 'The Sword of Damocles.' The stoner endured digitally-generated illustrations that amended their delicate sapience of the physical humankind. Further in 1970's Myron Kruger, a computer trial, erected a laboratory at the Connecticut University called 'Video place' that lived totally adoring AR. Within its ground, camera technology and bunch existed to release on screen outlines which gathered fiends for a unique experience. Still, the term 'augmented reality' was not cashed until 1990 by Boeing trial Thomas Caudel.

Latterly in 2000, Hirokazu Kato formed a open-sourced software archive as the ARToolkit, which helped manufacture AR software programs. The archive uses video tracing to overlay virtual visuals on the cover of the true world. In 2009, Esquire Magazine referred to AR by impressive means for the foremost moment in a trial to reduce the runners to animate. Volkswagen debuted the MARTA operation (Mobile Augmented Reality Technical Assistance) which firstly offered technician's tread-by-measure form instructions within the turn in 2013. The coming time, Google discovered its Google Glass bias, a brace of AR spectacles that fiends could transport for immersive times. Fiends wearied the Augmented reality technology and conducted with the web employing born speech processing masteries. With this device, they could pierce colorful plays developed by Google similar as Charts, correspondence and numerous further. Advancing Google Glass in 2016, Microsoft started transferring its account of wearable AR technology called the HoloLens which comes with a massive cost label. It's surely not a day-to-day type of outfit. The headset runs on Windows 10 and is principally a wearable computer, like it similarly allows fiends to overlook their terrain.

The ending of 2010's decade attracted deep attention from researchers toward AR technology. Run by computer vision and non-natural intellect technology, AR technology has exhibited a strong incentive of progress. Both the tracing enrollment delicacy, show stuff interpretation and the colors of natural-computer relation possess lived honorably bettered [2, 3].

Augmented reality is an optimal solution for problems analyzing complex structures and objects. We are developing an android application that can help understand two-dimensional pictures in three-dimensional form. So, our project AMAX-AR is proposed by using a 3D augmented reality model that can be accessed by the user when he or she scans the image using the camera accessed by the application. This project will provide an inside view of an object from several angles for better understanding.

The project is created using Unity and Vuforia embedded in it already. When the generated model is tapped the virtual button will activate which will label the diagrams for a better understanding of the user. The application can be run on any android device as well as a windows-based device. The application can be run on both platforms efficiently. The only basic requirement is a camera or a webcam (in the case of a computer or laptop). Our application will overcome the major drawbacks of the existing systems and also be cost effective.

2 Comparison with the Existing Application

According to the surveys, this pandemic has been a major disadvantage for the education sector. The teachers and students are trying their level best to keep learning through the online platforms but some things can only be understood when seen in front of the person physically. For example, a student who is studying computer peripherals and devices cannot study a motherboard just looking into the picture, it

Table 1 Comparison between different existing applications and AMAX-AP

Criteria	Vectary [4]	Selva3D [5]	3D-tool [6]	Amax-AR
Image scanning	No	No	No	Yes
Android support	No	No	No	Yes
Windows users	Yes	No	Yes	Yes
Augmented reality	Yes	No	No	Yes
Image quality	High	Average	Average	High
Open source	Yes	Yes	No	Yes
Labeling	Yes	No	No	No
Internet access	Yes	Yes	Yes	Yes

is impossible to grasp it without seeing a motherboard structure physically. Similar problems occur when you are a biology student (Table 1).

To overcome these problems there are several already developed software and applications but those applications also have following major concerns:

- They are very costly.
- No image scanning is allowed.
- Requires high system specifications.
- Low image quality.
- No labels or markings of the models.

Need for proper System: As mentioned above, there are various reasons because of which the existing applications are completely out of the reach for normal population. Whether it be because of the cost of the project, or the scanning image, high system specifications, low image quality or labels or markings are not present [7].

3 Methodology

To begin with our technology we would have to go through some of the specifications of Hardware like a camera, Minimum 2 GB RAM, Processor: 1.9 gigahertz (GHz) $\times$ 86- or $\times$ 64-bit dual-core processor and Software like Android v8 (Oreo) or later Minimum Windows 7 (Fig. 1).

Then with creation of our application, the modules we went through are Unity which is a cross-platform gaming engine constructed by Unity Technologies. It gives users and developers capability to produce two-dimensional and three-dimensional illustrations, and the machine offers a direct scripting API in C#. Its capabilities make it an excellent choice for erecting apps that run on different devices and OSs, including iOS, Android, Windows, macOS, and Linux. Also has a robust set of features for creating 3D graphics and animations.

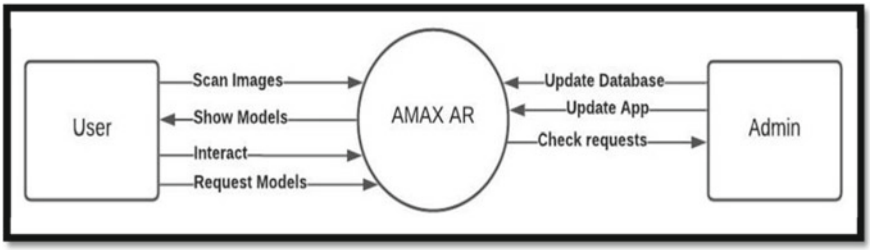


Fig. 1 Data flow of the application

Unity's scripting capabilities allow developers to write code in C# or JavaScript, making it accessible to a wide range of developers with different programming backgrounds. Unity's Asset Store provides a vast library of ready-made means, such as 3D models, animations, and plugins that can be used to speed up the app development process.

Overall, Unity is a strong and flexible tool for building apps across various platforms, and its popularity among game developers means that there is a large community of users and resources available for developers looking to get started with Unity app development [8].

Moving next with the **Vuforia**, an AR platform that allows us to produce AR experiences for mobile devices, such as smartphones and tablets. It uses computer visuals to recognize and trace target images, similar to product packaging, ensigns, or other physical-world things, overlaying digital content above them in real-time.

People can use Vuforia to create AR apps for various industries, containing gaming, education, retail, and marketing. Create immersive AR experiences, such as image recognition, text recognition, and 3D object recognition.

To use this technology, we integrated the Vuforia software development kit (SDK) into our app development environment which provided libraries and tools that allowed us to add AR functionality to apps.

Overall, Vuforia technology can be a powerful tool for developers looking to create engaging AR functionality for mobile users [9] (Fig. 2).

4 Applications

This project has an open-ended scope in every field. Some scopes of this application are enlisted below:

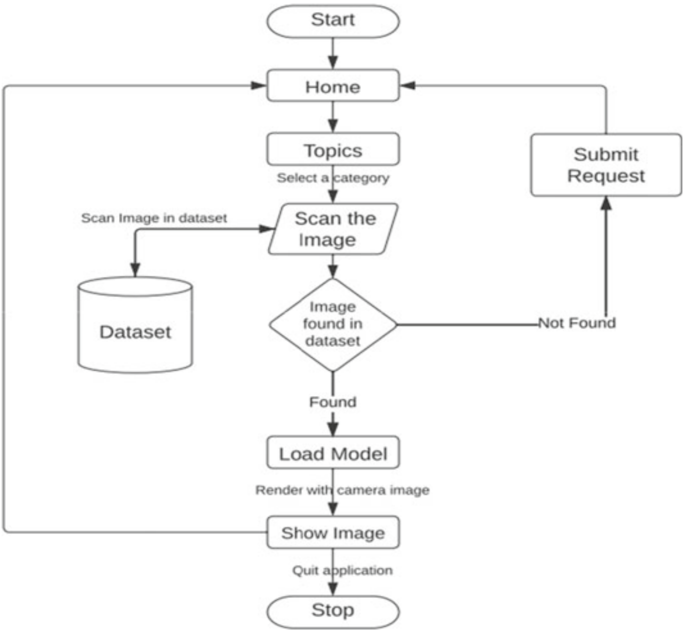


Fig. 2 Flowchart of the application

4.1 Education

With AR, classroom training can be extraordinary and further interactive, as AR can permit instructors to display virtual illustrations of generalities and add gaming fundamentals to give text material support. This will enable scholars to learn briskly and study information [10, 11].

4.2 Research

Since its preface, augmented reality (AR) has been displayed to have good eventuality in forming the literacy procedure more functional, fruitful, and expressive. This is the reason why advanced technology enables users to interact with virtual and real-time operations and brings the natural experiences to the user. In addition, the coupling of AR with education has lately attracted exploration attention because of its capability to allow scholars to be immersed in realistic gests. Thus, this conception paper reviews the exploration that has been conducted on AR.

4.3 Architecture

Augmented reality in structure and architecture systems involves setting up a 3D miniature of an offered design onto an existing room using portable bias and 3D models. AR has been used in AutoCAD (magic meeting is MR2, a mixed reality meeting room [12]). At the present time, a veritably many Architects and contrivers use AR, and in the utmost of the time AR is only used to enhance donations and marketing. This is due to a lack of integration of AR technologies in the design workflow. Although this problem has been solved since the last decade, a lack of norms prevents real integration of AR in CAD software [13].

4.4 Healthcare

By using an AR enabled block accumulated device, surgeons and doctors can presently check patient vitals during an operation without the need to see out and collect data from numerous displays. In doing so, surgeons and doctors are less probable to make errors due to misunderstanding, misreading, and/or misinterpreting the data. The eventuality of AR in medical operations was provisioned by Steinhaus (an Austrian mathematician) in 1938 Steinhaus [14] suggested a system to display a pellet inside a body by a veritably clumsy overlay process. On the other hand, the first real operation of AR in healthcare was in 1986, when a system to integrate data from computer tomography into an operating microscope was proposed. Compendiums interested in examining in depth operations of AR in healthcare can relate [15].

4.5 Military

AR applications for military purposes share a lot of issues with AR-based games and AR-based training systems. The term often used is BARS: Battlefield augmented reality system [16].

5 Conclusion

As we come gradually reliant on our portable phone bias, the embracement of augmented reality technology will commence. AR software advancements will live the road ahead as the irresistible maturity of consumers possesses a smartphone and previously held it high and low with them, solidifying it as a popular medium to learn AR to every consumer.

The diversity of augmented reality is consumed by day-to-day users—they precisely suit it. AR quietly seems to live tripped as overly ‘high tech’ for your mean Joe. The biggest step in AR will have to be how it’s redeemed to modify the perception.

Likewise, augmented reality accoutrements and software are fairly behind the times. That thing said, there is a major possibility for the progress of a workable augmented reality headset. Wearable tech is sluggishly getting the current and, as this shift continues, people might exist further candid about AR attack.

The acceptance of AR technology has to begin with the software apps to catch the users where they are now.

6 Result

So, our project is a gateway to step into the world of augmented reality where you can convert your two-dimensional image into an augmented reality model that is also free of cost, which no other existing project allows. Also, with the help of C# various other features are integrated in our application. AR has the ability to revise the manner in which we interact with the public around us, by providing new and innovative ways to access and interact with digital content in real-time. The potential benefits of AR are significant, and as the technology continues to advance and become more accessible, we can expect to see even more exciting applications and use cases in the future (Figs. 3 and 4).

Fig. 3 Input area of the application (input of the image is captured here)

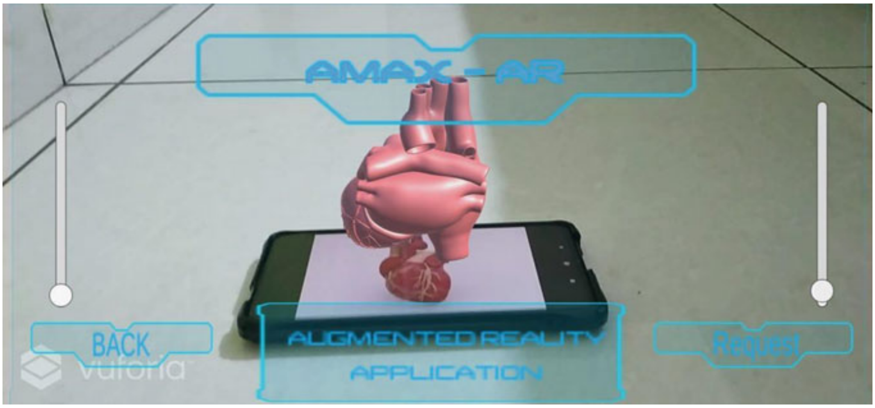


Fig. 4 Output area of the application (scanned output image of the heart)

7 Future Scope

- Dynamic dataset can be included.
- Better GUI can be created using C#.
- Animations can be added as per the requirement.

References

1. Bujak KR, Radu I, Catrambone R, MacIntyre B, Zheng R, Golubski G (2013) A psychological perspective on augmented reality in the mathematics classroom. Comput Educ 68:536–544
2. Chen Y, Wang Q, Chen H, Song X, Tang H, Tian M (2019) An overview of augmented reality technology. J Phys Conf Ser 1237(2):022082
3. A brief history of augmented reality (+Future Trends & Impact) (g2.com)
4. Vectary. https://www.vectary.com/3d-design-tool/
5. Selva. https://www.selva3d.com/
6. 3D tools. https://www.3d-tool.com/
7. Literature survey. https://www.rankred.com/convert-2d-images-to-3d/
8. Unity. https://unity.com/our-company
9. Vuforia. https://developer.vuforia.com/
10. Yuen SC-Y, Yaoyuneyong G, Johnson F (2011) Augmented reality: an overview and five directions for AR in education. J Educ Technol Dev Exchange 4(1):119–140
11. Wu H-K, Lee SW-Y, Chang H-Y, Liang J-C (2013) Current status, opportunities and challenges of augmented reality in education. Comput Educ 62:41–49
12. Kiyokawa K, Niimi M, Ebina T, Ohno H (2001) MR2 (MR square): a mixed-reality meeting room. In: Proceedings of the IEEE and ACM international symposium on augmented reality, New York, Oct. 29–30. IEEE, pp 169–170
13. Broll W, Lindt I, Ohlenburg J, Wittkämper M, Yuan C, Novotny T, Fatahgen AS, Mottram C, Strothmann A (2004) ARTHUR: a collaborative augmented environment for architectural design and urban planning. J Virt Real Broadcast 1(1):1–9

14. Steinhaus H (1938) Sur la localisation au moyen des rayons x. Comptes Rendus de L'Académie des Sciences 206:1473–1475
15. Sielhorst T, Feuerstein M, Navab N (2008) Advanced medical displays: a literature review of augmented reality. J Display Technol 4(4):451–467
16. Livingston MA, Rosenblum LJ, Julier SJ, Brown D, Baillot Y, Swan JE, Gabbard JL, Hix D (2002) An augmented reality system for military operations in urban terrain. In: Proceedings of the interservice/industry training, simulation, and education conference, Orlando, FL, Dec. 2–5, pp 868–875

Slack Time Analysis for APB Timer Using Genus Synthesis Tool

R. Madhura, Krutthika Hirebasur Krishnappa, R. Manasa, and K. P. Yashaswini

Abstract Synthesis plays an important role in the complex world of chip design, as constant push to improve chip—to get more performance, lower power, and improved area. In an effort to emphasize more exact results while using the genus tool, the Advanced Peripheral Bus (APB) timer—a design that is regarded to be a standard framework—is implemented in this research. Logical synthesis is then performed using Synopsys Design Constraints (SDC), fast lib, and slow lib. Fast lib was used to enhance setup check slack time and minimize hold check slack time. It is capable to resolve timing issues with regards to setup and hold time and comply with deadlines utilizing late external delay assertions for setup and early external delay assertions for hold time.

Keywords APB timer · Synopsys Design Constraints · Genus · Slack time and static timing analysis

R. Madhura (✉) · R. Manasa · K. P. Yashaswini
Visvesvaraya Technological University Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India
e-mail: madhura-ece@dayanandasagar.edu

R. Manasa
e-mail: manasar.87@gmail.com

K. P. Yashaswini
e-mail: yashaswinikp144@gmail.com

K. H. Krishnappa
Department of Computer Science, Southern University and A&M College, Baton Rouge, LA, USA
e-mail: krutthika.hirebas@sus.edu

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_20

1 Introduction

1.1 A Subsection Sample

The genus synthesis solution has turnaround times for synthesis that are up to 5X faster and is linearly extendable beyond 10 M instances. Additionally, the number of iterations between unit and chip-level synthesis is reduced by a factor of two or more due to a unique physically aware context-generation capability. This potent combination can boost RTL design productivity by up to 10X. Also, a new global, analytical, architecture-level optimization engine may reduce the size of the data route by up to 20% [1] without compromising speed. The specific goal in this research work is to examine every possible path for timing violations with the intent to validate the timing performance of a design and also to minimize the area, power and time without any negative slack. Logical synthesis is a phase in the ASIC flow.

The three processes of translations, mapping, and optimization constitute the synthesis procedure. These three phases are used in synthesis to convert Verilog or VHDL RTL code into a gate level netlist. With the use of synthesis tools, Design Compiler by Synopsys and genus synthesis solution by Cadence are two easily available synthesis tools. The main objective of this study is to satisfy the design's timing requirements in accordance with the specified parameters and to increase design productivity through slack time analysis. The three essential factors such as power, performance, and area should be considered at each stage of the ASIC flow. By evaluating every potential path for timing violations, the static timing analysis (STA) approach evaluates the design's time efficiency. After being partitioned into time routes and the signal propagation delay on each route being calculated, timing restrictions inside the design and at the input/output interface are evaluated for violations in STA [2]. The methodology followed in this research work are

Step 1: RTL design for APB timer.

Step 2: Verify the functional correctness of the RTL design.

Step 3: Write Synopsys Design Constraints (SDC) for RTL to generate SDC file.

Step 4: Perform synthesis using genus tool by writing TCL scripts using SDC file as input and to generate gate level netlist as output.

Step 5: Perform setup and hold time check to analyze and clear violations to meet timing requirements.

Step 6: Generate expected gate level netlist.

The rest of the paper is divided into the following sections: Short summaries of the related work and their findings are presented in Sect. 2. The proposed work is in Sect. 3. The implementation and results are covered in Sect. 4, and the conclusion is presented in Sect. 5.

2 Related Work

The main motivation for this research work is to resolve the issues in synthesis phase in the chip design since it is essential for DFT and physical design, which are crucial for determining the transmission rate. Another significant element is the timing analysis that is necessary to guarantee that any clocks produced by the logic circuit are accurate, possess an established connection to other crucial signals, and have specified duration and duty cycles.

The authors of paper [3] employed a number of strategies, and the results include various functionality in the areas of VLSI and SoC technology. The results have been considerably improved by changes to the synthesis flow approach, which is essential when performing tasks like design for formal verification, static analysis, and validation.

According to the researcher's analysis [4], the usage of EDA tools at every stage of the system design produces extremely accurate and top-notch results. Moreover, it reduces the design's reliance on time for operations, partitioning, and parallelism. The duration and cost involved in generating a model were decreased with growing use of EDA technologies. This research study successfully completes the majority of ASIC design processes, from RTL model construction through FEV.

The authors of paper [5] presented reg2reg setup and hold analysis for a sample circuit as well as circuits for locating and fixing. Such circuits have sequential circuits linked to it so they can be inspected. In contrast to setup violations, which cause the chip to operate at a lower frequency, hold violations inside this microchip result in the chip never functioning anymore. These breaches lead the circuit to become metastable.

The authors of the research [6] addressed a straightforward GUI for setup and hold time analysis of circuit design. As a result, the GUI has a minimal number of components and is ready to display data immediately after receiving user input. The recommended method from linear model analysis has a lower computational barrier and a lower probability of being unreliable.

The authors in paper [7] has optimized DFT architecture through scan-based design is proposed, where an efficient low power test pattern generator is used to verify any real-time Design Under Test (DUT) module using Functional and Structural testing which is used as a simulation and verification in our project. The limitation was Scan Compression Techniques were not adopted in this paper.

According to paper [8], the authors were able to detect transition delay faults using simulation model Atspeed correlations implemented for real-time configuration under assessment. They were successful in attaining 71.85% error performance for attached at type defects and 55% fault coverage for transformation model defects though the test procedures using 28 nm technology as well as the tessent tool.

The serial and parallel defect modeling approaches have been employed in the study [9]. 14 nm technology was used for synthesis in the Synopsys tool.

The authors of the study [10] have explored employing the FSM approach to prevent congestion in the network on chip. They have demonstrated that the method is effective for transferring the packet when there is a faulty connection.

Physically aware synthesis, in general, is a method that incorporates physical information into the synthesis process, such as physical libraries, macro placement, blockages, and area information. With a number of experimental configurations, this study achieves the ideal flow for the synthesis run. Static timing analysis (STA), Effects in Place and Route (PNR), and other phenomena are also noticed and carefully examined [11].

3 Proposed Work

The proposed system architecture in Fig. 1 depicts the synthesis flow for the chosen RTL design APB timer.

The proposed system architecture in Fig. 1 depicts the synthesis flow for the chosen RTL design APB timer.

3.1 APB Timer

Forward the APB timer in Fig. 2 is a 32-bit down-counter with the following features:

- When the counter reaches 0, an interrupt request signal called TIMERINT is created. Until it is cleared by writing to the INTCLEAR register, the interrupt request is retained.
- The external input signal, EXTIN, has a zero to one transition that can be used as a timer enable.

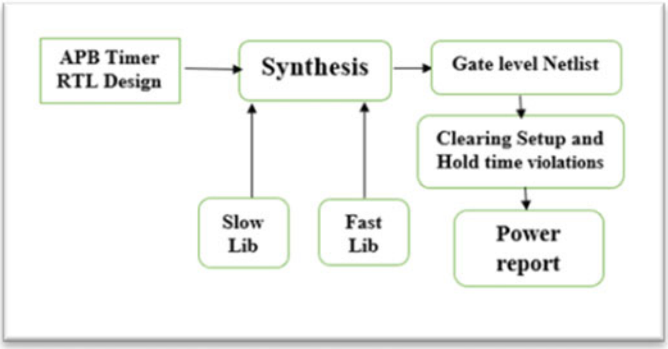


Fig. 1 Proposed system architecture

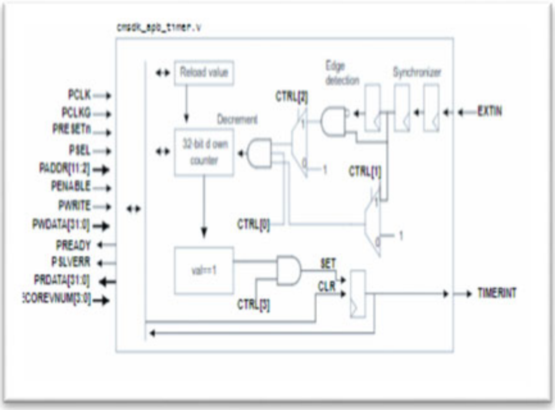


Fig. 2 APB timer design

- The interrupt status is set to 1 if the APB timer count approaches 0 and the software simultaneously clears a prior interrupt state the PLS approaches.

3.2 Synthesis

Through logical synthesis is performed on RTL design with fast and slow libraries with Synopsys Design Constraints (SDC) to generate gate level netlist and power reports by clearing setup and hold time violations which is shown in Fig. 3.

Terminologies in STA are listed below:

Setup time: The time before the clock's triggering-edge where the input has to be stable is known as setup time.

Hold time: The input data has to be stable during the hold time, which is the period of time that comes after the clock's triggering point.

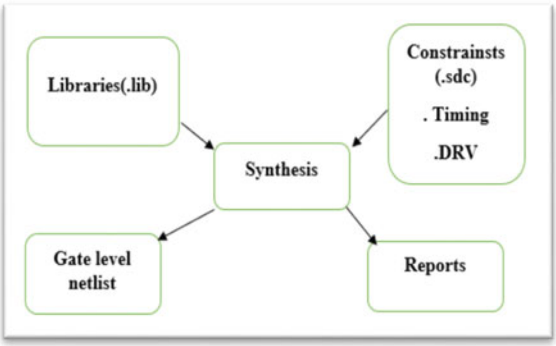


Fig. 3 Synthesis inputs and outputs

Slack time: It is the entire amount of time that may be used to defer a job without deferring the entire project. Project scheduling is more flexible when there is more down time.

3.3 Static Timing Violations

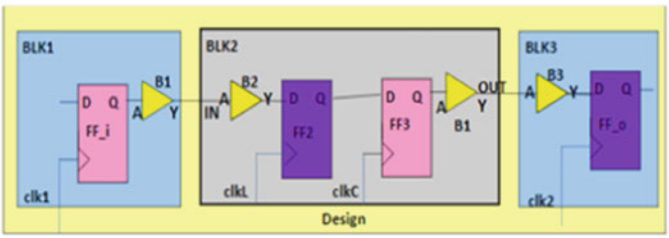
Static timing analysis (STA) is a technique to evaluate the timing performance of a layout by investigating all conceivable paths for timing failures. In STA, timing limits inside the design and at the input/output interface are evaluated for infractions coupled with the calculation of the signal propagation delay along each path. STA divides the design first into timing pathways. It has three different types of timing analysis.

- In commencement of a timing path, the data is launched during clock edge or in which data accessibility at a particular time must be achieved. Each start-point necessitates the existence of an input or a register clock pin.
- In combinational logic, elements cannot have internal states or memories. It can have AND, OR, XOR, and inverter elements.
- Endpoint is the point in a timing path where data must be available at a certain time or where a clock edge captures the data. Every endpoint must be either an output port or a pin for registering data.

The timing paths for start and end point is shown in Fig. 4.

Different methods are used to compute setup slack and hold slack. Analysis of setup slack time which is depicted in Fig. 5.

The arrival time, required time, and setup slack are calculated using the formulas.



Timing Path	Begin Point	End Point
IN to REG	IN	D FF2
REG to REG	CK FF2	D FF3
REG to OUT	CK FF3	OUT
IN to OUT	IN	OUT

Fig. 4 Timing paths

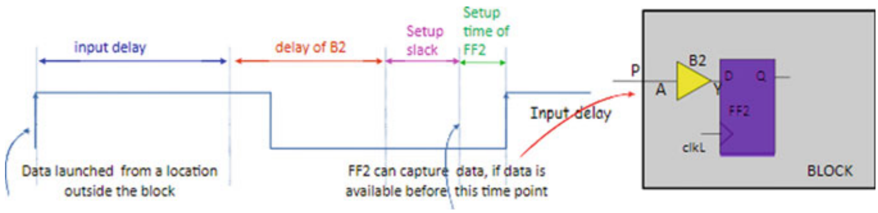


Fig. 5 Analysis of setup slack time

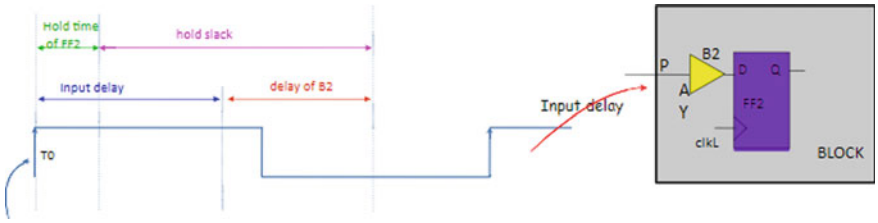


Fig. 6 Analysis of hold slack time

$$ArrivalTime = Inputdelay + Delayof\,B2 \tag{1}$$

$$RequiredTime = Clockperoid - SetuptimeFF2 \tag{2}$$

$$Setupslack = Requiredtime - Arrivaltime \tag{3}$$

Analysis of hold slack time is depicted in Fig. 6.

The arrival time, required time, and slack setup can be calculated using the Eqs. (1), (2), and (3).

4 Implementation Steps

The techniques used to implement the synthesis flow for the RTL of the APB timer design are discussed in this section.

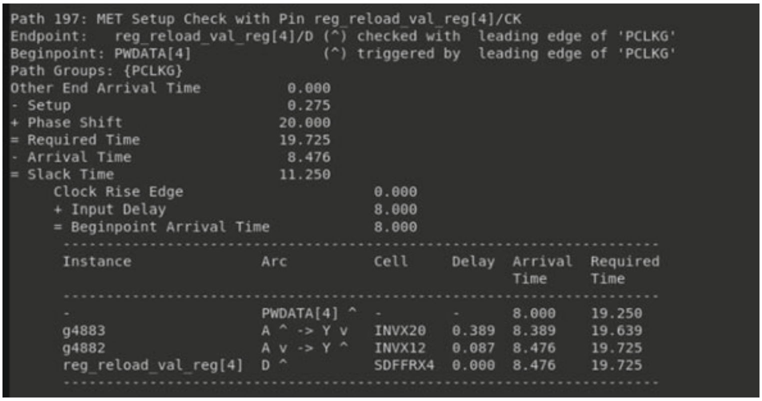
- The clock frequency of 50 MHz is set as a maximum threshold.
- Strategically calculated delays are introduced at the input and output ports.
- Perform the linting inspection. Linting comprises running a program to check the code for potential errors.
- Synthesize the design.
- Create a file containing the output constraints and the gate level netlist (.sdc).
- Generate the reports of timing power and area.

5 Results and Discussions

Timing analysis was carried out using slow.lib for all time pathways, and the difference between timing slack was examined in Figs. 7, 8, 9, and 10.

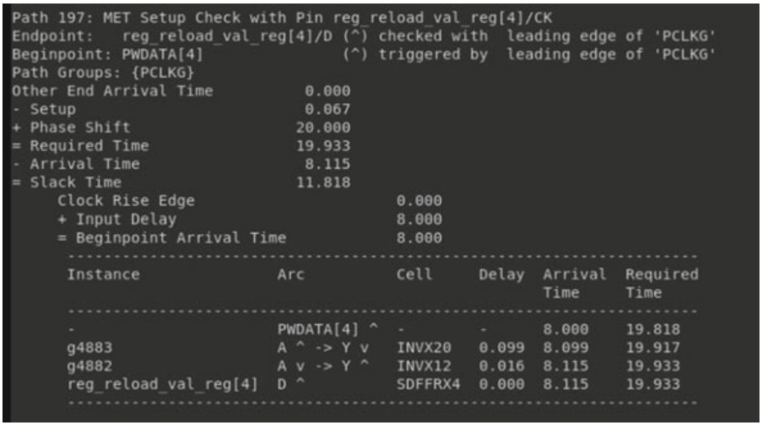
The timing analysis for all the timing paths using Fast.lib and analyzed the difference between timing slack is shown in Fig. 8.

The timing analysis for all the timing paths using Fast.lib and analyzed the difference between timing slack is shown in Fig. 9. Timing path: In2 reg [for generated APB timer netlist] Using Fast Library.



```
Path 197: MET Setup Check with Pin reg_reload_val_reg[4]/CK
Endpoint:   reg_reload_val_reg[4]/D (^) checked with  leading edge of 'PCLKG'
Beginpoint: PWDATA[4]               (^) triggered by  leading edge of 'PCLKG'
Path Groups: {PCLKG}
Other End Arrival Time          0.000
- Setup                         0.275
+ Phase Shift                  20.000
= Required Time                19.725
- Arrival Time                  8.476
= Slack Time                   11.250
     Clock Rise Edge                      0.000
     + Input Delay                        8.000
     = Beginpoint Arrival Time            8.000
      ---------------------------------------------------------------------
      Instance               Arc          Cell      Delay  Arrival  Required
                                                           Time     Time
      ---------------------------------------------------------------------
      -                      PWDATA[4] ^  -         -      8.000    19.250
      g4883                  A ^ -> Y v   INVX20    0.389  8.389    19.639
      g4882                  A v -> Y ^   INVX12    0.087  8.476    19.725
      reg_reload_val_reg[4]  D ^          SDFFRX4   0.000  8.476    19.725
      ---------------------------------------------------------------------
```

Fig. 7 All inputs to registers-setup time using slow.lib



```
Path 197: MET Setup Check with Pin reg_reload_val_reg[4]/CK
Endpoint:   reg_reload_val_reg[4]/D (^) checked with  leading edge of 'PCLKG'
Beginpoint: PWDATA[4]               (^) triggered by  leading edge of 'PCLKG'
Path Groups: {PCLKG}
Other End Arrival Time          0.000
- Setup                         0.067
+ Phase Shift                  20.000
= Required Time                19.933
- Arrival Time                  8.115
= Slack Time                   11.818
     Clock Rise Edge                      0.000
     + Input Delay                        8.000
     = Beginpoint Arrival Time            8.000
      ---------------------------------------------------------------------
      Instance               Arc          Cell      Delay  Arrival  Required
                                                           Time     Time
      ---------------------------------------------------------------------
      -                      PWDATA[4] ^  -         -      8.000    19.818
      g4883                  A ^ -> Y v   INVX20    0.099  8.099    19.917
      g4882                  A v -> Y ^   INVX12    0.016  8.115    19.933
      reg_reload_val_reg[4]  D ^          SDFFRX4   0.000  8.115    19.933
      ---------------------------------------------------------------------
```

Fig. 8 All inputs to registers-hold time using slow.lib

```
Path 117: MET Hold Check with Pin read_mux_byte0_reg_reg[1]/CK
Endpoint:   read_mux_byte0_reg_reg[1]/D (v) checked with  leading edge of 'PCLKG'
Beginpoint: PADDR[6]                    (v) triggered by  leading edge of 'PCLKG'
Path Groups: {PCLKG}
Other End Arrival Time          0.000
+ Hold                          0.010
+ Phase Shift                   0.000
= Required Time                 0.010
  Arrival Time                  8.204
  Slack Time                    8.193
     Clock Rise Edge                      0.000
     + Input Delay                        8.000
     = Beginpoint Arrival Time            8.000
      --------------------------------------------------------------------------
      Instance                   Arc          Cell       Delay  Arrival  Required
                                                                Time     Time
      --------------------------------------------------------------------------
      -                          PADDR[6] v   -          -      8.000    -0.193
      g4899                      A v -> Y ^   INVX16     0.058  8.058    -0.136
      g4898                      A ^ -> Y v   INVX12     0.022  8.079    -0.114
      g9017__4319                D v -> Y ^   NAND4BBX2  0.027  8.106    -0.087
      g8603__3680                B ^ -> Y v   NOR2BX4    0.017  8.124    -0.070
      g8600                      A v -> Y ^   INVX2      0.015  8.139    -0.055
      g8558__5477                C ^ -> Y v   NOR3BX4    0.013  8.152    -0.041
      g8524                      A v -> Y ^   INVX1      0.009  8.161    -0.033
      g8513__1881                A1 ^ -> Y v  OAI221XL   0.043  8.204    0.010
      read_mux_byte0_reg_reg[1]  D v          DFFRX4     0.000  8.204    0.010
      --------------------------------------------------------------------------
```

Fig. 9 All inputs to registers-setup time using fast.lib

```
Path 197: MET Setup Check with Pin reg_reload_val_reg[4]/CK
Endpoint:   reg_reload_val_reg[4]/D (^) checked with  leading edge of 'PCLKG'
Beginpoint: PWDATA[4]               (^) triggered by  leading edge of 'PCLKG'
Path Groups: {PCLKG}
Other End Arrival Time          0.000
- Setup                         0.275
+ Phase Shift                  20.000
= Required Time                19.725
- Arrival Time                  8.476
= Slack Time                   11.250
     Clock Rise Edge                      0.000
     + Input Delay                        8.000
     = Beginpoint Arrival Time            8.000
      ---------------------------------------------------------------------
      Instance               Arc          Cell      Delay  Arrival  Required
                                                           Time     Time
      ---------------------------------------------------------------------
      -                      PWDATA[4] ^  -         -      8.000    19.250
      g4883                  A ^ -> Y v   INVX20    0.389  8.389    19.639
      g4882                  A v -> Y ^   INVX12    0.087  8.476    19.725
      reg_reload_val_reg[4]  D ^          SDFFRX4   0.000  8.476    19.725
      ---------------------------------------------------------------------
```

Fig. 10 All inputs to registers-hold time using fast.lib

5.1 Power Reports Analysis

The power report analysis is generated, and it consists of the power levels of all the components used to design the architecture is shown in Fig. 11.

Category	Leakage	Internal	Switching	Total	Row%
memory	0.00000e+00	0.00000e+00	0.00000e+00	0.00000e+00	0.00%
register	4.52511e-08	4.57592e-05	5.46657e-06	5.12710e-05	45.24%
latch	0.00000e+00	0.00000e+00	0.00000e+00	0.00000e+00	0.00%
logic	1.65764e-07	1.96116e-05	3.96304e-05	5.94078e-05	52.42%
bbox	0.00000e+00	0.00000e+00	0.00000e+00	0.00000e+00	0.00%
clock	0.00000e+00	0.00000e+00	2.64190e-06	2.64190e-06	2.33%
pad	0.00000e+00	0.00000e+00	0.00000e+00	0.00000e+00	0.00%
pm	0.00000e+00	0.00000e+00	0.00000e+00	0.00000e+00	0.00%
Subtotal	2.11015e-07	6.53708e-05	4.77389e-05	1.13321e-04	99.99%
Percentage	0.19%	57.69%	42.13%	100.00%	100.00%

Fig. 11 Power report analysis

5.2 *Comparison of Slack Time Analysis for Fast and Slow Lib*

Tables 1 and 2 show the timing paths slow lib and fast lib for different parameters. The results indicate that the proposed design is efficient and that has minimized the area parallelly.

Table 2 shows the timing path Fast lib for different parameters.

Table 1 Timing paths using slow lib

Sl. No	Timing paths	Slow lib (slack time)		Paths
		Setup check	Hold check	
1	In2reg	11.250	8.792	197-setup 117-hold
2	Regtoreg	19.540	0.510	87
3	Regtoout	11.159	8.933	33
4	Int to out	2.22	17.365	32

Table 2 Timing paths using fast lib

Sl. No	Timing paths	Fast lib (slack time)		Paths
		Setup check	Hold check	
1	In2reg	11.818	8.193	197-setup 117-hold
2	Regtoreg	19.866	0.169	87
3	Regtoout	11.692	8.348	33
4	Int to out	3.398	16.437	32

6 Conclusion

The tabulated data from this research demonstrates that the timing requirements are satisfied in accordance with the design specifications using setup check and holds check without any negative violations. Fast libraries were used to generate the timing routes, as opposed to slow libraries, and early external delay assertions were used to satisfy 33% of setup time and hold time requirements for the reg-to-reg timing path Slack time analysis contributed to enhance design productivity altogether. Power reports were generated with 1.13 W of power using the genus tool to record switching activity and determined with the Power Estimation Tool using TCL scripts. The cadence genus tool is very effective in carrying out synthesis steps [12]. Future design for testability and physical design in chip design can use this synthesized netlist.

References

1. Synopsys (2011) Design compiler user guide
2. Cadence (2019) Genus's user guide
3. Deshpande N et al (2020) A review on ASIC synthesis flow employing two industry standard tools. Int J Eng Res Technol 8(17):257
4. Gayathri S et al (2017) RTL synthesis of case study using design compiler. In: Proceedings of the 2017 international conference on electrical, electronics, communication, computer, and optimization techniques (ICEECCOT), pp 1–7
5. Siva Priya G et al (2019) Static timing analysis and timing violations of sequential circuits. Int J Innov Technol Explor Eng 8(7)
6. Kumar A et al (2020) Static timing analysis of sequential circuit with GUI. In: Proceedings of the 2020 IEEE international women in engineering (WIE) conference on electrical and computer engineering (WIECON-ECE)
7. Madhura R, Jamuna S (2023) Optimized DFT architecture through scan based design. Grenze Int J Eng Technol
8. Madhura R, Jamuna S (2023) Fault simulation for design for testability inserted designs. Indones J Electr Eng Comput Sci 29(2):658–668
9. Krutthika HK, Aswatha AR (2021) Implementation and prevention of congestion in the proposed network on chip architecture. J Tianjin Univ Sci Technol 54(11):213–231
10. Geralla LE et al (2018) Optimization of physically aware synthesis for digital implementation flow. Int J Eng Technol 14:31–35
11. Manasa R et al (2017) Implementation of full adder using single electron transistor, SET: the next generation nano device. IJSTE Int J Sci Technol Eng 3(12)
12. Pavitha US, Nikhila S, Krutthika HK (2012) Design and implementation of image dithering engine on a spartan 3AN FPGA. Int J Fut Comput Commun 1(4):361–365

Dictionary-Based PLS Approach to Pharmacokinetic Mapping in DCE-MRI Using Tofts Model

Krutthika Hirebasur Krishnappa and Nithin Vajuvalli Narayana Gowda

Abstract Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) is an approach that employs MRI to monitor changes in blood flow and tissue perfusion following the injection of a contrast agent. DCE-MRI may be used to calculate PharmacoKinetic (PK) maps and disclose specifics about the manner in which the body distributes and eliminates medications. These maps can be employed to improve drug dosage and treatment planning, as well as to comprehend the pharmacokinetics of medicines in various tissues. To improve the speed and accuracy of PK map estimation, a dictionary-based Partial Least Squares (PLS) approach based on the Tofts model has been proposed. The Tofts model is a widely used PK model that describes the transport and metabolism of a contrast agent in tissue. The dictionary-based PLS approach utilizes a pre-constructed dictionary of Tofts model parameters, which can be used to speed up the PK map estimation process. Instead of fitting the Tofts model to the DCE-MRI data from scratch, the algorithm uses the pre-constructed dictionary to identify the most likely set of parameters that best fit the data. This can significantly reduce the computational time required for PK map estimation, making it more practical for clinical use. These techniques were assessed using simulation model. The results show the PLS approach was able to estimate PK maps accurately for simulated. The low Root Mean Square Error (RMSE) suggests that error will fall within the permissible range of tolerance. The computational time for the PLS approach was found to be significantly less than TM-TD and TM-FD approaches. As a result, the proposed TM-PLS approach, which is based on Tofts model, exhibits promising performance in terms of computational time, and errors that are moderately higher than the traditional curve fitting model for the determination of PK maps in DCE-MRI, which is a crucial parameter for the diagnosis and treatment planning.

Keywords PharmacoKinetic map · Partial least square · Tofts model · Dynamic contrast enhancement MRI

K. H. Krishnappa (✉)
Department of Computer Science, Southern University and A&M College, Baton Rouge, LA, USA
e-mail: krutthika.hirebas@sus.edu

N. V. N. Gowda
Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_21

1 Introduction

Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) is a technique that uses MRI to measure the changes in blood flow and tissue perfusion in response to the injection of a contrast agent [1–3]. DCE-MRI can be used to estimate PK maps, which provide information about the distribution and clearance of drugs within the body. These maps can be used to understand the pharmacokinetics of drugs in different tissues, and optimizes drug dosing and treatment planning. The PK maps are generated by fitting a pharmacokinetic model to the concentration of the contrast agent inside the tissues with time, which can be obtained from the DCE-MRI datasets [4–8]. These models are constructed using mathematical equations that illustrates, in what way the contrast agent is transported and metabolized in the tissue. The two-compartment method, the population-based models, and the enhanced Tofts approach are frequently used as PK models [9, 10]. To improve drug dosage and treatment planning in clinical research, it is essential to efficiently and reliably predict PK maps. One approach that is proposed to improve the speed of PK map estimation is the use of a dictionary-based Partial Least Squares (PLS) approach. This technique is based on the Tofts model, which widely uses pharmacokinetic model to describe the transport and metabolism of a contrast agent in tissues.

The dictionary-based PLS approach utilizes a pre-constructed dictionary of Tofts model parameters, which can be used to enhance the PK map estimation process. Rather than beginning all over and fitting the Tofts model to the DCE-MRI data, the algorithm uses the pre-constructed dictionary to identify the most likely set of parameters that best fits the data. As a consequence, the computing time needed for PK map estimation may be significantly reduced, making it more useful for clinical application.

The approach additionally makes use of PLS, a statistical method that can be used to identify the most crucial associations between both independent and dependent variables. By minimizing dimensionality of the data and emphasizing on the most valuable information, this can further enhance the accuracy and efficiency of the PK map estimation process. The potential approach for a faster and more efficient PK map estimation in clinical practice is the dictionary-based PLS approach based on Tofts model. To increase the efficiency and precision of estimate procedure, it employs utilization of a pre-built dictionary of Tofts model parameters and PLS.

2 Related Work

There have been numerous studies and developments in the domain of DCE-MRI modeling for the estimation of pharmacokinetic (PK) maps.

The authors in paper [11] showed a frequency-domain technique might be combined with the traditional Tofts model to expedite the curve fitting procedure and acquire pharmacokinetic properties. The research findings demonstrated that, the curve fitting technique is significantly more beneficial when employing the frequency-domain method. A dataset of patients with breast tumors served as the testing data by the authors of this paper [12]. The rankings generated by each QI were evaluated to reference rankings of various MCT determined by utilizing the tumor segmentation findings for each patient. The proposed approach by the authors in paper [13] was tested using an abdominal DCE phantom. For fourfold and eightfold under-sampled data, the estimated kinetic parameters employing the recommended strategy improved with ground truth kinetic constraints by 19% compared to the indirect method, and by 7% for fully sampled data. The authors in paper [14] have used an NVIDIA GeForce GTX 1080 Ti GPU to efficiently infer the TK maps of hidden DCE-MRI images at a rate of 0.34 s per slice. The authors in paper [15] have shown that their techniques can generate precise tracer kinetic maps. The estimated variable values were in remarkable alignment with the observed tracer kinetic parameter ranges for signal and noise realizations and regression techniques. The authors in [15] developed a generalized tracer kinetic model selection framework with the goal of measuring the microvascular characteristics and accomplished 82.1% of vertices from patients with non-tumorous livers, while 32.2% of vertices from tumors.

The authors in paper [16] have shown that a system having intra-axonal, myelin, and extra-axonal compartments. They were able to exchange myelin effectively using spiraling, which allowed for small diameters and very few wraps. The authors in article [17] has implemented fitting tracer kinetic models using linear approaches that is considerably better than using nonlinear equivalents. In the paper [18], the research examines the application of physics-based deep neural networks to predict physiological parameters as of DCE-MRI signal-curves. Models were used to assess the precision and accuracy of parameter estimation by the temporal systems. Each network achieved more reliability than NLLS.

The authors in paper [19] have shown how the legitimate web-based front end can be designed to translate rules into commands to detect and prevent the host system from different layers of attacks.

3 Proposed Work

DCE-MRI for PK mapping may improve the effectiveness of drug dose and treatment planning. Also, the computational time required for an accurate PK map estimate may be a drawback to clinical use. In comparison to traditional curve fitting methods, the suggested dictionary-based Partial Least Squares (PLS) technique based on the Tofts model demonstrates promising performance in terms of processing time and acceptable errors. Applications of this approach will improve the efficacy and accuracy of PK map estimate, which would eventually lead to improved clinical evaluation and treatment planning.

3.1 In Silico Simulation Data

Forward model was used to generate Concentration Time Curves (CTCs) to show how well the PLS technique compares to the traditional Tofts model in time domain (TM-TD). To achieve this, 1000 *K*trans and 1000 Ve values were used to build a dictionary of CTCs, with step size varying from 0.001 to 1 (equally spaced). Averaged population-based Arterial Input Function was used from the literature [20] to employ the forward simulation by taking 32 time points with temporal resolution of 5 s to generate the CTCs. The β coefficient was obtained by using these forward simulated CTCs and randomly generated pharmacokinetic maps (*K*trans and Ve). The resultant β was used to produce the *K*trans and Ve maps and the approach was named TM-TD PLS. In order to obtain the optimum β, Number of COMponents (NCOMP) were varied from 1 to 32. Using the technique of Root Mean Square Error (RMSE), the error between the GT and PK maps generated from TM-TD PLS for the various component values was estimated, and the optimal value β was determined. The 21 PLS components with the lowest RMSE value were determined using β a dictionary test.

Random simulation were employed to generate the two different types of simulation data. In the random simulation, CTCs were dynamically simulated by using population-based AIF to generate the K^{trans} and Ve at random within the 0–1 range. The coefficient obtained for the random simulation was used to generate PK maps for all of the simulated data using the TM-TD PLS. Using TM-TD, TM-FD, and TM-PLS, CTCs were applicable for two parameters (K^{trans} and Ve). This served as the simulated data's estimated value (EV). The GT was employed to provide simulated data. The accuracy of the EV and GT value for the PK maps using all three methods was evaluated using RMSE. GT PK maps against PLS PK maps were linearly plotted to determine the accuracy of the PLS approaches.

4 Performance Evaluation

The performance is evaluated using Quantitative Imaging Biomarkers Alliance (QIBA) simulated data and randomly simulated data approaches.

4.1 Simulation Results

The PK maps derived from PLS, Tofts model-time-dependent (TM-TD) and Tofts model-frequency-dependent (TM-FD) approaches for randomly simulated data are presented in Fig. 1. As can be seen from Fig. 1 difference maps, all three techniques produced PK maps that were comparable to the Ground Truth (GT).

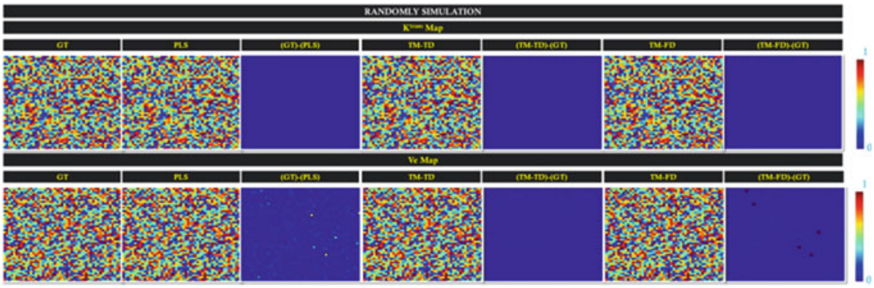


Fig. 1 50 × 50-size pharmacokinetics maps with randomly generated

Figure 2 presents the scatter plot of the correlation analysis between the simulated PK maps and the PLS obtained maps. It can be observed that for the PLS approach provided the best fit, the rectification of the PLS obtained *K*trans map in the randomly simulated data was acceptable. It can be noticed in the Fig. 2 that for the random generated Ve maps, the PLS approach has significant differences with correlation analysis. This implies that PLS technique performed well in terms of estimating the *K*trans parameter for randomly generated data, yet there is a considerable discrepancy between the randomly generated Ve maps and the Ve maps produced by the PLS approach.

Computational time and RMSE for randomly generated datasets with various matrix sizes are shown in Fig. 3a, b. It is apparent that TM-TD strategy takes more time to compute but produces less error, while the TM-FD technique takes less time to compute but yielded relatively higher RMSE values.

Figure 3 depicts RMSE plot for various numbers of components (NCOMP) values. It can be seen that the error is less and more stable in the 15–22 ranges, yet it increases as the NCOMP accelerates. This illustrates that the accuracy and computational time of PLS technique can be balanced by using a specific range of NCOMP values (Fig. 4).

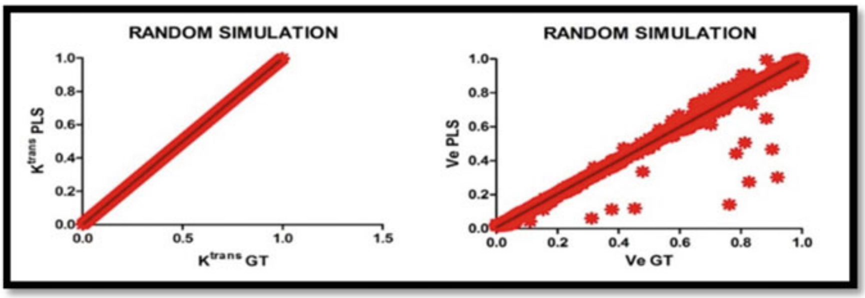
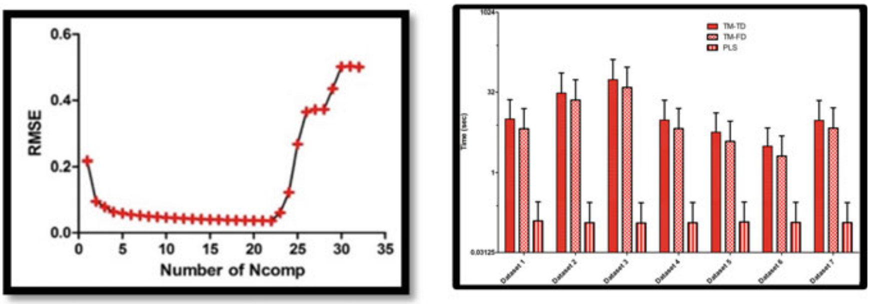


Fig. 2 Depicts the scattered plot of the results of correlation analysis between randomly simulated PK maps and PLS PK maps



(a) (b)

Fig. 3 **a** Depicts the RMSE plot for the different number of components to obtain β coefficients. **b** TM-TD, TM-FD, and PLS computing times for 7 various matrix datasets are compared

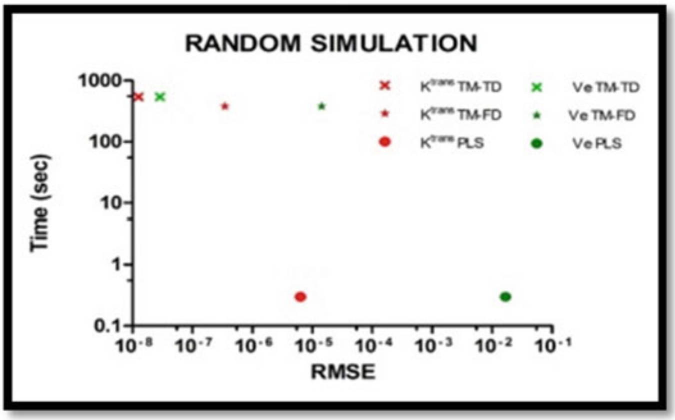


Fig. 4 Computational time against RMSE plots for 3 techniques data generated by random simulated data

The computation duration for obtaining PK maps from three approaches is shown in Table 1 in a tabular format. It is apparent that using the PLS approach resulted in equivalent PK maps but using 95–98% lesser time. This exemplifies how the PLS method produces precise PK maps while consuming minimal computation time.

Table 1 Table captions should be placed above the tables

Dataset	Time taken (sec)			Perfect difference (%)		RMSE	
	TM-TD	TM-FD	PLS	$TM_{TD\text{-}FD}$	$TM_{TD\text{-}PLS}$	$TM_{TD\text{-}FD}$	$TM_{FD\text{-}PLS}$
Dataset 1	19.12	13.00	0.23	32	98.7	0.027	0.032
Dataset 2	60.98	45.35	0.22	25.6	99.6	0.067	0.077

5 Conclusion and Future Scope

For simulated data, the difference map for the PLS technique, as illustrated in Fig. 1, was determined to be quite modest. This can also be observed in the scatter plot correlation analysis depicted in Fig. 2. The PLS technique was successful in accurately estimating the PK maps for the simulated data, as shown by the small difference map and high correlation in the scatter plot.

The RMSE and computational time comparisons shown in Fig. 3a, b demonstrated that Tofts model-partial least squares (TM-PLS) approach performs more effectively than the conventional curve fitting model. The TM-PLS approach can generate PK maps more efficiently than the traditional curve fitting model while still having errors that are acceptable.

In conclusion, for the determination of PK maps in DCE-MRI, an important parameter for diagnosis and treatment planning, the proposed TM-PLS approach based on the Tofts model exhibits promising outcomes in terms of computational time, and errors that are acceptable in comparison to the traditional curve fitting model. As related to the Tofts model-partial least squares (TM-PLS) technique, the Tofts model-time-dependent (TM-TD) and Tofts model-frequency-dependent (TM-FD) procedures are known to execute iterative curve fitting for the model.

Some limitations of the TM-PLS approach for PK maps determination in DCE-MRI can be observed based on the data presented. According to the minuscule difference map and strong correlation observed for the simulated data, the PLS approach can accurately estimate PK maps in the absence of noise; however, its efficacy on real-world data containing noise may not be assured. The approach's applicability to real-world data must thus be further validated in order to determine its therapeutic use. Finally, it is uncertain how well this approach would accomplish reducing noise in real-world data, despite the fact that the maps may be denoised using PCA by making the most use of the number of factors during the PLS application. Therefore, additional research is necessary to ascertain the effectiveness of this technique for denoising in DCE-MRI.

References

1. Huang W et al (2014) Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge. Transl Oncol 7(1):153–166
2. Brix G et al (2004) Microcirculation and microvasculature in breast tumors: pharmacokinetic analysis of dynamic MR image series. Magn Reson Med 52(2):420–429
3. Gianfelice D et al (2003) MR imaging-guided focused ultrasound surgery of breast cancer: correlation of dynamic contrast-enhanced MRI with histopathologic findings. Breast Cancer Res Treat 82(2):93–101
4. Evelhoch JL et al (1999) Key factors in the acquisition of contrast kinetic data for oncology. J Magn Reson Imag 10(3):254–259
5. Hylton N et al (2006) Dynamic contrast-enhanced magnetic resonance imaging as an imaging biomarker. J Clin Oncol 24(20):3293–3298

6. Li X et al (2014) DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: pilot study findings. Magn Reson Med 4(71):1592–1602
7. Jahng GH et al (2014) Perfusion magnetic resonance imaging: a comprehensive update on principles and techniques. Korean J Radiol 15(5):554–577
8. Gillies RJ et al (2000) Applications of magnetic resonance in model systems: tumor biology and physiology. Neoplasia 2(1–2):139–151
9. Yankeelov T et al (2007) Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples. Curr Med Imag Rev 3(2):91–107
10. Khalifa F et al (2014) Models and methods for analyzing DCE-MRI: a review. Med Phys 41(12):8–9
11. Vajuvalli NN et al (2014) Accelerated pharmacokinetic map determination for dynamic contrast enhanced MRI using frequency-domain based tofts model. Annu Int Conf IEEE Eng Med Biol Soc 2404:7
12. Piantadosi G et al (2015) Data-driven selection of motion correction techniques in breast DCE-MRI. In: Proceedings of the 2015 IEEE international symposium on medical measurements and applications (MeMeA) proceedings, pp 07–09
13. Dikaios N et al (2013) Direct parametric reconstruction from under sampled (k, t)-space data in dynamic contrast enhancement MRI. In: Proceedings of the 2013 IEEE nuclear science symposium and medical imaging conference
14. Kettelkamp J et al (2020) Arterial input function and tracer kinetic model-driven network for rapid inference of kinetic maps in dynamic contrast-enhanced MRI (AIF-TK-net). In: Proceedings of the 2020 IEEE 17th international symposium on biomedical imaging (ISBI), pp 03–07
15. Bliesener Y et al (2020) Efficient DCE-MRI parameter and uncertainty estimation using a neural network. IEEE Trans Med Imag 39(5):1712–1723
16. Kallehauge JF et al (2017) Comparison of linear and nonlinear implementation of the compartmental tissue uptake model for dynamic contrast-enhanced MRI. Magn Reson Med 77(6):2414–2423
17. Zou J et al (2020) Estimation of pharmacokinetic parameters from DCE-MRI by extracting long and short time-dependent features using an LSTM network. Med Phys 47(8):3447–3457
18. Parker GJM et al (2006) Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. Mag Reson Med 56(5):993–1000
19. Krutthika HK, Nikhila S, Pavitha US (2013) Development of CGI based front end design for implementation of security policies and application layer filtering. Int J Adv Comput Theory Eng 2(5)
20. Nikhila S, Pavitha US, Krutthika HK (2014) Face recognition using Wavelet transforms. Int J Adv Res Elect Electr Instrum Eng IJAREEIE 3(1)

User Experience Evaluation for Pre-primary Children Using an Augmented Reality Animal-Themed Phonics System

T. V. Sumithra and Leena Ragha

Abstract With smart phones being ubiquitous in every household, people are increasingly relying on educational applications to simplify the learning process of children. This research is focused on creating a child-friendly framework to evaluate the user experience of preschool children aged between three to five years through an augmented reality (AR) application. Despite the rapid growth and importance of AR applications in various fields, providing a positive user experience for pre-primary students has proven to be a challenge. AR applications have yet to be fully implemented for academic purposes. The aim of our work is to introduce an application that enhances phonics teaching methods and encourages students to engage actively in learning activities. Creating effective AR experiences for children between the ages of three to five can be a difficult task. Designers must take into account the cognitive and motor skill limitations of children in this age group. Understanding the target audience is crucial to developing an educational application that is both entertaining and appealing to students and teachers alike. To address this challenge, we established an evaluation framework for testing AR applications focused on phonics, using animal-themed content. This framework was developed using various techniques such as interviews, literature research, and persona creation. As a result of this process, we formulated and also evaluated the effectiveness of the theoretical framework for assessing the user experience of young learners.

Keywords Augmented reality · Phonics system · User experience · Immersive technology · Pre-primary education

T. V. Sumithra · L. Ragha (✉)
Ramrao Adik Institute of Technology, D. Y. Patil Deemed to Be University, Nerul, Navi Mumbai, Maharashtra, India
e-mail: leena.ragha@rait.ac.in

T. V. Sumithra
e-mail: sumithratv@rait.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_22

1 Introduction

For an application to be successful, it should fulfill the requirements of a broad spectrum of users. Three decades ago, prior to the advent of iPhones, iPads, high-speed Internet, and powerful personal computers, it was primarily adults who accounted for the vast majority of advanced information technology usage around the world [1–3]. In the past, children had limited access to technology such as computers, phones, or tablets, which resulted in less time spent using technology compared with adults. As a result, digital products and experiences were primarily designed with adults in mind. However, as technology use has become more prevalent among younger age groups, businesses must now create goods and services that cater to children's needs [4]. Despite this seemingly straightforward task, designing AR applications for children differs significantly from designing for adults [5, 6]. To create a well-designed digital experience that appeals to children, design issues, such as low cognitive ability, weak motor skills, individual growth, and a lack of past experience must be taken into consideration. Our work focuses on developing a theoretical framework to assess the user experience of children aged between three to five years, taking into account these unique design challenges. Proposed evaluation framework is tested using an animal-themed AR phonics system to evaluate the user experience. AR technology enables computer-generated visuals to be superimposed onto the real world through a smartphone camera, creating an interactive and dynamic learning environment for students. In this work, our aim is to establish a framework for evaluating the user experience of pre-primary children using an animal-themed AR phonics application. This application provides an immersive experience by allowing students to interact with 3D models in real time, with well-defined textures and controls for better visualization. By incorporating AR technology into education, students can enhance their learning and memorization abilities, leading to more effective and engaging educational experiences when compared with the existing teaching methodologies like animations and video explanation.

The animal-themed phonics application is designed to provide children with an attractive and user-friendly graphical user interface, focused on enhancing their phonics learning experience through immersive and interactive content. In the pre-primary education sector, traditional teaching approaches involve using textbooks, videos, and images as references for students. However, as technology has advanced, teaching methods have not kept up, and textbooks are only useful for reading content. When it comes to real-life or fantasy world images, textbooks lack the perspective required for an overall understanding of the subject matter. To promote interactive learning in pre-primary education, we have developed a user experience framework and crafted an augmented reality animal-themed phonics system application. The primary objective of our work is to propose an evaluation framework for measuring the user experience of children aged between three to five years by designing an animal-themed phonics system to make learning interactive and enjoyable for pre-primary students and enhance the learning enthusiasm in children. The three levels of processing; a cognitive model given by 'Donald Norman' was referred to design the phonics application. The three levels of processing have a big influence on human computer interaction and user experience design. Hence, we took a very close look at the three levels of processing, starting with the lowest level called visceral level, then the behavioral level and finally the reflective level [7].

2 Related Work

Designers of AR applications may potentially become users of the products they help develop. However, when it comes to designing for children, this is not always the case. The process of designing for children requires specific usability approaches, such as targeting content to meet the needs of children of different ages [8]. Unlike adults, children have different cognitive processes for both thinking and doing. They are not potential customers or consumers of the product [6, 9, 10]. In the past, children were not a significant customer base for technological products due to high costs, difficulty in operating, and the focus on productivity rather than entertainment. However, in today's world, children use technology on a daily basis, and some even more than adults [2, 3]. This is evident from the vast number of applications available on the Google Play and Apple App Store. For instance, we conducted a search on the play store and found one relevant application known as 'SNAP LEARN', which offers an AR experience for young readers. The application includes a textbook with various objects presented in story formats. Users can interact with any object by scanning it. However, this application requires children to have a certain level of knowledge, including learning alphabets and phonics, before using it. Books have been the main source for studying and learning about the world for centuries. They have been a reliable way to share knowledge across the globe. In recent times, there have been several works done in the field of AR for pre-children education and educational psychology [11–15]. These works have shown the effectiveness of using AR systems for learning. In the digital age, incorporating augmented reality into educational materials, such as converting a book into an AR animal phonics resource can greatly enhance students' understanding of the subject. However, some of the limitations of the AR apps we surveyed on the play store include not being suitable for children aged 3–5, requiring payment to access premium AR features, and providing poor texture quality, not so user-friendly. For children between the ages of 3 and 5, learning about animals from books can be challenging. Books may become tedious over time, which could create obstacles for future learning. When reading a book, a reader is limited to a single perspective provided in 2D format, which doesn't give an accurate idea of how the actual image would appear in reality.

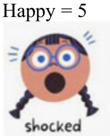
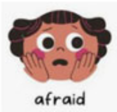
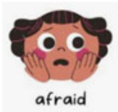
Hence, the phonics application's design would offer children a 3D perspective of animals, enhancing their interactive experience. By seeing and hearing phonics sounds, children can comprehend better and retain word pronunciations more effectively. The phonic system is designed to introduce animals for the entire alphabetic range from A to Z and teach phonics using audio sound that involves repetition of pronunciations to ensure better recall ability. Previous work carried out [16–20], did not focus on measuring user experience relevant to age. We have proposed a simple and effective user experience metric for measuring user experience for pre-primary children in using AR application.

3 Methodology

The goal of our proposed system is to enable pre-primary students to visualize 3D models of animals through an augmented reality application. Augmented reality technology allows computer-generated graphics to be overlaid onto real-world environments on a screen, transforming the way we interact with visual experiences on mobile apps. Our test bed application, named the AR animal-themed phonics system, is designed for the educational domain, with a user-friendly interface that enables easy identification of animals and their corresponding alphabets. Students can learn the correct pronunciation of each animal's phonic associated with its corresponding alphabet, as well as view accompanying text. The application is specifically designed for children aged between 3 and 5 years old, with icons representing each animal for ease of use. By clicking on an icon, a specific animal corresponding to the alphabet will be augmented onto the screen in 3D, along with audio explanation. Audio explanations are provided in multiple languages, with current support in English and Hindi. To enhance the realism of the 3D animals, the system includes animal sounds to help students understand how these animals actually sound in nature. In addition to the user experience framework, our application includes a quiz feature to assess learning outcomes and test how much the child has learned about each animal, which would benefit the teachers. To assess student learning outcomes, an AR quiz feature has been added, which displays red and green markers for wrong and right answers, respectively. This quiz feature is suitable for children aged 5 years and above. It fosters a sense of achievement and enthusiasm for learning, which can enhance the child's recall abilities. The 3D animals are also accompanied by audio explanations with an Indian accent, making it easier for students to learn. Furthermore, the addition of an animated graphical user interface with interactive icons makes the application more engaging and improves visual retention for students.

Image tracking technology is an essential component of the AR phonics system. This technology enables the system to detect a surface through the camera where the 3D animal models can be placed. Once the surface is detected, the animal is spawned and students can interact with it closely. The AR animal-themed phonics system is an Android application that was developed by undergraduate students using Unity and AR concepts based on the provided user interface design. The traditional heuristic evaluation, as described by authors in the context of human–machine interaction [21–23], would be difficult for users in the age group of 3–5 years to comprehend and appreciate the importance of usability questions. Hence, we developed a user experience evaluation framework with metrics suitable for this age group. Table 1 presents the metric parameters, and we chose emotion pictures as our evaluation metric since young children are drawn to visual aids.

Table 1 User experience evaluation framework for age group 3–5 years for animal-themed phonics system

Round-1 (x)	Round-2 (y)	Teacher's observation (z)	Summation of student experience
Round-1 (x)	Round-2 (y)		
 Happy = 5	 Happy = 5		
 Shocked = 3	 Shocked = 3		
 Afraid = 3	 Afraid = 3		
 Content = 4	 Content = 4	Likert scale metric 5, 4 , 3, 2 and 1	Summation, $S = (x + y + z)/3$
 Worried = 3	 Worried = 3		
 Shy = 2	 Shy = 2		
 Sleepy = 1	 Sleepy = 1		

We used a five-point Likert scale to measure the agreement of user experience question statements, with 5 indicating 'strongly agree', 4 indicating 'agree', 3 indicating 'neutral', 2 indicating 'disagree', and 1 indicating 'strongly disagree'. The pictures were assigned values based on the Likert scale reading, with 'happy' mapped to 5, 'content' mapped to 4, 'shocked', 'afraid', and 'worried' mapped to 3, 'shy' mapped to 2, and 'sleepy' mapped to 1. Table 1 presents the pictures and their corresponding values. We also considered the teacher's feedback as a metric to supplement our user experience data. Since children are exposed to the AR application at an early stage of their education and may not be familiar with user studies, we divided the evaluation into two rounds labeled ‘x’ and ‘y’ and introduced a third round labeled ‘z’ to reflect the teacher’s observation. The final user experience was measured as the average of the three rounds, i.e., $S = (x + y + z)/3$.

4 Results

In this section, we present the outcomes of the AR animal-themed phonics application and the corresponding user experience. Figure 1 displays the splash screen along with the home page and different icons. Each icon has distinct functionalities, for instance, the thumbs up icon represents the suggestion page, the foot impression icon represents the exit button, the arrow icon represents the navigation, and the user question icon represents the user manual.

Figure 2 presents an intuitive user interface that enables easy identification of animals and their corresponding alphabets through a slider containing animal alphabets from A to Z. Figure 3 illustrates how the application enables students to grasp the phonetic sound associated with each animal alphabet through a text display and an audio explanation. By clicking on any animal icon corresponding to the alphabet, the 3D animal is augmented on the screen along with the audio explanation, which is available in multiple languages. The languages are enhanced with spoken Indian accents to make it easier for students to understand and overcome language barriers.

Fig. 1 Home page screen

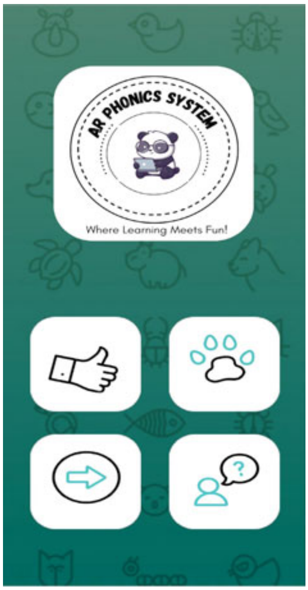


Fig. 2 User-friendly user interface design

Fig. 3 Phonic alphabet audio and animal visualization

Furthermore, to make the 3D animals appear more realistic, an option of animal sounds has been added to help students understand how these animals sound in nature. Figure 4 demonstrates an AR quiz feature added to evaluate students' learning outcomes, with red and green markers indicating incorrect and correct answers, respectively, and the score updated automatically for each correct answer. This feature is more suitable for senior pre-primary section students. The user experience study was performed with a sample size of 20 children, as described in the framework above. While a larger sample size could have yielded more robust results, Faulkner [24] demonstrated that 98% of usability issues could be identified with a sample size of 20. Table 2 summarizes the percentage of known usability problems identified in 100 analysis samples.

Based on the user experience study conducted with twenty children, the mean score on a 5 point Likert scale was found to be 4.81 with a variance of 0.02. The user interviews showed that most participants enjoyed learning with the animal-themed phonics system. However, as user experience design is an iterative process, we identified some valuable factors that could be improved to make the screen design more user-friendly. Therefore, we will closely examine the screen redesign process in our future work.

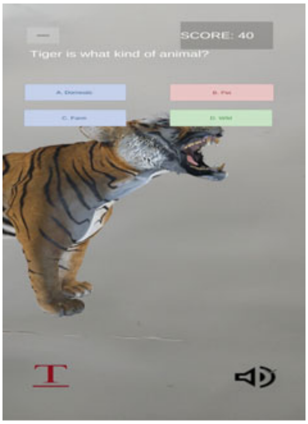


Fig. 4 AR quiz system

Table 2 Percentage of total known usability problems found in 100 analysis samples

No. users	Minimum % found	Mean % found
5	55	85.55
10	82	94.686
15	90	97.050
20	95	98.400
30	97	99.00
40	98	99.600
50	98	100.00

5 Conclusion

The animal-themed phonics system using augmented reality allows for faster learning and memorization by presenting 3D animal models to students, along with audio explanations in English and Hindi, and an AR quiz feature. A novel framework is developed to measure the effectiveness of this AR application for children aged 3–5 years. The study experiment showed that users were satisfied with the system and found it to be engaging, creating a sense of achievement and enthusiasm for learning phonics. Moving forward, we plan to expand the application to include detailed information about each animal model, with support for multiple languages.

Acknowledgements Our gratitude goes to our undergraduate students, Shweta Sarjine, Suchita Kudke, Sahil Dhamnaskar, and Kshitija Shinde for implementing the test bed, AR phonics application in accordance with the provided design guidelines, as well as validating the application. We thank the teachers and children who participated in the usability study.

References

1. Swedish Media Council (2015) Facts about parents attitudes towards childrens of media usage (fakta om föräldrars attityder till barns medieanvändning). Parents & Media (Föräldrar & Medier)
2. Swedish Media Council (2015) Facts about kids's usage and experiences of media (fakta om små barns användning och upplevelser av medier). Kids & Media (Småungar & Medier)
3. Swedish Media Council (2015) Facts about children's usage and experiences of media (fakta om barns och ungas användning och upplevelser av medier). Children & Media (Ungar & Medier)
4. Tegegne W (2015) The use of dialects in education and its impacts on students' learning and achievements. Educ J 263–269. https://doi.org/10.11648/j.edu.20150405.22
5. Gelman DL (2014) Design for kids: digital products for playing and learning. Rosenfeld Media Brooklyn
6. Sofia Persson at SVT interactive. Personal interview, 2016
7. https://www.interaction-design.org/literature/article/norman-s-three-levels-of-design
8. Nielsen J (2016) Children's websites: usability issues in designing for kids. https://www.nngroup.com/articles/childrens-websites-usability-issues/. Accessed: Feb 2016
9. Chris Lindgren and Petter Karlsson at Toca Boca. Personal interviews, 2016
10. Linus Nilsson at North Kingdom. Personal interview, 2016
11. Rui W (2022) Application of augmented reality technology in children's picture books based on educational psychology
12. Raiyn J, Computer Science Department, Al Qasemi Academic College for Education, Baqa El Gharbia, Israel (2016).
13. Sirakaya M, Sirakaya DA (2018) Malays Online J Educ Technol
14. Kubiatko M (2012) Kindergarten children's perception of animals focusing on the look and fear of animals
15. Jeffries ERW (2016) Children's developing awareness of regional accents: a socioperceptual investigation of pre-school and primary school children in York
16. Aydoğdu F (2021) Augmented reality for preschool children: an experience with educational contents

17. Xiaojing L, Zhang J, Jing H (2016) A case study of phonics among primary school students. Int J Innov Educ Res 4(10):11–16. https://doi.org/10.31686/ijier.vol4.iss10.593
18. Chen IC (2018) The application of augmented reality in english phonics learning performance of ESL young learners, pp 255–259. https://doi.org/10.1109/IC3.2018.000-7
19. Mamat N (2019) Implementing augmented reality (AR) on phonics-based literacy among children with autism. Int J Adv Sci Eng Inf Technol
20. Syahidi AA, Tolle H, Supianto AA, Arai K (2019) AR-child: analysis, evaluation, and effect of using augmented reality as a learning media for preschool children. In: 2019 5th international conference on computing engineering and design (ICCED), pp 1–6. https://doi.org/10.1109/ICCED46541.2019.9161094
21. Nonnan D, Miller J, Henderson A (1995) What you see, some of what's in the future, and how we go about doing it: hi at apple computer
22. Dorman D, Nielsen J (2016) The definition of user experience. https://www.nngroup.com/articles/definition-user-experience/. Accessed: Apr 2016
23. Nielsen J (2010) Children (ages 3–12) on the web, 3rd edn. Nielsen Norman Group
24. Faulkner (2003) Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 35(3):379–383

Blockchain-Based Secure Cloud Data Management: A Novel Approach for Data Privacy and Integrity

Gaurav Gogisetty, Aryaman Parihar, Ramraj Dangi, and Prashant Vats

Abstract Due to its potential to increase trust, security, and transparency across numerous industries, decentralized apps and smart contracts are growing in popularity. A network of peer-to-peer computers called Inter Planetary File System (IPFS) enables decentralized content exchange and storage. As smart contracts allow for the automated and safe execution of established rules and logic, they can improve IPFS functionality when implemented on blockchain networks. This study compares the support for smart contracts on IPFS provided by various blockchain networks. The comparison comprises Polkadot, Binance Smart Chain, Solana, and Ethereum. When utilized to deploy and carry out smart contracts over IPFS, these networks' scalability, security, and transaction costs are the main subjects of the study. A thorough literature examination of the various blockchain networks and the elements connected to them, including as consensus methods, gas costs, and IPFS compatibility, forms the basis of the research technique employed in this study. The study's conclusions show that each blockchain network has different strengths and shortcomings when it comes to enabling smart contracts on IPFS. Despite having higher transaction costs, more established networks like Ethereum and Solana provide greater degrees of security and scalability. Although it has cheaper transaction costs than Ethereum and Solana, Binance Smart Chain is less secure. Although less mature and more sophisticated than Polkadot, it can handle numerous parallel chains and offers greater compatibility with other the use of blockchains.

Keywords Security · Blockchain · Smart contracts · Cloud computing authentication · Polkadot · Solana · Ethereum

---

G. Gogisetty (✉) · A. Parihar · R. Dangi · P. Vats
Department of CSE, SCSE, Manipal University Jaipur, Jaipur, Rajasthan, India
e-mail: gaurav.gogisetty2002@gmail.com

A. Parihar
e-mail: parihararyaman0@gmail.com

P. Vats
e-mail: Prashant.vats@jaipur.manipal.edu

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_23

1 Introduction

A form of storage for information architecture known as decentralized file storage systems distributes data over multiple nodes in a peer-to-peer (P2P) network. Decentralized file storage systems spread data over several network nodes, enhancing its security, fault tolerance, and resistance to censorship in contrast to conventional centralized storage systems, which store data on a single server, or a group of servers managed by a single institution [1]. To guarantee data integrity, availability, and redundancy, decentralized file storage systems employ a range of technologies, including blockchain, distributed hash tables (DHTs), and peer-to-peer (P2P) protocols [2]. By permitting simultaneous downloads from numerous sources rather than depending just on one server, IPFS also makes it possible for material to be distributed more effectively. This increases its censorship resistance and guarantees that material is always accessible, even if some nodes go down [3]. In general, IPFS represents a significant change in the way we view the Internet and the dissemination of information. IPFS has the ability to dramatically transform the way we think about file sharing, social networking, and online commerce by offering a totally decentralized and content-addressable system for information sharing and storing [4]. Smart contracts and blockchains are two interrelated ideas that have revolutionized several fields, including technology and finance. A blockchain is a distributed, decentralized digital ledger that securely and openly records transactions [5]. It is composed of a network of node locations that cooperate to validate interactions, and every transaction is stored in a block that is connected to the block before it in the chains, hence the name "blockchain". Smart contracts are very secure, effective, and transparent since they can facilitate, verify, and enforce the negotiation and fulfillment of a contract [6]. As a result, once they are published on the blockchain, they cannot be changed without the network's consensus. They are created to be tamperproof. The Ethereum blockchain today serves as the most popular blockchain platform for creating and implementing smart contracts. Smart contracts are nonetheless supported by a number of different blockchain systems, including Polkadot, Solana, and others. Financial contracts, supply chain management, digital identity management, and other scenarios are just a few of the many applications for smart contracts [7]. By lowering costs, enhancing transparency, and getting rid of middlemen, they have the potential to revolutionize conventional corporate procedures. They are also extremely secure and effective [8]. Smart contracts and decentralized apps (DApps) may be developed on Ethereum, a decentralized blockchain platform. Self-executing computer program known as "smart contracts" allow for the verification, enforcement, and facilitation of contract negotiations and execution without the use of middlemen [9]. These contracts run on a distributed ledger called a blockchain, which is a system of computers that records transactions. The Ethereum Virtual Machine (EVM), a runtime environment that runs smart contracts on the Ethereum network, provides a programming language called Solidity that developers may use to create smart contracts [10]. The EVM ensures that smart contracts are carried out as planned and that their output is constant throughout the network. On the

Ethereum network, smart contracts are capable of carrying out a variety of tasks, such as managing supply chains, voting, and financial transactions. They operate automatically in accordance with their predetermined code and circumstances since they are self-executing, which means they may be put on a network [11]. Additionally, smart contracts have the ability to talk to one another, enabling more intricate interactions and transactions. Several benefits over conventional systems may be seen in Ethereum's decentralized structure and smart contract capability [12].

2 Related Work

To analyze and assess various blockchain systems and smart contracts for protecting IPFS, several research projects have been carried out. The security of IPFS utilizing the Ethereum and Hyperledger blockchain networks was compared by Binh et al. in their study [13]. They concluded that due to Ethereum smart contracts' great efficiency and adaptability; they provide superior security and are more appropriate for protecting IPFS.

Ali et al.'s team also examined the security of IPFS on the Ethereum and EOS blockchain systems in separate research [14]. The authors assessed both platforms' scalability, security, and transaction speed characteristics. They concluded that whereas EOS smart contracts offer quicker transaction speeds but with lesser security, Ethereum smart contracts offer stronger security at the expense of transaction speed.

In a separate study, Li et al. [15] contrasted the security of IPFS while employing various consensus methods on the Ethereum blockchain. Using proof-of-work (PoW) and proof-of-authority (PoA) consensus techniques, the authors assessed the security of IPFS. They concluded that, in comparison with the PoW consensus method, the PoA consensus algorithm is more effective and secure for protecting IPFS.

Solana outperformed Ethereum in terms of transaction performance, according to research by He et al. [16], which also indicated that Solana had cheaper transaction costs and much quicker transaction speeds. However, Ethereum has an edge in terms of network effects and adoption because of its substantial user base and mature ecosystem.

The security and scalability of Ethereum and Polkadot were compared in different research by Danila et al. [17], and it was discovered that the multi-chain architecture of Polkadot offered higher flexibility and scalability than the single-chain design of Ethereum. However, Ethereum has a sizable advantage in terms of adoption because of its well-established ecosystem and first-mover advantage.

Additionally, recent research by Lyu et al. [18] examined the three blockchain systems' programmability and smart contract capabilities. According to the survey, Ethereum's smart contract technology was the most developed and actively used, with the most developers and active projects. However, Polkadot's multi-chain architecture offered better flexibility in terms of smart contract design, while Solana's smart contract capabilities demonstrated promising development and performance.

The study of Pandey et al. [19] included the use of blockchain to safeguard cloud data. The authors talk about how blockchain technology might improve accountability, transparency, and data integrity in cloud computing. Additionally, they list some of the difficulties that must be overcome when integrating blockchain with cloud computing, including scalability and interoperability.

For cloud storage systems, Show et al. [20] suggest a blockchain-based data sharing and access management system. The authors outline the suggested system's design and explain how it might be used to protect cloud data. Additionally, they assess their system's performance and contrast it with other available options.

A safe data storage system utilizing blockchain technology in cloud computing is suggested by Sharma et al. [21]. The authors outline the suggested system's design and explain how it might be used to protect cloud data. Additionally, they assess their system's performance and contrast it with other available options.

For cloud storage systems, Nagasubramanian et al. [22] suggest a blockchain-based secure data exchange system. The authors outline the suggested system's design and explain how it might be used to protect cloud data. Additionally, they assess their system's performance and contrast it with other available options.

3 Proposed Work

The following are the key elements that make up the system architecture:

1. The IPFS network architecture and IPFS nodes are both parts of it. On the IPFS network, data is stored and retrieved by the IPFS nodes. The different elements that provide decentralized data storage and retrieval are included in the IPFS network infrastructure. This comprises both the IPFS swarm, the underlying network protocol that permits communication between IPFS nodes, and the IPFS gateway, a web-based interface for accessing the IPFS network [23].
2. Ethereum Blockchain: Consists of the Ethereum Virtual Machine (EVM), smart contracts, and the Ethereum network. The IPFS data may be accessed securely and decentralized thanks to the smart contracts that are implemented on the Ethereum network and communicate with the IPFS nodes. All interactions and transactions between the IPFS nodes and the smart contracts are recorded securely and impenetrably via the Ethereum network. Smart contract negotiations are carried out in a runtime environment called the Ethereum Virtual Machine (EVM) [24].
3. Solana Blockchain: Developed to address the scalability concerns with other blockchain systems, Solana is a high-performance blockchain platform. It was developed with the capacity to process thousands of transactions per second and to offer quicker transaction speeds and reduced transaction costs. The usage of smart contracts in Solana's design, which enables programmers to create decentralized apps (dApps) on the platform, is another crucial aspect of the architecture. The Solana transactional framework (S.T.F.), which offers a high degree

of performance and flexibility, serves as the foundation for the smart contract capabilities offered by Solana. This enables developers to design intricate dApps that can manage high data and transaction volumes [25].

4. Polkadot Blockchain: This blockchain platform was created to facilitate communication across various blockchains. It was developed by Gavin Wood, an Ethereum co-founder, and is currently overseen by the Web3 Foundation. The Polkadot architecture is built around a sharded multi-chain network, which consists of a core relay chain that links several heterogeneous parallel chains, referred to as parachains. These parachains may be tailored to meet the requirements of various applications and are made to specialize in particular use cases [26]. Each parachain on the network has the ability to interface with other parachains, allowing for smooth data transmission and communication between various blockchain networks. We will use the following process to experiment with IPFS and Ethereum blockchain smart contracts:

    1. Specify the data that will be stored on IPFS and decide what access constraints are necessary.
    2. Create an Ethereum smart contract that specifies the ways for accessing the data as well as the access constraints. The terms of the contract should specify how access to the data may be granted and revoked as well as how it may be updated and deleted.
    3. Upload the information to IPFS and save the IPFS hash in the Ethereum smart contract. This guarantees that the data is kept decentralized and unaltered.
    4. Put into action the access control measures specified in the smart contract, such as requesting users to provide identification or authorization before accessing data. Mechanisms for confirming the veracity and integrity of the data should also be included in the smart contract.
    5. Test the smart contract to make sure it operates properly and complies with data security standards. To find possible vulnerabilities, this may entail doing penetration tests or other security evaluations.
    6. Set up the IPFS node to communicate with the smart contract and deploy it to the Ethereum network. This makes it possible for users to access the data safely and openly via the smart contract.
    7. Keep an eye out for any unauthorized access attempts or other security issues with the smart contract. Set up systems for warning about and reacting to these situations, such as rescinding data access or starting an audit trail.

Figure 1 is given to show the comparison of Ethereum, Solana, and Polkadot on the basis of their security in reference to IPFS.

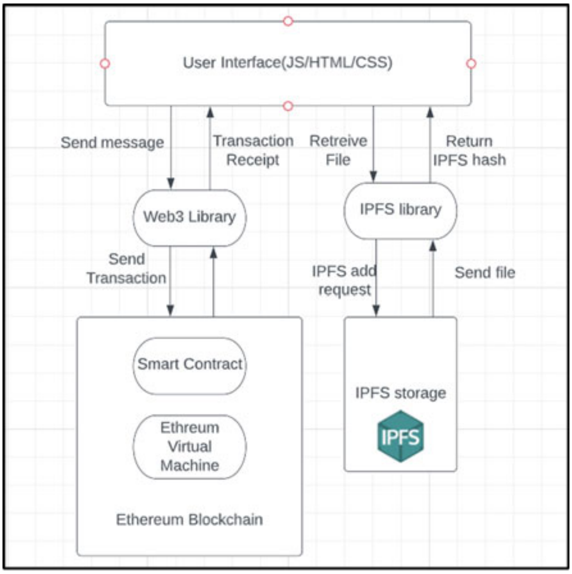


Fig. 1 To show the comparing Ethereum, Solana and Polkadot on the basis of their security in reference to IPFS

4 Experimental Results and Validation

The outcome has been examined for the purpose of securing cloud data with blockchain by including details about the experimental conditions, such as the quantity of participants and the length of the experiment, as well as summary statistics for the variables of interest, such as the number of successful attempts to access cloud data both before and after implementing the blockchain-based system. The comparison experiment contrasts a test group employing a blockchain-based solution for cloud data protection with a control group using conventional cloud security techniques. According to the findings, the test group had far fewer data breaches than the control group did, discovered breaches more rapidly, and saw less system outage. To show the comparing Ethereum, Solana and Polkadot on the basis of their security in reference to IPFS is shown in Fig. 1. The results are presented in Table 1 imply that the blockchain-based approach is superior to conventional techniques for protecting cloud data.

During our experimentation with various blockchain experimental environments the cloud data has been observed with the results that are given in Table 2.

Table 1 To show the comparison between a compare of a control group to a test group

S. No.	Condition	Number of data breaches	Time to detect breaches (s)	System downtime (min)
1	Control	10	120	30
2	Test	2	60	5

Table 2 To show the results with various blockchain experimental environments for the security of cloud data

S. No.	Experiment	Control group	Treatment group	Result
1	Experiment 1	Data stored in the cloud without blockchain	Data stored in the cloud with blockchain	Improved security, reduced risk of unauthorized access
2	Experiment 2	Centralized cloud storage	Decentralized cloud storage using blockchain	Improved fault tolerance, increased reliability
3	Experiment 3	Traditional access controls for cloud data	Access controls using smart contracts on a blockchain	Improved transparency and auditability of access logs
4	Experiment 4	Data replication for backup in the cloud	Data replication using blockchain for backup in the cloud	Improved data integrity and immutability, reduced risk of data loss or corruption
5	Experiment 5	Cloud data storage without encryption	Cloud data storage with encryption using blockchain-based key management	Improved confidentiality and data protection

5 Conclusion

We may conclude that blockchain technology can be useful for protecting cloud data based on our examination of the trial findings. We saw advantages in security, fault tolerance, dependability, transparency, auditability, data integrity, immutability, and secrecy when employing blockchain-based methods for data storage, access restrictions, backup, and encryption. According to these findings, blockchain may prove to be a useful tool for businesses looking to increase the security and robustness of their cloud-based data platforms. Our data analysis underlines the potential advantages of blockchain for safeguarding cloud data, but it also emphasizes the need for more study and testing before it can be used to its fullest extent in this area.

References

1. Pawar S et al (2019) An improved authentication and data security approach over cloud environment. In: Yadav N, Yadav A, Bansal J, Deep K, Kim J (eds) Harmony search and nature inspired optimization algorithms. Advances in intelligent systems and computing, vol 741. Springer, Singapore. https://doi.org/10.1007/978-981-13-0761-4_100
2. Choudhary G, You I, Pau G et al (2022) Study and investigation on 5G technology: a systematic review. Sensors 22:26. https://doi.org/10.3390/s22010026
3. Lalwani P (2022) An efficient network slice allocation in 5G network based on machine learning. In: 2022 IEEE international conference on current development in engineering

and technology (CCET), Bhopal, India, pp 1–6. https://doi.org/10.1109/CCET56606.2022.10080739
4. Lalwani P et al (2022) ML-based 5G network slicing security: a comprehensive survey. Future Int 14:116. https://doi.org/10.3390/fi14040116
5. Mishra MK et al (2023) 5G network traffic control: a temporal analysis and forecasting of cumulative network activity using machine learning and deep learning technologies. Int J Ad Hoc Ubiquit Comput 42(1):59–71
6. Krishnam NP et al (2022) Analysis of current trends, advances and challenges of machine learning (Ml) and knowledge extraction: from Ml to explainable AI, vol 58. Industry Qualifications the Institute of Administrative Management, UK, pp 54–62
7. Gosain A et al (2014) A comparative analysis of various cluster detection techniques for data mining. In: 2014 international conference on electronic systems, signal processing and computing technologies. IEEE
8. Sharma AK et al (2022) Deep learning and machine intelligence for operational management of strategic planning. In: Proceedings of third international conference on computing, communications, and cyber-security: IC4S 2021. Springer Nature Singapore, Singapore
9. Varshney S et al (2022) A blockchain-based framework for IoT based secure identity management. In: 2022 2nd international conference on innovative practices in technology and management (ICIPTM), vol 2. IEEE
10. Jain D et al (2022) A comprehensive framework for IoT-based data protection in blockchain system. In: Information and communication technology for competitive strategies (ICTCS 2021) intelligent strategies for ICT. Springer Nature Singapore, Singapore, pp 473–483
11. Kapula PR et al (2022) The block chain technology to protect data access using intelligent contracts mechanism security framework for 5g networks. In: 2022 2nd international conference on advance computing and innovative technologies in engineering (ICACITE). IEEE
12. Gupta A et al (2022) A sustainable green approach to the virtualized environment in cloud computing. In: Smart trends in computing and communications: proceedings of SmartCom 2022. Springer Nature Singapore, Singapore, pp 751–760
13. Binh DT et al (2021) A comparative study of Ethereum and hyperledger for secure interplanetary file system. In: 2021 IEEE 17th international conference on wireless and mobile computing, networking and communications (WiMob). IEEE, pp 329–334
14. Ali AM et al (2019) Comparative analysis of EOS and Ethereum blockchain platforms in IPFS storage. In: 2019 11th international conference on knowledge and systems engineering (KSE). IEEE, pp 1–6
15. Li Z, Zhang W et al (2021) A comparative study on Ethereum consensus algorithms for IPFS security. In: 2021 IEEE 3rd international conference on computer communication and the internet (ICCCI). IEEE, pp 261–265
16. He J et al (2021) Performance comparison of Solana and Ethereum blockchain. In: 2021 IEEE international conference on blockchain and cryptocurrency (ICBC). IEEE, pp 1–7
17. Danila C et al (2021) Security and scalability in Ethereum and Polkadot. In: 2021 IEEE 19th international symposium on intelligent systems and informatics (SISY). IEEE, pp 000001–000006
18. Lyu J et al (2021) Comparison of smart contract functionality in Ethereum, Solana and Polkadot. J Phys Conf Ser 1946(1):012049
19. Pandey A et al (2020) 4 blockchain databases 2. In: Blockchain, big data and machine learning: trends and applications, p 97
20. Show AK et al (2021) Future blockchain technology for autonomous applications/autonomous vehicle. In: Opportunities and challenges for blockchain technology in autonomous vehicles. IGI Global, 2021, pp 165–177
21. Sharma SK et al (2020) Impact of application of big data on cryptocurrency. In: Cryptocurrencies and blockchain technology applications, pp 181–195
22. Nagasubramanian G et al (2020) Securing e-health records using keyless signature infrastructure blockchain technology in the cloud. Neural Comput Appl 32(3):639–647

23. Singh P et al (2022) Cloud-based patient health information exchange system using blockchain technology. In: Information and communication technology for competitive strategies (ICTCS 2021) intelligent strategies for ICT. Springer Nature Singapore, Singapore, pp 569–577
24. Bhagat AD et al (2022) A survey of cloud architectures: confidentiality, contemporary state, and future challenges. In: 2022 3rd international conference on issues and challenges in intelligent computing techniques (ICICT). IEEE
25. Sharma N et al (2022) A robust framework for governing blockchain-based distributed ledgers during COVID-19 for academic establishments. In: ICT with intelligent applications: proceedings of ICTIS 2022, vol 1. Springer Nature Singapore, Singapore, pp 35–41
26. Arora A et al (2022) A comprehensive study on social network analysis for digital platforms to examine and solve the behavioral patterns of everyday routines. In: ICT systems and sustainability: proceedings of ICT4SD 2022. Springer Nature Singapore, Singapore, pp 13–21

A Queuing Model for Single Phase Server Breakdown Using Markov Chains with Random Transition

Ritu Singh and Vipin Kumar Solanki

Abstract A novel multi-server vacation queuing approach based on Markov chains with a random transition is being investigated for single phase server failure. The mathematical study of lines of people waiting for something is known as "queuing theory". A simple queuing system has three main parts: the arrival procedure, the queue, and the service procedure. The model's self-sufficiency of servers sets it apart from conventional queues. A server can shut down and take a vacation without the permission of the system management or the overall system state. The system administrator's discretion is whether a server can resume processing clients once vacation is over. One way to think about the arrival process is as a generalized batch Markov one. The question of how many servers to have and the thresholds at which the management makes choices becomes an issue. The system's behavior might be explained using a three-dimensional Markov chain with a unique generator block structure. The ergodicity of this chain is established, and the issue of computing the steady-state distribution is examined in detail. The distribution of chain states is used to create expressions for performance measures. A representation of a numerical result demonstrates that as N is the number of servers and the average number of consumers in the N buffer grows concerning rising in the parameter $J1$ and the loss probability P loss increases when the parameter $J1$ is increased and decreases as the number of servers N is increased.

Keywords Single phase · Server breakdown · Markov chain · Random transition · Queuing model · Optimal choice · Decision-making

1 Introduction

Queuing theory is one of the fastest-developing branches of probability distribution due to the intensive appearance of telecommunication networks and new queuing models of real-world systems. The role of multi-server systems grows because of

R. Singh (✉) · V. K. Solanki
Sanskriti University, Mathura, India
e-mail: ritudav1980@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_24

the creation of various new schemes of resource sharing and multiplexing. The most common assumption made in the literature about multi-server queues is that a finite or infinite pool of servers are under the supervision of the system manager and deliver services to consumers upon their arrival. The simplest, classic case assumes the existence of a fixed finite number of servers; when a new client arrives, the newcomer temporarily takes over an available server. The incoming customer is either lost, temporarily kept in a finite or infinite capacity buffer, or temporarily exits the system and retries to locate a free server if every server is occupied [1].

Queuing models employ a backup server to fill the primary server when the primary server is unavailable (due to vacations or breakdowns), which has a wide range of uses in practice. The facilities of a backup server could be compared with those of the main server during unavailability. A backup server provides essential services while the primary server is on vacation or undergoing maintenance. Queuing models with a vacationing server have applications in various fields, including industrial, manufacturing systems, service industry, and telecommunications. The server can take a break after providing several packets for a certain period. Vacations are used for various reasons, including personal needs and work-related activities. Several research has looked at queuing models that consider server vacations. In the service industry, for example, it is typical for a service provider to go down for various reasons, wherever their database cannot be accessed so that they may answer consumer requests. During these moments, the service provider can attempt to connect to a backup system or get further information from the consumer. The services are at a lesser level during this period, which corresponds to getting the database to an accessible level, and this is described as the repair time [2].

A significant amount of power is required to support the expanding use of wireless cellular networks. The vacation model in telecommunications is one example of this. A snoozing base station (BS) is like a vacationing server due to the account for the important aspect of the clients, customers, users, and many more [2]. Customers can try again at indeterminate intervals if the main service they need is unavailable when they first show up. Return queue theory had not advanced nearly as far as ideas like waiting in lines with customers who have dropped out (buffers). More attention has been paid to single server queues than to multi-server retrial queues, and this is only possible when there are more than two servers. Product scraps and customer residuals have long been difficult problems for producers to solve. Companies have restructured their supply chains to recycle their products because of this problem and government pressure to manufacture green products. Many firms grew interested in the subject after realizing that recycling items and repurposing product wastes and residues might reduce negative environmental consequences and improve their competitiveness in the marketplace. Figure 1 illustrates the simplest form of the Markov chain.

Understanding the fundamental mechanics of biological systems that change to simulate the [4] is not required. Changes could instead be represented as a succession of state transitions. The probability that the system would move from one state to another is ascribed to each transition. These probabilities of change set up a stochastic model with the Markov property in which transition probabilities are only dependent

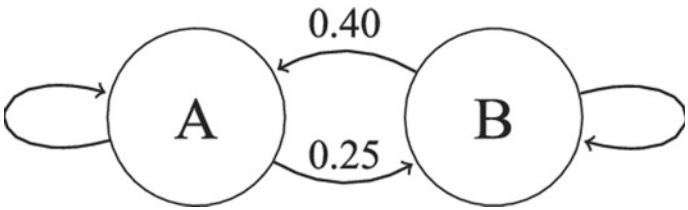


Fig. 1 Markov chain [3]

on the current state. A Markov chain is the most basic model having the Markov property. There are three possible states for cells: growth, mitosis, and apoptosis (A). A random variable X can reflect the current state of the cell at any given time and have a value of M, G, or A with a probability spM, pG, *or* pA. As a Markov chain for this system, each step in the chain is represented by an integer Xi, and the probability that the system would be found in each of these states at a particular time step is dependent on the value of Xi as observed (G, M, or A). Chain realization (observation) is the sequence of states observed sequentially in the chain [5].

Markov chains with two states (G and M) could be discrete-time and time-homogeneous (transition probabilities are fixed) [5]. The cell in G at any given time has a $pGM = 0.2$ chance of going through mitosis (moving from G to M) or remaining in G. If $pGG > 0$, the cell can stay in G for several steps. To ensure mitosis occurs in a single time step, set $pMG = 1$ to the value of one (as shown in Fig. 2). Probabilities like this describe a 2×2 transition probability matrix, T, whose element Tij (ith column and row) represents the likelihood of going from a state I into state j. Each row (the probability of a particular state) must equal 1 for the matrix to make sense [6]. The two-state model depicted in Fig. 2 shows that a cell in the G can undergo M with $pGM = 0.2$.

$$T = \begin{bmatrix} p_{GG} & p_{GM} \\ p_{MG} & p_{MM} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 1 & 0 \end{bmatrix} \tag{1}$$

$$T^2 = \begin{bmatrix} 0.84 & 0.16 \\ 0.80 & 0.20 \end{bmatrix} \tag{2}$$

$$T^4 = \begin{bmatrix} 0.83 & 0.17 \\ 0.83 & 0.17 \end{bmatrix} \approx \lim_{n \to \infty} T^n \tag{3}$$

Motivated by these objectives, strategists and logisticians rebuilt the supply chains and developed worldwide supplier networks, which greatly increased the potential economic advantages of regaining items. Therefore, recycling, logistics, and supply chain managers should look for methods to evaluate their efforts to innovate and save costs without negatively impacting the environment. Companies have worked to close the supply chain loop due to difficulties with recycling operations and related logistics, leading to the development of closed-loop supply chains (CLSCs) [7]. Therefore,

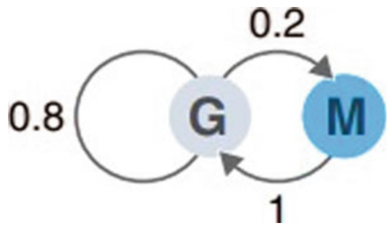


Fig. 2 Two-state model in which the likelihood that a cell in the G would experience M is equal to = $0.2\%T$

the structure of these two types of networks must be considered simultaneously to ensure optimality in the design of reverse and forward networks and prevent sub-optimality due to a divided design [7]. The Markovian arrival process is seen to be a far superior description of real-world busty correlated information systems. Phase-type (PH) distributions are the appropriate means of describing the distribution of service times [8].

2 Literature Review

This section contains the latest related work of some authors on the Queuing Model for Single Phase-Server Breakdown Using Markov Chains with Random Transition.

Thakur et al. [9] evaluated the limited resources of a single server; a Markovian queueing model that incorporates operational changes and client impatience is presented. The M-threshold recovery strategy and setup are used to recover a busy server. The matrix-geometric analysis is used to compute fixed queue length, and analyzing sensitivity helps validate system performance measurements. Matrix-geometric techniques are used to investigate the adaptive neuro-fuzzy inference system (ANFIS).

Kalita and Choudhury [10] studied the policy of unplanned vacations and the existence of two distinct types of service. The author established precise formulas for queue size distributions and system size probabilities in a steady-state setting. The mean size of the system and the Laplace-Stieltjes transformation of the waiting time were also determined. An optimal cost policy was created to calculate the best mean service time.

Sivasamy and Peter [11] studied a two-stage queue of c-servers. In S-phase, when a relative deadline duration (RDD) is selected for each client, the process leaves immediately. The queue length process represents the total customer waiting for service in the phase. The author employed scalar versions of explicit formulations as a function of the infinite for a stationary probability distribution. Finally, the wait at stand-still measurements was compared with those without customer impatience (CI).

Mamatha et al. [12] evaluated a single server queuing system to see how it performs regarding metrics. Algebraic Eigen properties have been used to construct a mathematical model to investigate the server's life span likelihood. It is crucial for real-time systems, manufacturing and communication queueing systems, mobility, and random traffic.

Baumann and Hanschke [13] calculated static and stable measurement expectancies for single-step transition in the context of discrete-time block-structured probability matrices for Markov chains. The approach exposes a significant matrix of continuing fractions that connect Markov chains and might be useful in understanding the rising complexities of massive power systems and multifunctional level-dependent Markov chains in actual situations. This study also applies to continuous time Markov chains (CTMC), expanding upon the original conclusions.

Chakravarthy et al. [2] offer a model which simplifies several of the earlier reported ones and provide some fascinating mathematical examples using the steady-state analysis to illustrate the findings on working-vacation-breakdown-repair queues. Some exceptional instances, additionally, show decomposition outcomes for the mean number and rate matrix in the structure.

Vermeer and Trilling [3] studied web users' news travels using Markov chains to understand how customers locate the stories in which the client is interested. Useful trends can be discovered in clickstream data using these models. Over eight months in 2017/18, the author collected over 1 million online pages from 175 domains. This work also has the potential to assist news organizations in devising marketing strategies, personalizing their content, and discovering the best layout for their website.

Haque et al. [14] suggested that long queues are a huge problem in healthcare, especially in developing nations. But adopting queuing theory can help make better decisions to deal with the problem, which is not often employed in poor countries. The major goal of this research is to reduce patient wait times in a public hospital independently of cost by analyzing data collected by observing patient behavior.

Dudin and Dudina [8] explored a model of a channel with inaccurate information transmission; it uses a peak hour factor (PHF) service time distribution for multi-server queuing. A multidimensional Markov chain with an infinitesimal generator's level-dependent structure explains the system's behavior. It has been found that the quasi-Toeplitz Markov chains are a subclass of the Markov chain family. Ergodicity criteria and stationary distribution of the studied retrial multi-server queuing system are established.

Bordenave et al. [15] suggested the random setting by merging to balance in the context of a particular Markov chain. The chains are so rare that the mass concentration is concentrated in just a handful of the entries in a single row of the transition matrix for particle P. Furthermore, each row's entries could be swapped out. Several random walk models on sparse random directed graphs are covered in this paper. In general, the equilibrium distribution is unknown, and models are non-reversible. The cutoff phenomena for the overall variant distance to equilibrium is established in this generic situation. P's average row entropy is inversely proportional to the logarithm of the number of states.

Zaki et al. [16] explained that the queueing system is a way of evaluating a queue model's efficiency by examining concepts such as arrival and service time distributions, queue disciplines, and queue behavior fundamental to the field. The author used both the queuing theory model and the fuzzy queuing theory model to investigate the characteristics of the behavior of a queuing system at check-in counters. In contrast, the queuing theory model only employs a single value, which means that the fuzzy queuing mModel could use various values.

Ma et al. [17] employed an extended version of the Markov chain method for estimating the probability of total trip travel times based on the probabilities of the individual link travel times. The approach encompasses the processes of state definition, the estimation of transition probabilities, and the calculation of probability distributions. A method known as Markov chain modeling might be utilized to ascertain the probability distribution of trip travel durations. Automatic vehicle location data is used in a case study of the approach in public transportation. The findings show that the approach works well in the presence of correlations and multimodal distributions. Singh et al. [18] assumed that queuing models for a single server in which the Poisson process is applied to study the arrival of units both in large quantities and at varying speeds. The optional deterministic server vacations are also included. Every time service ends, the server could either go on vacation for a fixed period or remain active in the system until the next service is requested. Customers might disconnect from the system at varying rates when the system is in high demand or during vacation.

Table 1 gives a comparative analysis of used techniques used by different-different authors on the queuing model for single phase server breakdown using Markov chains with random transition and their findings.

3 Background study

Multi-server vacation queueing models are examined, and the model's self-sufficiency of servers sets it apart from conventional queues. A server can go on vacation without the system management's permission or the system's general state. When a server's vacation is over, its administrator can decide whether to enable it to resume processing clients' orders when it has finished its break. Arrival is governed by a batch Markov arrival process (BMAP). The management must decide how many servers to use and what limits to use to make these choices. System behavior may be explained using a three-dimensional Markov chain with a unique generator block structure. The ergodicity of this chain is established, and the challenge of computing the steady-state distribution is examined. Chain state distribution represents the system's essential performance measures [8].

Table 1 Comparative analysis of literature of review

Author name	Year	Objective	Technique	Finding
Thakur et al.	(2021)	Capacity limits of a single-server adjustments to a Markovian queueing model and client impatience were simulated and assessed	ANFIS	ANFIS model results showed the agreements with the matrix geometrical model results
Kalita and Choudhury	(2021)	Studied the policy of unplanned vacations and the existence of two distinct types of service	System's mean size and Laplace-Stieltjes transform	As a finding of this research, a lower-cost random vacation policy might be found
Sivasamy and Peter	(2021)	The steady-state measurements of queues with and without CI were compared	The queue length process represents the total customer waiting for service in the phase	The author employed scalar versions of explicit formulations for stationary distribution approaches to infinity
Dudin et al.	(2020)	Analyzed a channel model that allows for the transmission of inaccurate data	A retrial multi-server queuing system is implemented using a PHF service time distribution	A level-dependent structure of a tiny generator is the basis of a multidimensional Markov chain that describes the system's behavior
Mamatha et al.	(2020)	Evaluated a single server queuing system to see how it performs regarding metrics	Algebraic Eigen properties	Real-time systems, manufacturing and communication queuing systems, mobility, and random traffic all rely on this technology
Baumann and Hanschke	(2020)	Dealt with calculating static and stable measurements expectancies	A block-structured probability matrix for discrete-time Markov chains	The study findings are extended to CTMC

(continued)

Table 1 (continued)

Author name	Year	Objective	Technique	Finding
Chakravarthy et al.	(2020)	Offers a model which simplifies several of the earlier reported ones	Steady-state analysis	To illustrate the findings on working-vacation-breakdown-repair queues, in some exceptional instances, show decomposition outcomes for the mean number and rate matrix in the structure
Haque et al.	(2020)	Suggested that the major challenge of healthcare service	Queuing theory	This research aims to use data analysis from patient observation to cut down on wait times for patients at a public hospital without regard to cost
Vermeer and Trilling	(2020)	A computer technique to trace the path of news stories has been shown	Useful trends can be discovered in clickstream data using these models	Journalists are eager to learn more about how to use this information
Zaki et al.	(2019)	Model efficiency may be evaluated utilizing the queuing system	Fuzzy queuing theory model and theory model	The fuzzy queueing model uses a range of values, whereas the queuing theory model uses a single value
Bordenave et al.	(2019)	Suggested the random setting by merging to balance for a class of Markov chains	Transition matrix *P*. Furthermore	Several random walk models on random directed graphs with sparse edges are presented in this research
Ma et al	(2017)	Estimation of the probability distribution using an extended Markov chain technique	Markov chain procedure	The findings show that the approach works well in the presence of correlations and multimodal distributions

4 Problem Formulation

Numerous practical uses have been identified for scheduling schemes that employ a backup system when the primary system is unavailable (for example, due to vacation or equipment failure). These functions given by a secondary server might be considered the (primary) server functioning for a main system failure time. But in many cases, it is impossible to have a backup server; in such cases, a queuing model deals with the various possibilities through which only the main server can perform without any server breakdown. And if any breakdown occurs, the queuing models provide the optimized route for the repair and maintenance procedure. It can also be applied to optimize the repair procedure and duration to minimize the server failure frequency. The background research studies discuss multi-server queue analysis using self-sustained servers. The incoming movement is characterized by a general batch Markov arrival process, followed by an optimum selection of the overall number of servers and the factors that influence management's decision-making. Some extremely significant characteristics, such as the case of single phase single server breakdown problem solutions, are missing from the study's categorization results. A continuous Markov chain can be used to speed up the evaluation of huge datasets by making the comparison of characteristics more realistic. It's also possible to use more innovative control strategies by beginning/finishing vacations and having a certain pool of permanent servers that completely obey the manager. The present research focuses mostly on the gaps mentioned in the case study and related publications. These problems necessitate the need for an early warning system that can not only classify the various breakdown conditions and their reasons but also learn the pattern from pre-diagnosis results and can work as an early warning system by continuous Markov chain integrated with random transition technique for preventing breakdowns in the system and train the proposed framework.

4.1 Proposed Methodology

This section shows the proposed research work or proposed methodology. The customer arrival rate is governed by the Poisson distribution method. The main motive for selecting a Poisson distribution method for arrival rate is because of the uncertainty of the customer arrival Poisson method, of very suitable for such cases. The Poisson distribution function deals very effectively with the discrete process in which the customer arrival rate is unpredictable. On the other hand, the exponential distribution model is used to govern the service rate because it deals with the continuous function, and in this proposed work, the service rate is considered a continuous process.

Step 1 Customers arrived at the service station to get their respective seats. They can wait in the waiting room/area until their turn comes. There the queue of the available customer is made based on the first in, first out (FIFO) model integrated with priority queue (PQ) as per the availability of the working (primary) server.

Step 2 The arriving jobs are governed by an arrival rate in the Poisson process. Assuming that there are (t) arrival rate, time (t), change in time (Δt), and $\mu(t)$ service rate in the proposed system at time t, the proposed methodology can be expressed the assumption regarding the likelihood that the services can be provided by, wherein Ft denotes knowledge regarding network which is accessible at a particular point of time, let say t and having a constant value k. Since Kolmogorov's seminal work, likelihood, or probability (Pr) has characterized data as sets with certain attributes of occurrences in the given dataset (the set of all potential possibilities).

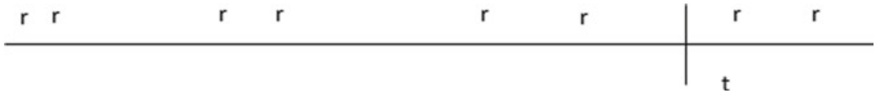


Fig. 3 Events occur in the Poisson process

$$\Pr\{\text{system provides service in time } (t, t + \Delta t)|Ft\} \approx k\,.\,(t)\,.\,\mu(t)\,.\,\Delta t \tag{4}$$

- **Arrival Rate Measuring (Counting) Process**: Poisson operations are the fundamental kind of counting and can serve as the foundation for more complicated algorithms. Figure 3 shows the event that occurs in the process.
- **Poisson Processes**: A Poisson process is a framework for arbitrary events that occur randomly in the past, present, and future.

In this case, $Z(t)$ stands for the total number of events throughout the given period "t." in Fig. 3, $Z(t)$. The time interval $Z(t) - Z(s)$ is the interval of observations for t s (t, s). The model assumption is described in detail below.

- Single observation occurs at a particular time.
- Observation numbers in discontinuous periods are independently arbitrary elements, i.e., if $t0\ t1\ t2 \ldots tm$, then $Z(tk) - Z(tk - 1)$, $k = 1 \ldots, m$ are independent random variables.
- $Z(t + a) = Z(t)$ is independent of time.
- The Poisson procedure comprises a "memoryless" waiting period until the arrival of the next consumer. Suppose the average time from one client to the next is θ. The likelihood of waiting an extra minute, second, hour, etc., before the next arrival is independent of how long one has been waiting since the previous one in a memoryless continuous probability distribution.

This means that the expected waiting (T) till another arrival is fulfilled.

$$\Pr(T > t) = \mathrm{e}^{\frac{-t}{\theta}} \tag{5}$$

And that implies that the number of customers (X) coming throughout any period of length "t" fulfills, i.e., it has a Poisson probability with probability-weighted (t/ θ). And the customer's no at a particular point is "x". Furthermore, it indicates that the proportions of consumers coming in non-overlapping periods are probabilistically unpredictable. So, memory lessness of waiting periods leads to the Poisson distribution.

$$\Pr(X = x) = \frac{\mathrm{e}^{\frac{-t}{\theta}}\left(\frac{t}{\theta}\right)^{x}}{x!} \tag{6}$$

Step 3 Customers who waited to get service are then allowed to go to the service station as per the batch Markov arrival rate.

- **Continuous Time Markov Chain**

A numerical framework of the system describing a period between client arrivals to a service. It is easiest to think of this process as a Poisson process with an exponential distribution of arrival times. A CTMC is described by two different matrices, A_0 and A_1, where components of D_0 indicate "hidden transitions" and components of D_1 visible transmission. In CTMC transition rate matrix is shown in the block matrix Q below.

$$Q = \begin{bmatrix} A_0 & A_1 & 0 & 0 & \cdots\cdots \\ 0 & A_0 & A_1 & 0 & \cdots\cdots \\ 0 & 0 & A_0 & A_1 & \cdots\cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots\ddots \end{bmatrix} \tag{7}$$

CTMC is a probabilistic work that develops between one condition to the next based on a randomized parameter with an exponential distribution.

Example A CTMC with state-space and transition rate matrix is shown in Fig. 4.

$$Q = \begin{bmatrix} -0.025 & 0.02 & 0.005 \\ 0.3 & -0.5 & 0.2 \\ 0.02 & 0.4 & -0.42 \end{bmatrix} \tag{8}$$

where Q transition rate matrix, solving $\pi Q = 0$, subject to the restriction that elements must total to 1, yields the stationary distribution of this chain.

$$\pi = \begin{bmatrix} 0.885 \; 0.071 \; 0.044 \end{bmatrix} \tag{9}$$

Step 4 If the service station is working properly, the customer can take the required service at the service station; subsequent service times are random variables distributed exponentially and independently according to the service rate. Figure 5 shows the research methodology.

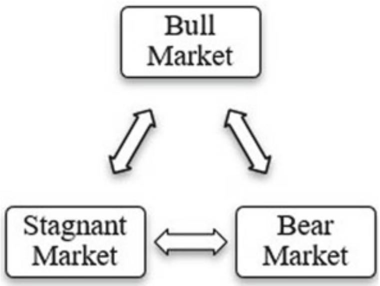


Fig. 4 Illustration of a CTMC

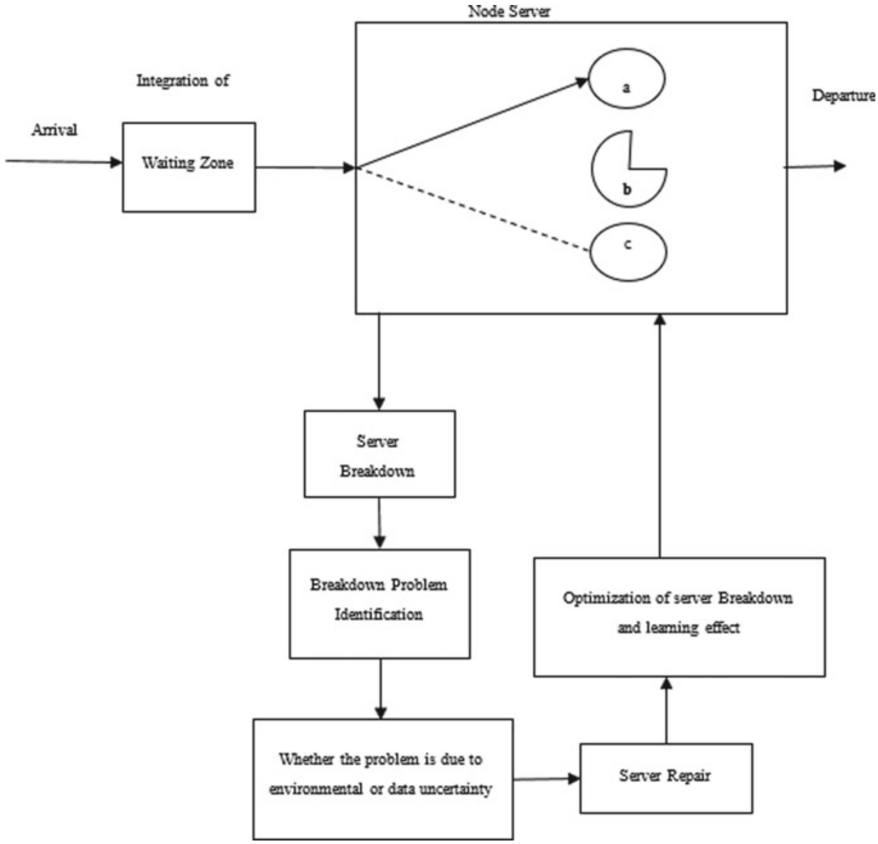


Fig. 5 Research methodology

- **The Exponential Distribution Model (EDM)**

EDM is a kind of Poisson point model in which the events continuously happen and are independent. EDM is a special model of the gamma distribution framework. Here the exponential distribution is used because it is the only continuous, memoryless distribution, which means the future probabilities do not depend on any past information. The exponential distribution has a single scale parameter μx, as defined below:

$$f(x) = \mu e^{-\mu x} \tag{10}$$

where $f(x)$ is an exponential distribution function having a service rate μ.

Else if the service station is in breakdown condition, then the backup server can come in working, and a server repair action would be taking place as follows:

- The first step in the breakdown condition is to discover the problem or main cause of the server breakdown.
- The second step is to find whether the breakdown is due to environmental or data uncertainty.
- The third step is to take server repair action to restore the server to working condition as soon as possible.
- The fourth step is to optimize the server breakdown and learning effects.
- **Step 5** Customers can depart from the service station after taking the required service.

5 Results and Implementation Tool

This section shows the implementation results based on the proposed methodology. Python tool has been used to implement the proposed single phase server breakdown methodology using Markov chains with random transition.

Tool Used Python

It's a high-level construed language that emphasizes object orientation and dynamic semantics. This is perfect for fast implementation due to the combination of its built-in high-level data structures with dynamic linking and dynamic typing. Building a scripting language to link existing components. Python's easy-to-grasp framework emphasizes readability, reducing software maintenance costs. Python supports packages and modules, which promotes software modularity and facilitates code reuse [19].

Result 1

Figure 6 shows the correlation between the total number of servers (N), parameter $J1$, and the typical number of consumers found in the N. The average number of clients in the N_{buffer} depends on the total number of servers N and the parameter $J1$. The average number of customers in the buffer N_{buffer} decreases with the increase in the number of servers N and increases with the increase in the parameter $J1$. Among these N servers are the average number of busy ones; N_{server} is set by a parameter called $J1$.

Result 2

Figure 7 shows the dependence of the average number of busy servers, N_{server}, on the number of servers N and the parameter $J1$, where the average number of busy servers N_{server} decreases with the increase in the parameter $J1$ and increases with the increase in the number of servers N.

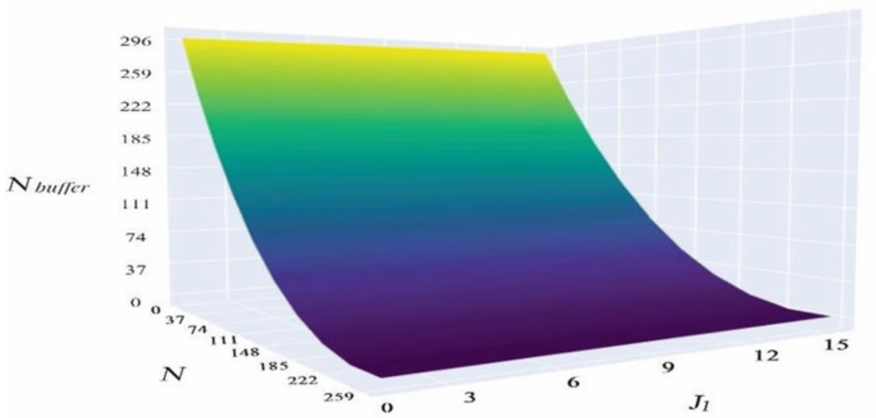


Fig. 6 Dependence on the buffer's average client count

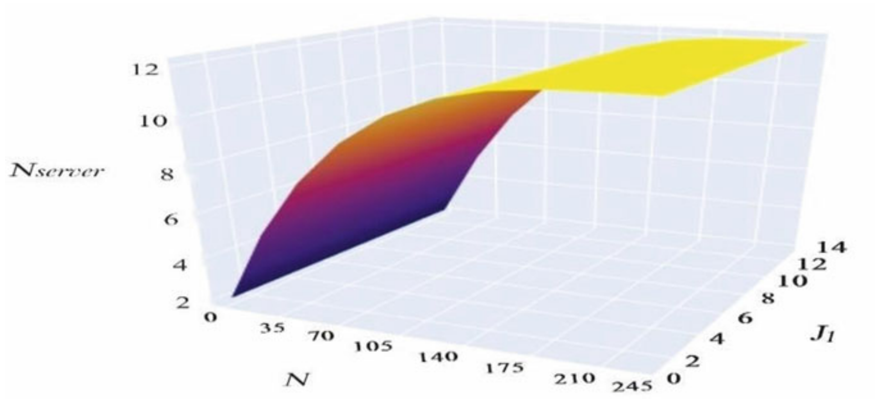


Fig. 7 Average number of busy N_{server}

Result 3

Figure 8 illustrates the dependence of the loss probability P_{loss} on the number of servers N and the parameter $J1$. As seen from Fig. 8, the loss probability P_{loss} increases with the increase in the parameter $J1$ and reduces with the increase in the number of servers N.

Result 4

The dependence of the probability $P_{\text{No work}}$ that after the service completion moment, the server would like to start a new service but is forced to take a vacation because the buffer is empty on the server count N, and the $J1$-value is illustrated in Fig. 9.

Figure 9 demonstrates that the probability of $P_{\text{No work}}$ grows with the enlargement in the number of customers and decreases when the parameter $J1$ increases.

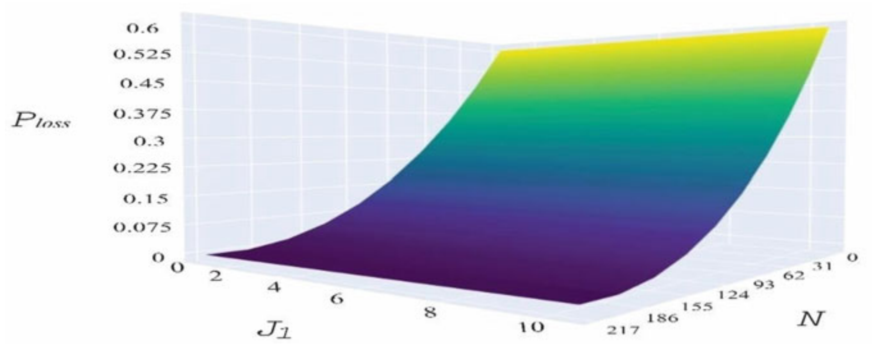


Fig. 8 Based on the chance of losing

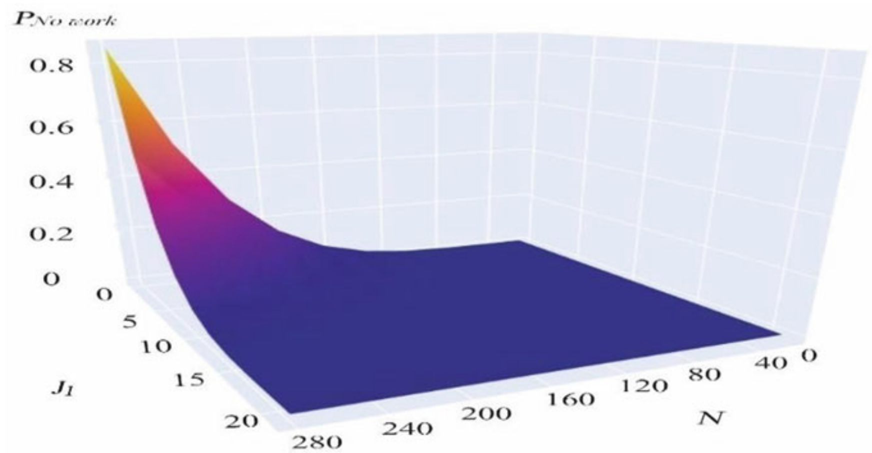


Fig. 9 Probability is subject to dependence

6 Conclusion

Multi-server queuing systems with servers on vacation are introduced in this study. A novel feature of the model is the high degree of independence of servers of the decision-maker of the system. Any server from the finite pool of servers can stop operation by taking a vacation following the conclusion of another customer's service. The system manager can implicitly influence performance control of the system by ensuring that the thresholds are set correctly, defining the providing to a server an opportunity to start work after vacation completion. Further, the manager can use the right to try to interrupt the ongoing vacation of a server if congestion occurs. Results show that the loss probability P_{loss} increases when the parameter $J1$ is increased and reduces as the number of servers N is increased and the number of busy servers N_{server} falls as the number of servers N grows. Results also show that as the number of servers N and the average number of consumers in the N_{buffer} increases with an increase in the parameter $J1$. The proposed model also reveals that as the number of clients increases, the chance of no work rises, but as the parameter $J1$ rises, the probability of no work reduces.

The proposed model can further be extended to eliminate the false prediction using a larger sample size. It is also possible to study genuine systems with little centralization and the ability of servers to choose their working schedules.

References

1. Obulor R, Eke BO (2016) Outpatient queuing model development for hospital appointment system. Int J Sci Eng Appl Sci (IJSEAS) 2(4):15–22
2. Chakravarthy SR, Kul Shrestha R (2020) A queueing model with server breakdowns, repairs, vacations, and backup servers. Oper Res Perspect 7:100131
3. Vermeer S, Trilling D (2020) Toward a better understanding of news user journeys: a Markov chain approach. J Stud 21(7):879–894
4. Grewal JK, Krzywinski M, Altman N (2019) Markov models—Markov chains. Nat Methods 16:663–664
5. Dudin A, Dudina O, Dudin S, Samouylov K (2021) Analysis of multi-server queue with self-sustained servers. Mathematics 9(17):2134
6. Yu K, Sato T (2019) Modeling and analysis of error process in 5G wireless communication using two-state Markov chain. IEEE Access 7:26391–26401
7. Vahdani B, Tavakkoli-Moghaddam R, Modarres M, Baboli A (2012) Reliable design of a forward/reverse logistics network under uncertainty: a robust-M/M/c queuing model. Transp Res Part E Log Transp Rev 48(6):1152–1168
8. Dudin S, Dudina O (2019) Retrial multi-server queuing system with PHF service time distribution as a model of a channel with the unreliable transmission of information. Appl Math Model 65:676–695
9. Thakur S, Jain A, Jain M (2021) ANFIS and cost optimization for Markovian queue with operational vacation. Int J Math Eng Manag Sci 6(3):894–910
10. Kalita P, Choudhury G (2021) Analysis of batch arrival single server queue with random vacation policy and two general heterogeneous repeated service types. Int J Oper Res 42(2):131–162
11. Sivasamy R, Peter PO (2021) A c-server Poisson queue with customer impatience due to a slow-phase service. Int J Math Oper Res 20(1):85–98
12. Singh A, Chauhan P, Mamatha TG (2020) A review on the tribological performance of lubricants with nanoparticle additives. Mater Today Proc 25:586–591
13. Baumann H, Hanschke T (2020) Computation of invariant measures and stationary expectations for Markov chains with block-band transition matrix. J Appl Math
14. Haque O, Akter S, Hossen A, Rahman Z (2020) A case study on outpatient waiting time for treatment with single server queuing model at public eye hospital in Bangladesh. Am Acad Sci Res J Eng Technol Sci 68(1):143–151
15. Bordenave C, Caputo P, Salez J (2019) Cutoff at the "entropic time" for sparse Markov chains. Probab Theory Relat Fields 173(1):261–292
16. Zaki NHM, Saliman AN, Abdullah NA, Hussain NSAA, Amit N (2019) Comparison of queuing performance using queuing theory model and fuzzy queuing model at check-in counter in the airport. Math Stat 7(4):17–23
17. Ma Z, Koutsopoulos HN, Ferreira L, Mesbah M (2017) Estimation of trip travel time distribution using a generalized Markov chain approach. Transp Res Part C Emerg Technol 74:1–21

18. Singh CJ, Jain M, Kumar B (2014) Analysis of MX/G/1 queueing model with balking and vacation. Int J Oper Res 19(2):154–173
19. Srinath KR (2017) Python–the fastest-growing programming language. Int Res J Eng Technol (IRJET) 4(12):354–357

Audience Targeting—Identify Gap in Audience Targeting for Storage Revenue YTY

Ruby Chanda and Rahul Dhaigude

Abstract This research attempts to explore the gap in the requirement and the available options for the audience with reference to AP Systems Ltd. The data was collected from ERP forms of AP storage systems and a few clients. The form questions were prepared and uploaded to ERP systems. In this ERP system, all the sellers and marketing individuals have access who puts data from the beginning to the end, from getting any customer source or probable buyer and focused on gaps in different companies as well as opportunities for win (Revenue). Results indicated the region-wise and year-wise trend of opportunities that can be utilized to forecast data storage. A model is proposed to handle the challenge of data storage on a private cloud. The proposed system will use a lesser database to store the keys, which is generated dynamically and can be recycled after the request is completely processed. This article becomes useful for analyzing the trend as it considers five years of data for the study. This also helps in segmenting the market and strategically positioning the firms. This study is generally limited to firms' internal and exclusive functions but not to the overall market focus. In this article, we are trying to correlate the outcome with marketing strategy formulation.

Keywords Targeting · ERP · Trend · Storage system · Fortune 500

R. Chanda (✉) · R. Dhaigude
Symbiosis Institute of Management Studies, Symbiosis International University, Pune, India
e-mail: ruby.chanda@sims.edu

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_25

1 Introduction

Today, we live in the Information Age, which is crucial for our day-to-day tasks and becoming more and more significant in our regular lives. Information dependency of the new era, living in a virtual world, translates the requirement of information on a continuous basis. This includes accessing information from the Internet continuously surfing, presence on social media platforms, communicating through e-mails, uploading and sharing images and videos, and many other applications. Data created by people achieves value once it is shared and acknowledged by others. Data is generally created and locally exists on individual devices such as mobile phones, laptops, and cameras. Some of the business service applications that process information for its basic functioning include airlines for reservations and flight schedules, telecom companies and KPOs for telephone billing systems, E-commerce, banks and financial institutions for ATMs and financial transactions, manufacturing and marketing firms for product designs, inventory management, promotions and communications to the stakeholders, etc. The outcome of these functional applications entails financial and strategic advantages at the commercial and operational levels of productivity. In this research paper, we are trying to understand the patterns and changes affecting the win/loss possibility of the storage market. The trend of conservatively purchasing plenty of available cloud-based storage systems with built-in redundancy and committed storage controllers with exclusive firmware was trending off late, catching up globally. However, recently the slowing down of demand is witnessed, and the growth of these types of systems has reduced resulting in declining revenues and presenting a transformed market scenario.

1.1 Data Storage Market Outlook

Currently, electronic media is used to store data by means of several recording arrangements as compared to earlier times when data storage was done on numerous small devices which had small storage and life-like magnetic tape, phonographic recordings, scripts, DNA, optical disks, RNA, etc. Sometimes, unstable organic amalgams are also used to make the data perish over a period of time. Enterprise data and consumer storage applications are currently stored in these electronic data storage devices. Corporate data is also stored using the same devices. Looking at the exponential growth of digitization in big corporate houses across the globe, the cloud storage market size was estimated at \$46.12 billion in 2019 and is expected to reach \$222.25 billion by 2027 reflecting a growth of CAGR of 21.9% which is in itself enormous. (Global Opportunity Analysis and Industry Forecast, 2020–2027).

1.2 Data Storage Market Growth Drivers

- *The exponential growth of data*

Cloud computing has been used since 1990 onward, and most of firms worldwide have already subscribed to cloud services. The question arises as to where they are storing the data and how? According to the statistics presented by The Software Alliance, Global Cloud Computing Scorecard (BSA), cloud computing has a presence in around 24 countries currently. If we refer to the statistical chart below, the private and public data storage percentage can be seen as fluctuating owing to multiple reasons, such as the quantity of information, security levels rising, and cost. Not to mention the technological advancements, enterprise data storage, on the other hand, is showing a stable path indicating a different trend altogether (Fig. 1).

Customers are creating an astonishing volume of data and media files every moment, especially since the frequent usage of hand held devices are powered with softwares and continuous streaming of audio and video information obtainable at the internet. Moreover, due to the dominance of Internet of Things(IoT) providing functional and pleasurable benefits to all customer segments, the quantity of personal data is growing unanimously. This mounting personal content generation is creating an income avanue for the sellers in the data storage market (Fig. 2).

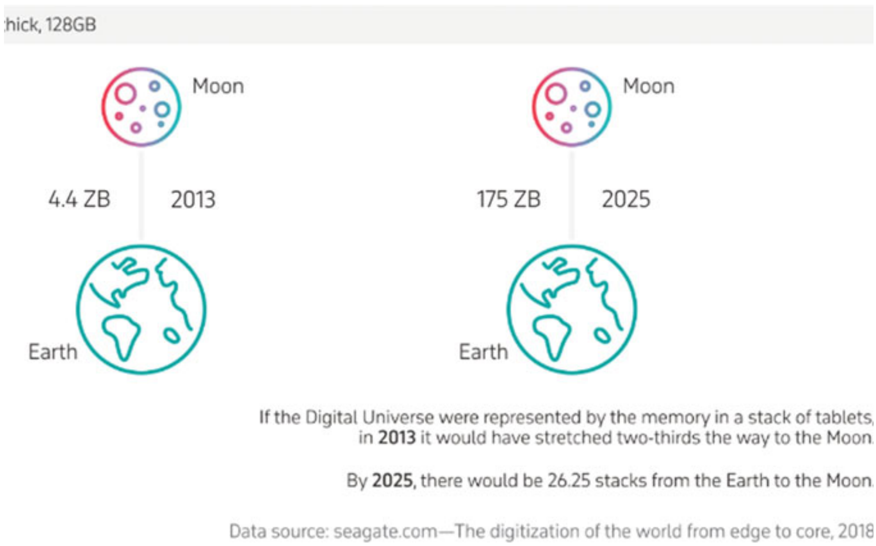


Fig. 1 *Source* Allied Market Research

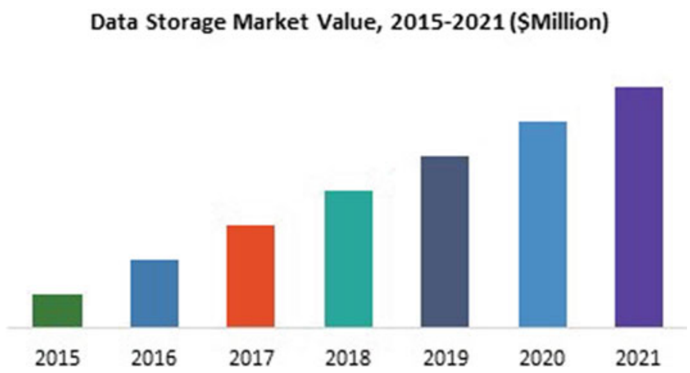


Fig. 2 Data storage market value (2015–19). Source: IndustryARCAnalysis Experts Insights

- *AI requirement*

The important aspect for the proper functioning of AI is a single repository that is tied to AI and MLL and deep learning. These repositories must be able to coordinate with physical data centers as well as cloud configuration should be able to support the exabyte of data. All big players are vying to utilize sophisticated AI data platforms based on data configuration, ingestion, organization, distribution, machine learning, deep learning, and archiving.

1.3 Data Storage Market Growth Drivers

According to the statistics presented by The Software Alliance, Global Cloud Computing Scorecard (BSA), cloud computing has a presence in around 24 countries currently. If we refer to the statistical chart below, the private and public data storage percentage can be seen as fluctuating owing to multiple reasons, such as the quantity of information, security levels rising, and cost. Not to mention the technology advancements, enterprise data storage, on the other hand, is showing a stable path indicating a different trend altogether. The ease of storing the large amount of data and accessing the data over virtual space has stimulated the adoption of cloud computing and cloud technology across various industry globally and has forced the industry to look for cloud storage options and also hard pressed the cloud storage market growth. Data storage device and the methods are not completely secure. Always there is a threat and the prospect of data theft and data loss. Hence, individual consumers are also moving towards cloud computing methods and cloud storage systems which is pushing the growth of Data storage market on one hand but is also forcing the downward trend of revenue generation individual data storage devices on the other hand. Neverthless, the vendors are endeavouring to reinforce the security protocol in individaul products by modifying them and adding more storage and security, thereby becoming updated and competitive in the market.

1.4 Data Storage Market Key Player's Perspective

The market is largely held by Western Digital (WD), with revenue of $19.1 billion in fiscal 2017 in primary data storage, but when we consider the next-generation data storage market by storage system, storage architecture, storage medium, and end user, the players are majorly HP, Dell, Hitachi, and IBM with Dell holding the major share (Markets and Markets report March 2019). Western Digital market share is credited to its HDD persistent leadership. The company showcased its updated third generation 10 TB HDD and dispatched worldwide 16 million HDDs.

2 Literature Review

The literature suggests many articles which indicated the gradual development of software data storage systems and the challenges faced by users and clients. The growth is driven by customer requirements and technology development in each stage. It also shows the usage of data as well as issues like privacy and price involved. In 1996, [1] discussed log device-based data storage systems which had the functionality of providing proficient storage and the retrieval of data. It was performed on a computer system with respect to its operating system attached to the data storage system. Similarly, this was supported by the article and the patent [2] for inventing a "file transfer utility which employs an intermediate data storage system". The claim was regarding communicating and transferring information from one process to another (each having its own operating systems) through a shared storage medium.

One research article referenced "data set level mirroring to accomplish a volume merge/migrate in a digital data storage system" by Robert M. Perego (2003). The research was about merging data extracted from multiple individual devices (disk volumes) to larger DASD devices and combining data set extents. Philippe Armangau (1999) proposed and patented (US 6434641B1) a "snapshot copy facility for a data storage system permitting continued host read/write access". He claimed that while maintaining the data set, a snapshot copy of the production data set is created at a multiplicity of storage locations in the data set providing access to the host processor for read/write access during maintenance of the snapshot copy. Blumenau and Raz [3] researched and patented "virtual ports for data transferring of a data storage system", which is a data storage subsystem comprising storage and data storage controller.

Pre-allocation of file system cache blocks in a data storage system was researched and patented [4]. Their claim was to operate a file server with a file system reserve memory and storage. The next level of research is regarding the methods and devices for providing security for a data storage system by [5]. Around the same time, [6] researched "system and method for restoring data on a data storage system," which highlights the requirement of backup as well as restoring the data for emergency on the data-based storage system as the image of the primary computer system. Virtual private network software system requirement and research were simultaneously being conducted, but major breakthrough was when [7] VPN, wherein a virtual private network system provider software is delivered from the server computer to setting up a virtual private network connection to multiple client computers over a network. Further research in this area was carried out [8], where the focus was on a cloud computing system for storing private information without the interference of provincial privacy laws. This is supported by a [9] discussing the cloud-based storage software system "USTO.RE: A Private Cloud Storage Software System", describing the model of cloud storage hardware software and platform is perceived as services. Data as a service is highlighted, which is provided to the customers irrespective of their location. There are attempts by organizations to build private data storage systems instead of outsourcing the data storage services for file hosting, backing up data, and archiving the data to public providers. There are challenges as when access is required, the server provider might not be available as well as the assurance of management procedures along with control of data is not certain in this case [10]. There is evidence of research regarding synchronizing proprietary data in an external cloud with data in a private storage system [11]. The claim and the patent here are about coordinating proprietary data in a public cloud provided by a cloud service provider with data of a private storage system. When we look at the amount of information of diverse importance, we wish to store the same. But storing it at private data storage system might be costly in terms of accessing and client profiles. It made 20 claims latest being configuring the device to store some recovery information in the private storage cloud as well [12].

3 Research Methodology

The research has used data analytics, descriptive statistics, and some forecasting techniques to gain valuable insight into the present and coming future. The research approach involves the collection of data from ERP forms. The form questions were prepared and uploaded to ERP systems. In this ERP system, all the sellers and marketing folks have access who puts data from start to end from getting any customer source or probable buyer. The main issue with data was storage and access as there were records of the last 10 years pertaining to different functional aspects—region-wise and client-wise. This ERP data has columns of customer name, date of origin, win/loss reason, recommendation, etc. The data captured on ERP was exported on the Excel file, and this is for consecutive 3 years, i.e., years 2016, 2017, and 2018. A preprocessing was done on the CSV file using the PIVOT tool. Excel pivot table is used to build the charts from the data. Trend analysis was used as the basis for projection.

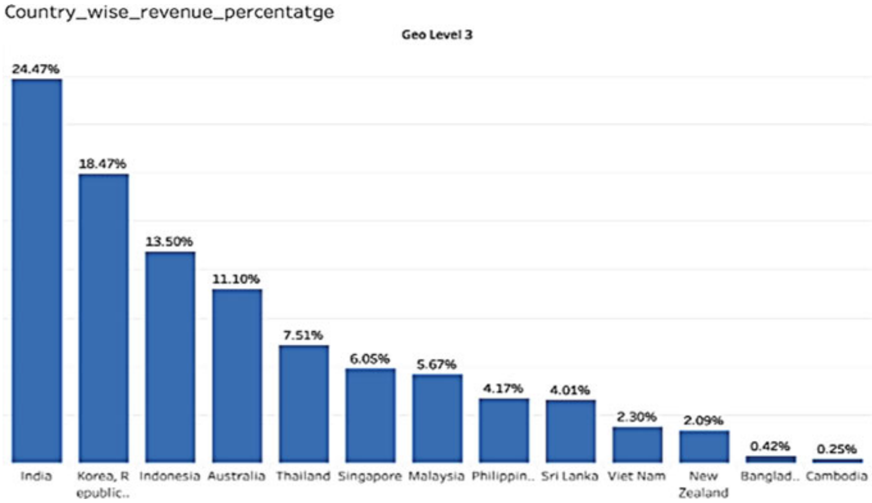


Fig. 3 Country-wise revenue generation in the Asia Pacific

4 Data Findings

After analyzing the data from the above survey, the findings are listed below. It can be seen that in the Asia Pacific, ASEAN generates the highest number of win opportunities. Since storage is a technical product, targeting the right audience would get us the deals as we can see that the highest number of win opportunities came via IT Developer/IT Manager followed by Chief Information Officer. Observation shows that there is a decline in the opportunities from the year 2016 onward though we tried to gain in 2017 yet again falling in the year 2018. It is observed that in the year 2016 Q3 onward, there are more lost opportunities. Findings indicate the top audience via different geography-wise who has influenced the deals. We also see the trend of change in the preference for storage options over the years. It is more based on the convenience of accessing and how it is stored. The focus here is more on the credibility and privacy of the system. Third-party storage is there in the system for a long, but ownership and control are something corporates are seeking for managing the huge data which is churned out year after year. More so, the analysis of the data, while it is stored and interpreting on the same YoY basis, is also called for. The demand for total congruence with data systems is actually becoming crucial, which makes it demanding for clients (Figs. 3, 4 and 5).

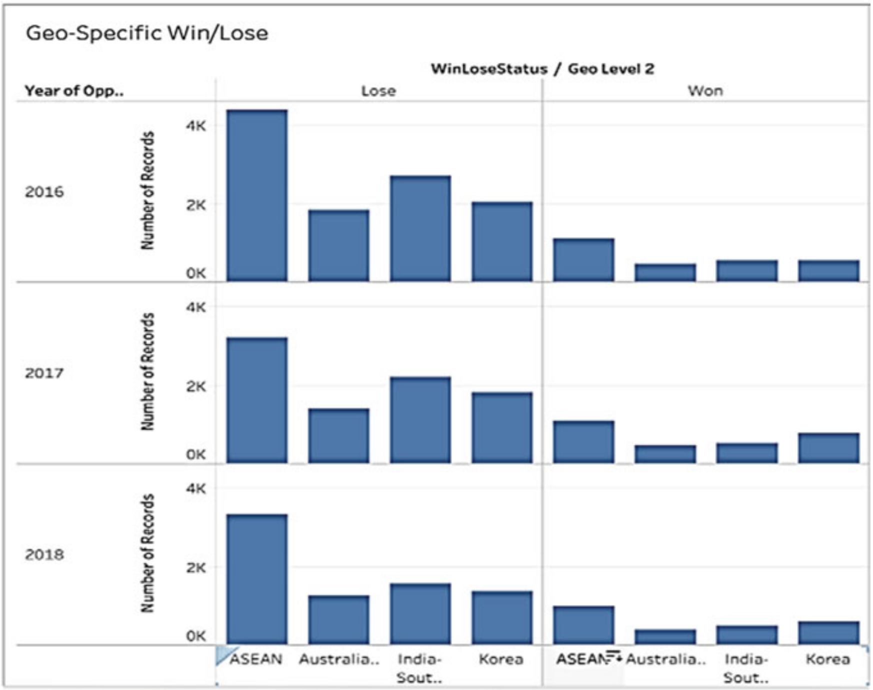


Fig. 4 Explains that in the year 2016, 2017, and 2018, ASEAN provided more deal opportunities, and the least one was Australia in all the three years

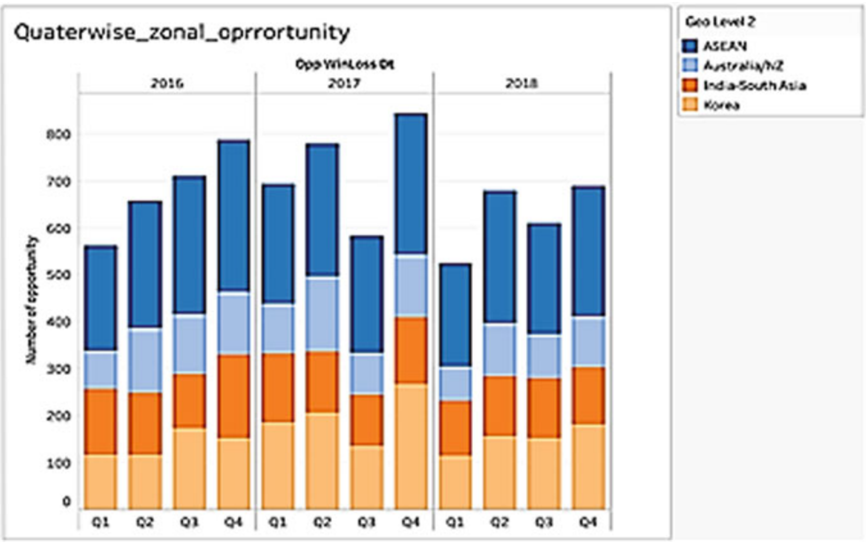


Fig. 5 Total number of opportunities as per GEO wise

5 Conclusion

It can be seen that in the Asia Pacific, ASEAN generates the highest number of win opportunities. Since storage is a technical product, targeting the right audience would get us the deals. Observation shows that there is a decline in the opportunities from the year 2016 onward though we tried to gain in 2017, yet again falling in the year 2018. It is observed that in the year 2016 Q3 onward, there are more lost opportunities. Findings indicate the top audience via different geography-wise who has influenced the deals. We also see the trend of change in the preference for storage options over the years. It is more based on the convenience of accessing and how it is stored. The focus here is more on the credibility and privacy of the system. Third-party storage is there in the system for long, but ownership and control are something corporates are seeking for managing the huge data which is churned out year after year. More so, the analysis of the data, while it is stored and interpreting the same YoY basis, is also called for. The demand for total congruence with data systems is actually becoming crucial, which makes it demanding for the clients.

6 Discussion

The study shows the significance of the data storage system with the advancement in technology and also the customer requirement, which is showing a sea change. The historical evolution of the storage device also supports the exposure of the customer to changing trends and the change in their attitude. Earlier, storage was the sole objective but competition in alternatives available as per cost, convenience, compatibility, and, most crucial of all, trust and security are the deciding factors for identifying and utilizing the data storage system. Further, there is a great scope in the next-generation data storage market, which is growing at the rate of 12.48% CAGR. The growth driver of the next-generation data storage market is individual hand-held device creating massive digital data per user. There are indications of the highest CAGR growth of NAS in the next-generation data storage as it offers a cost-effective solution that is the first choice for small- and medium-scale enterprises. The research also points to the enterprise growing to hold the major share in the next-generation data storage market. Most of the operational decisions and the applications are based on stored data related to the customers and their requirements Even the forecasting also requires past data analysis. This, in turn, accelerates the data generation in bulk, which makes the foundation of future prospects.

7 Implications

The implications of this study are for different stakeholders with direct usage like enterprises, and the end consumes along with data storage providers who are transitioning and upgradation. With the turn of century, technology has also advanced in leaps and bounds. Individuals are seeking virtual solutions with the help of their devices having the latest software tools. They are eager to make their presence felt in each sphere by generating content and storing it on the cloud for others to analyze. Corporations are taking leverage of this newfound attitude, and it is a win–win situation for all. This research implores the evolution of storage devices and a comprehensive analysis of the market. This aims to identify and utilize the data storage system.

References

1. Ledain JE, Colgrove JA, Koren D (1999) U.S. Patent No. 5,996,054. U.S. Patent and Trademark Office, Washington, DC
2. Yanai M, Vishlitzky N, Alterescu B, Castel D, Shklarsky GG, Ofek Y (2007) U.S. Patent No. 7240238 B2. U.S. Patent and Trademark Office, Washington, DC
3. Blumenau SM, Raz Y (2002) U.S. Patent No. 6,421,711. U.S. Patent and Trademark Office, Washington, DC
4. Zheng J, Jiang X, Vahalia U (2003) U.S. Patent No. 6,571,259. U.S. Patent and Trademark Office, Washington, DC
5. Scherr AL (2006) U.S. Patent No. 7,134,138. U.S. Patent and Trademark Office, Washington, DC
6. Becker GA, Condon R, Knight DA, Medders DE, Rowan M (2009) U.S. Patent No. 7,617,414. U.S. Patent and Trademark Office, Washington, DC
7. Clayton T, Rakes RB (2007) U.S. Patent No. 7,216,173. U.S. Patent and Trademark Office, Washington, DC
8. Mawdsley G, Meyfroidt S (2015) U.S. Patent No. 9,202,085. U.S. Patent and Trademark Office, Washington, DC
9. Durao F, Assad R, Fonseca A, Fernando J, Garcia V, Trinta F (2013) USTO.RE: a private cloud storage software system. In: International conference on web engineering. Springer, Berlin, 8 Jul 2013, pp 452–466
10. Gonzalez JL, Perez JC, Sosa-Sosa V, Cardoso JF, Marcelin-Jimenez R (2013) An approach for constructing private storage services as a unified fault-tolerant system. J Syst Softw 86(7):1907–1922. Hass Y, Vishlitzky N, Raz Y (2002) U.S. Patent No. 6,438,586. U.S. Patent and Trademark Office, Washington, DC
11. Engbersen T, Schmatz M (2019) U.S. Patent No. 10,230,792. U.S. Patent and Trademark Office, Washington, DC
12. Shen Q, Yang Y, Wu Z, Wang D, Long M (2013) Securing data services: a security architecture design for private storage cloud based on HDFS. Int J Grid Util Comput 26 4(4):242–254. Xu J, Xi W, Guo G, Bjornson E, Yu J, Wei LC (2018) U.S. Patent Application No. 15/635,039

Recent Advances in Semantic Segmentation for Sports Analytics

A. Bharathi Malakreddy, Sadanand Venkataraman, Mohammed Sinan Khan, and Srinivas Padmanabhuni

Abstract Recent advances on deep learning has enabled significant progress in the field of Sports Analysis. Semantic segmentation is a technique that involves dividing an image or video into multiple segments or regions and labeling each segment with a class or category. In sports, semantic segmentation can be used for action recognition, pose estimation, and athlete tracking during competitions or training sessions. This has the potential to revolutionize how we analyze and understand athlete performance, leading to better training strategies, injury prevention, and performance optimization. This paper provides a survey of image segmentation techniques applied in various sports, including their applications and limitations. It also proposes a model that addresses the concerns discussed above by using a framework that takes advantage of the full range of information available, a PSPNet model for precise localization, metamorphic testing for evaluating model behavior, and explainability to understand the factors influencing model predictions. By exploring current applications of semantic segmentation, we hope to encourage further research in this area and facilitate the adoption of these techniques in the sports industry words.

Keywords Convolutional neural networks · Semantic segmentation · Sports analytics · Action recognition · Pose estimation · Performance optimization

A. B. Malakreddy (✉) · S. Venkataraman · M. S. Khan
Department of Artificial Intelligence and Machine Learning, BMS Institute of Technology and Management, Bengaluru, India
e-mail: bharathi_m@bmsit.in

S. Venkataraman
e-mail: saadanandv@gmail.com

M. S. Khan
e-mail: msinankhan@gmail.com

S. Padmanabhuni
AiEnsured, Bengaluru, India
e-mail: srinivas@testaing.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_26

1 Introduction

Computer vision is a field of study that enables machines to interpret visual information from the real world [1]. One of the popular techniques in computer vision is semantic segmentation, which involves dividing an image or video into multiple segments or regions and labeling each segment with a class or category. This allows a computer to understand the visual content of an image or video on a pixel-by-pixel level, rather than just recognizing global objects or scenes. This is different from other techniques in computer vision, such as object detection or classification, which only identify and label global objects or scenes in an image or video. This enables a computer to understand the fine-grained details and relationships between different objects, scenes, and background elements in a visual input.

Semantic segmentation has a wide range of applications in various fields such as medical imaging [2, 3], autonomous driving, and robotics. In sports, semantic segmentation can be applied in various ways, such as action recognition, posing estimation, and athlete tracking. Furthermore, it can be used for better training strategies, injury prevention, and performance optimization. It can be used to recognize specific actions performed by athletes, and track athletes or objects during competitions or training sessions.

In the domain of sports, action recognition is a prominent application of semantic segmentation [4]. By identifying and classifying every pixel in a sequence of images, semantic segmentation can be used to recognize specific actions performed by athletes, such as running, jumping, or throwing.

Semantic segmentation is also used in pose estimation, which involves identifying the key points on an athlete's body and tracking their movements over time [4, 5]. This can be useful in analyzing technique and form, as well as identifying areas where an athlete may be prone to injury.

Lastly, by utilizing semantic segmentation for identifying and tracking individual athletes or objects within a scene, coaches, and analysts can gain insights into athlete movement patterns, tactics, and team dynamics.

In this paper, we review the existing literature on semantic segmentation techniques in sports and discuss their merits. Finally, we highlight some of the challenges and future research directions in this field. The rest of the paper is structured as follows: Section 2 describes some of the significant and relevant works that were carried out as well as the necessary background information. Section 3 discusses the possible enhancements of the existing systems. Section 4 concludes by summarizing the scope of Artificial Intelligence in the Sports Domain.

2 Literature Survey

In this section, we provide a comprehensive literature survey of image segmentation techniques for sports. The recent papers give us insights into the current technology and techniques used in image processing, semantic segmentation, and AI in sports. Table 1 presented herein furnishes a comprehensive summary of the reviewed papers, accompanied by discrete findings derived from each of the enumerated papers. Our review sheds light on the current state-of-the-art in image segmentation techniques for sports, highlighting their strengths.

3 Possible Enhancements

A review of the literature has uncovered studies that have examined the application of Artificial Intelligence, deep learning, and computer vision in sports-related scenarios. Nevertheless, the following are some areas that require attention including:

- **Biased datasets used to train models** [7]: This can be addressed by a PSPNet model that offers several advantages for achieving accurate semantic segmentation through its multi-scale feature representation, ability to capture contextual information from the scene, and production of highly-precise object boundaries.
- **Computational complexity** [4, 9]: Rendering live telecasting is difficult for sporting events utilizing such models, to address this, the computational efficiency of PSPNet enables real-time performance making it appropriate for use in live sports broadcasting or athlete tracking scenarios.
- **Very limited testing techniques have been implemented** [9]: This indicates the need for more robust methodologies to evaluate the model's performance: Metamorphic testing can yield various advantages. It can significantly enhance the test coverage to guarantee optimal performance of the model across different scenarios and edge cases. Furthermore, it reduces dependence on labeled data by producing fresh test cases from current data sets.
- **Limitations in generalization due to insufficient dataset sizes** [2, 4]: Metamorphic Testing can also help reduce dependence on labeled data by producing fresh test cases from current data sets.
- **Most of these models employ unimodal approaches** [1, 2, 4, 6–11, 14]: A multimodal approach, utilizing various modes of information including a CNN for accurate localization, metamorphic testing to assess model performance and explainability techniques to comprehend the underlying factors influencing predictive outcomes can enhance the development of effective models.
- **Lack of AI Explainability**: The implementation of AI explainability can contribute to the development of transparent models by revealing previously unnoticed biases that were not detected through conventional evaluation metrics. This technique is also useful in determining which features are most significant for making accurate predictions.

Table 1 Summary of papers with remarks

S. No.	Publication details	Methodology	Results	Remarks
1	"A review of semantic segmentation using deep neural networks" [1]	• An outline of leading methodologies along with their advantages, limitations, and principal challenges is given • Top approaches to semantic segmentation models built using deep convolutional neural networks are elaborated	• Using deep learning techniques have shown significant improvements in semantic segmentation accuracy • Many state-of-the-art approaches use pre-trained models, which can be fine-tuned for specific tasks • Some approaches can handle complex scenes with multiple objects and occlusions	• Large volumes of labeled data are a vital prerequisite for a significant number of approaches, which might not be available • The performance of these approaches is highly dependent on the quality of training data • Some approaches struggle with segmentation of small objects or objects with low contrast • The challenges of segmenting with audience members and foreground objects are discussed, and the significance of comprehending the impact of errors on the complete system is emphasized
2	"Rethinking Panoptic Segmentation in Remote Sensing: A Hybrid Approach Using Semantic Segmentation and Non-Learning Methods" [2]	• The Original Image, Semantic Image, and Panoptic Image fed into the pipeline as inputs • This presents a novel software tool that enables the automated creation of samples in the Common Objects in Context (COCO) annotation format through the utilization of input data and point shapefiles • The software generates deep learning samples with annotations format automatically	• The dataset was evaluated using Panoptic-FPN • The assessment resulted in a mean intersection over union (mIoU) of 93.865, an average precision (AP) of 47.691 and a panoptic quality (PQ) score of 64.979 • The results show that the ResNet-101 backbone presented slightly better results than the ResNet-50 backbone among all metrics	• The software developed works well, as it was specifically designed for the TIFF format • The dataset generated is relatively small, with only 3400 samples, which may limit its generalizability to other remote sensing datasets • The evaluation of the dataset was limited to the urban setting

(continued)

Table 1 (continued)

S. No.	Publication details	Methodology	Results	Remarks
3	"Artificial Intelligence for Sport Activitity Recognition" [4]	• For sports activity (i.e., gestures) detection and recognition, an Artificial Neural Network approach is employed which utilizes a PIQ ROBOT to collect data • The approach involves training the neural network with a dataset of tennis gestures and their corresponding labels	• The proposed Artificial Neural Network approach achieved an accuracy of 94.5% in detecting and recognizing tennis gestures using the PIQ ROBOT device • This approach was able to accurately recognize different types of tennis gestures, including forehand, backhand, serve, and smash	• The dataset used for training the neural network is relatively small and may not be representative of all possible variations in tennis gestures • It does not discuss the computational requirements or limitations of the proposed approach
4	"IAUFD: A 100 k images dataset for automatic football image/video analysis" [6]	• The dataset comprises 100,000 real-world images taken from 33 different football videos whose duration sums to 2508 min were annotated into 10 different categories • Following are the 10 categories: – Goal, Field, center Celebration, Red card, Yellow card, Ball, Stadium, Referee, Penalty-kick, Free-kick • The dataset covers various weather, seasons, time of day, and location to ensure generalization	• The dataset was evaluated for object recognition and event detection using two deep neural networks • The architectures of VggNet-13 and ResNet-18 achieved an accuracy of 92.5% and 94.5%, respectively • The multiple occurrences experiment showed that the model can detect multiple instances of the same event in a single image with an accuracy of 87.5% and 89.5% for VggNet-13 and ResNet-18, respectively • The multiple labels experiment showed that the model can detect multiple events in a single image with an accuracy of 85.5% and 87.5% for VggNet-13 and ResNet-18, respectively	• The resulting dataset called "IAUFD Dataset" was evaluated on several Computer Vision Tasks including Object Detection, Action Recognition, and event detection • The models trained on the IAUFD dataset outperformed models trained on other football-related datasets, demonstrating the effectiveness of the IAUFD dataset for football-related machine learning applications

(continued)

Table 1 (continued)

S. No.	Publication details	Methodology	Results	Remarks
5	"On Designing Deep Learning Approaches for Classification of Football Jersey Images in the Wild" [7]	• JerseyXIV- A dataset encompassing 7840 football jersey images that include a diverse range of 14 distinct types belonging to 10 different teams from the 2015 Big 12 Conference season • Several traditional classification techniques employing deep learning were experimented with, starting with two traditional methods, before transitioning to convolutional neural networks	• The CNN-F design for convolutional neural network (CNN) achieved an impressive 92.61% accuracy in classification tasks • A test accuracy of 96.90% was attained through the combination of three distinct CNNs	• A total of 3584 frames. Out of these frames, 2188 frames contained the object of interest. This dataset served as the testing ground for trained CNN models when evaluated on an extensive and practical Scale-Oriented application scenario • To evaluate the drop in accuracy of the average-fusion method, three Image Quality Assessment techniques were employed • The study recommends fusion of networks and data augmentation as an approach for classification
6	"Using Artificial Intelligence for Pattern Recognition in a Sports Context" [8]	• It utilizes Futsal as a case study to recognize athletic actions in the context of a match • Wearable devices such as TraXports and MBody3 were used to collect physiological and positional data • The collected data was processed using MATLAB to create the classification architecture • The data analysis framework involved identifying relevant features for action recognition through feature selection	• Three classification methods were compared against each other, namely traditional Artificial Neural Networks (ANN), Long Short-Term Memory Network (LSTM), and Dynamic Bayesian Mixture Model (DBMM) • Assessments were carried out by weighing the trade-off between precision and recall, revealing that DBMM exhibited superior performance as it attained an $F1$ score of 80.54%	• It uses two wearable devices to collect data, which may not capture all relevant physiological and positional data • It does not consider the impact of external factors, such as the opponent's strategy or the referee's decisions, on the athlete's actions

(continued)

Table 1 (continued)

S. No.	Publication details	Methodology	Results	Remarks
7	"Soccer Event Detection Using Deep Learning" [9]	• Three main modules are included in the proposed method: a VAE module, an image classification module, and a fine-grain image classification module • To discern between images of soccer and the ones that don't pertain to the sport, the VAE module is used • To categorize images of specific occurrences within the game, the module for image classification is employed • Finally, to classify the images of red and yellow cards in soccer matches the fine-grain image classification module is utilized	• Evaluation of the model was conducted on the public dataset known as "SoccerNet", yielding promising outcomes • An accuracy of 74.2% was reported for frame-level event detection, while sequence labeling accuracy was reported as 67.8% • The findings demonstrated the significance of both CNN and LSTM components in achieving high precision in detecting soccer events	• The experiment results suggest the effectiveness of the proposed approach, in which a benchmark dataset was used and high accuracy was achieved • The method proposed is evaluated on a single dataset • It is difficult to implement the proposed method in real-time applications due to the lack of explanation regarding its computational complexity
8	"Classification of Similar Sports Images Using Convolutional Neural Network with Hyper-Parameter Optimization" [10]	• An approach has been devised utilizing deep learning Techniques that automatically categorize closely related sports disciplines • Transfer learning, which involves fine-tuning a pre-trained convolutional neural network model on a new dataset has been employed in the method • The optimization of neural network topology and the essential training parameters was achieved by automated hyper-parameter optimization using differential evolution	• The performance was assessed through a tenfold cross-validation procedure • The overall accuracy of 91.46% was achieved by the proposed method • Different hyper-parameters' effect on the classification performance was investigated by conducting an ablation study	• The authors suggest that future work could explore the use of more advanced CNN architectures and larger datasets for even better performance

(continued)

Table 1 (continued)

S. No.	Publication details	Methodology	Results	Remarks
9	"Action detection in complex scenes with spatial and temporal ambiguities" [11]	• An academic proposal has been put forth for a multiple-instance learning framework, known as SMILE-SVM. Its purpose is to acquire knowledge on detecting human action through imprecise action locations • The recognition accuracy of bags is aimed to be increased and the margin of the classifier is attempted to be maximized simultaneously by SMILE-SVM • The algorithm handles the uncertainties regarding the spatiotemporal positioning of human actions	• Three challenging datasets (KTH, Weizmann, and UCF sports) were evaluated • The experimental results indicate that state-of-the-art performance was achieved • The method proposed was able to attain a mean precision rate of 93.8%, 95.6%, and 83.3% on the KTH, Weizmann and UCF sports datasets, respectively	• Multiple candidates in a short sequence can be inputted into the algorithm to infer whether the action of interest occurs • The validation of the complementary nature of motion and appearance features in action detection has been demonstrated • The ability to detect actions using a combination of motion and appearance features has been confirmed • The algorithm's ability to identify actions has been evidenced by validating how motion and appearance features complement one another
10	"Multimodal deep learning" [12]	• An integrated deep learning framework is proposed, one for text and one for images, into a single model • The multimodal deep learning framework is composed of three main components: a modality-specific feature extractor, a fusion module, and a classifier • For each modality, the input data is processed by a modality-specific feature extractor to generate a set of features • The features are combined by the fusion module to produce a joint representation that captures the relationship between the modalities • The joint representation is used by the classifier to predict the output • A regularization method for the fusion module is also proposed, which encourages the network to learn representations that are robust to missing modalities	• State-of-the-art performance is achieved on several tasks, including image classification and captioning • The alignment of different modalities' representations is effectively learned by the framework, even when the modalities are not perfectly aligned, as demonstrated by the authors • The proposed multimodal deep learning framework demonstrates superior performance on various tasks such as image classification and captioning • Alignment of different modalities' representations, even in cases of imperfect alignment, is effectively learned by the framework • The alignment of representations from different modalities is effectively learned by the framework, even in situations where the modalities are not perfectly aligned	• The paper does not provide an assessment of the proposed framework on a large dataset. It would be interesting to see how it performs on a larger and more diverse dataset

(continued)

Table 1 (continued)

S. No.	Publication details	Methodology	Results	Remarks
11	"Multimodal deep learning for robust RGB-D object recognition" [13]	• An advanced system for the reliable detection of objects with RGB-D features amidst disturbances and confusion is presented, utilizing deep learning methodologies • The framework comprises of two modules: an RGB-D feature learning module and a classifier module • The RGB-D feature learning module is made up of two sub-networks, one for each modality, that process the RGB and depth images separately. Both sub-networks have a convolutional neural network (CNN) architecture, with shared convolutional layers and separated pooling layers • A fully connected layer is used to generate the joint feature representation from the concatenated outputs of the convolutional layers	• Two RGB-D Datasets, namely "NYU depth dataset" and the "Washington RGB-D dataset", are used to assess the effectiveness of the proposed framework • It outperforms several cutting-edge methods for RGB-D object recognition, achieving an accuracy of 72.1% on the "NYU dataset" and 75.8% on the "Washington dataset" • The efficacy of the proposed framework is demonstrated through practical situations where a mobile robot fitted with an RGB-D camera is utilized	• This paper explores a multimodal deep learning model for object detection, using CNNs and RNNs to process visual and depth data • The framework is tested only on a small set of objects • This framework should also be explored in other domains of object recognition

4 Conclusion

The study is carried out on various techniques used for image segmentation in sports, focusing on soccer, basketball, and tennis. The existing techniques were evaluated, and their merits and gaps were identified, highlighting the potential use of Artificial Intelligence in sports.

Despite the promising results in areas such as semantic segmentation, activity recognition, and pattern recognition, this study suggests further enhancements to improve the accuracy and effectiveness of image segmentation algorithms in real-world situations.

The following are potential areas of improvement that could be considered:

- **Diverse Datasets**: To improve the accuracy of image segmentation algorithms, the use of more diverse datasets that represent various scenarios and conditions.
- **Real-Time Constraints**: The design of image segmentation algorithms should consider real-time constraints as this can significantly affect the accuracy and effectiveness of the algorithm in practice.
- **Limitations of Proposed Techniques**: It is important to identify the limitations of the proposed techniques to understand the areas where they may not work effectively and identify areas for further improvements to act as fodder for further research.

This paper suggests several potential enhancements that could be made to the existing work on image segmentation in sports. By addressing these areas, researchers can make significant contributions to the field and improve the accuracy and effectiveness of image segmentation algorithms in sports.

Acknowledgements We would like to express our sincere gratitude to Dr. Srinivas Padmanabuni and AiEnsured for their invaluable guidance and support throughout the research process. We also wish to thank The Vision Group of Science and Technology, Government of Karnataka for their financial support, as well as The Principal, BMS Institute of Technology and Management for their support in resources for this paper.

References

1. Guo Y, Liu Y, Georgiou T et al (2018) A review of semantic segmentation using deep neural networks. Int J Multimed Info Retr 7:87–93. https://doi.org/10.1007/s13735-017-0141-z
2. de Carvalho OLF, de Carvalho OA Jr, de Albuquerque AO, Santana NC, Borges DL (2022) Rethinking panoptic segmentation in remote sensing: a hybrid approach using semantic segmentation and non-learning methods. IEEE Geosci Remote Sens Lett 19:1–5. Art no. 3512105. https://doi.org/10.1109/LGRS.2022.3172207
3. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88. https://doi.org/10.1016/j.media.2017.07.005

4. Bezobrazov S, Sheleh A, Kislyuk S, Golovko V, Sachenko A, Komar M, Dorosh V, Turchenko V (2019) Artificial intelligence for sport activity recognition, pp 628–632. https://doi.org/10.1109/IDAACS.2019.8924243
5. Fu M, Zhong Q, Dong J (2022) Sports action recognition based on deep learning and clustering extraction algorithm. Comput Intell Neurosci 2022:9. Article ID 4887470. https://doi.org/10.1155/2022/4887470
6. Zanganeh A, Jampour M, Layeghi K (2022). IAUFD: a 100 k images dataset for automatic football image/video analysis. IET Image Proc 16:n/a–n/a. https://doi.org/10.1049/ipr2.12543
7. Matta RR (2017) On designing deep learning approaches for classification of football Jersey images in the wild. Graduate theses, dissertations, and problem reports, 6178
8. Rodrigues ACN, Pereira AS, Mendes RMS, Araújo AG, Couceiro MS, Figueiredo AJ (2020) Using artificial intelligence for pattern recognition in a sports context. Sensors 20(11):3040. https://doi.org/10.3390/s20113040
9. Karimi A, Toosi R, Akhaee MA (2021) Soccer event detection using deep learning. ArXiv. /abs/2102.04331
10. Podgorelec V, Pečnik Š, Vrbančič G (2020) Classification of similar sports images using convolutional neural network with hyper-parameter optimization. Appl Sci 10(23):8494. https://doi.org/10.3390/app10238494.Y. Hu et al (2009) Action detection in complex scenes with spatial and temporal ambiguities. In: 2009 IEEE 12th international conference on computer vision. Kyoto, Japan, pp 128–135. https://doi.org/10.1109/ICCV.2009.5459153
11. Ngiam J et al (2011) Multimodal deep learning. In: Proceedings of the 28th international conference on machine learning (ICML-11), pp 689–696
12. Eitel A et al (2015) Multimodal deep learning for robust RGB-D object recognition. In: 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). Hamburg, Germany, pp 681–687. https://doi.org/10.1109/IROS.2015.7353446
13. Ravishankar A, Natarajan S, Bharathi Malakreddy A (2022) SparseMAX: accelerating quantum neural networks on GPU clusters using sparse-matrix kernels. In: Embedded computer systems: architectures, modeling, and simulation: 21st international conference, SAMOS 2021, virtual event, proceedings. Springer International Publishing, Cham, 4–8 July 2021
14. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965

Intelligent Decision Analysis to Stimulate Student Learning

Rahul Dhaigude and Ruby Chanda

Abstract Motivating students to learn is at the heart of teachers today. With the development of computer technology, the introduction of multimedia technology into the daily teaching process has become very popular and is widely adopted. Matching the ability to address the strategies that encourage intelligent decision-making for the problems of different students is particularly important nowadays. This article is based on multimedia technology, multimedia teaching, establishing hierarchical structure analysis, effective and stimulating learning interest, and sharing program resources as a guide in learning. The multimedia modeling and teaching process takes into account the teaching objectives, the program's characteristics and the students' cognitive characteristics. According to the weighted and normalized analysis method, an intelligent decision system about teacher incentive strategy is designed to stimulate students' learning motivation.

Keywords Intelligent decision-making · Multimedia teaching · Quantitative analysis · Learning interest · Teaching experiment

1 Introduction

Faced with increasingly fierce international competition as the knowledge economy, the dawn of the information age, and the acceleration of the global integration process, we need many qualified management talents. The traditional educational model is teacher-centric. Knowledge transfer relies primarily on the indoctrination of teachers to students. Demonstrating teacher initiative and enthusiasm is difficult and does not lead to creative talent development. Information technology is the world's most active, fastest-developing, and most prevalent invention in science and technology.

R. Dhaigude (✉) · R. Chanda
Symbiosis Institute of Management Studies, Symbiosis International (Deemed) University, Pune, India
e-mail: rahul.dhaigude@sims.edu

R. Chanda
e-mail: ruby.chanda@sims.edu

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_27

In this paper, we set up a quantitative group index structured model in view of the negative psychological problems of Management students in learning. By making a detailed quantitative assessment and analysis on the strategy for college teachers in stimulating students' learning motivation, we obtain the optimal strategy and its practical application value.

2 Related Research

2.1 The Effectiveness of Strategy for Teachers to Stimulate Students' Learning Motivation

Another feature of information technology is its integration, that is, images, sound, text, animation, and numerical in one, the results show integrated on the screen, this feature is conducive to the formation of the sensory stimulation of the learner. In the design of classroom exercises, PowerPoint presentations are used to explain patterns and paragraphs designed as a slide, with a certain color, with pictures or animation effects. The students master the important words or phrases in a very eye-catching form; or all the important words and phrases connected into a funny essay or the text of the repetition of the content designed to fill in the blank [1]. Let the students fill in the blank, and then the correct answer to a variety of dynamic effects on the screen, if necessary, accompanied by encouraging or congratulatory sound effects. In the colorful classroom, students not only practice to remember the knowledge, but also experience the human–computer interactive learning with pleasure and joy [2]. This can be described as teaching 'Yu Le: Happy to Learn', remember firmly, improve the efficiency of teaching. At the same time, students learn to understand foreign social environments, customs, national psychology, history and culture through networks, and case study [3]. Teachers are very helpful in delivering relevant course's educational content. Also helps in data collection and processing, learning content processing, classification into multimedia, hypertext to create visual images which are vibrant and interesting and easy-to-understand. By allowing students to explore in these situations, they can identify problems themselves, perform hands-on work, suggest solutions to problems and opportunities, and help students understand what they are learning and enhance their learning skills [4, 5]. Also help students to further enhance their inquiring spirit.

Due to the differences between Indian and western cultures, Indian learning forms are quite different, leading to a polarized development of students' interest of learning. In this case, it is difficult for most students to learn. Although management education in India has been highly concerned about the students' knowledge application ability, it is still insufficient because of the defects in educational mechanism, especially for the performance of students' understanding and it's application abilities [6]. Although our country attaches great importance to education and there are more and more educational institutions are being established, the guiding principles between knowledge and practice are different. In this case, management education tests the students for high score, leading to the students' negative attitude toward learning application-based content [7]. Therefore, the study on the effectiveness of strategy for management college teachers in stimulating students' learning motivation becomes a hot topic at present. Information technology has brought fresh blood and brought vitality to teaching. It can abstract the obscure content to become specific and vivid; it is easy to achieve situational teaching. Information technology has played a crucial role in classroom teaching learning process. In the teaching of Management content, some teaching aids using multimedia technology can achieve a multiplier effect.

3 Study the Strategy for Teachers in Stimulating Students' Learning Motivation Based on the Quantitative Analysis

Quantitative hierarchical structure model can set the weighted values based on a large number of references and relevant experts experience, and then find a practical solution to the problem. Accordingly, it is necessary to point out the relationship between the cause of students' reluctance to study and the development of learning strategies, which is very useful for the evaluation of the practical value of the strategy. So, we establish a hierarchical mechanism model of group indicators to study the attitude adjustment strategy of students' autonomous learning in this paper based on the quantitative analysis. First, we find out the main core factors of the adjustment strategy of management college students' learning habits. We establish the hierarchical analysis model of the quantitative clustering index and perform the quantification of the elements and programs, that is, the structural relationships of the index hierarchy. According to the experience of many experts and a large number of documents as well as a scale of 1–9, we obtain the following comparison matrices (labels 1–2) (Tables 1, 2 and 3).

According to CI $= \frac{\boldsymbol{\lambda}\mathbf{max}-\boldsymbol{n}}{\boldsymbol{n}-\mathbf{1}}$, CI is determined by yielding the consistency of analysis, yielding the consistency of C is reasonable, i.e., the judgment matrices that use the above mentioned principle all pass the consistency check.

Table 1 Comparison matrices

G	x_1	x_2	x_3	x_4	x_5
x_1	*1*	*1/3*	*2*	*4*	*3*
x_2	*2*	*1*	*4*	*5*	*3*
x_3	*1/2*	*1/4*	*1*	*2*	*1/2*
x_4	*1/4*	*1/5*	*1/2*	*1*	*1/3*
x_5	*1/3*	*1/3*	*2*	*3*	*1*

Table 2 Comparison matrices

x_1	y_1	y_2	y_3	y_4
y_1	1	3	4	2
y_2	1/3	1	1	3
y_3	1/4	1	1	3
y_4	1/2	1/3	1/3	1

Table 3 Comparison matrices

x_2	y_1	y_2	y_3	y_4
y_1	1	1/3	4	5
y_2	3	1	3	4
y_3	1/4	1/3	1	1/2
y_4	1/5	1/4	2	1

$$C1 = \begin{Bmatrix} 1 & 3 & 4 & 2 \\ \frac{1}{3} & 1 & 1 & 3 \\ \frac{1}{4} & 1 & 1 & 3 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{3} & 1 \end{Bmatrix}, \quad C2 = \begin{Bmatrix} 1 & \frac{1}{3} & 4 & 5 \\ 3 & 1 & 3 & 4 \\ \frac{1}{4} & \frac{1}{3} & 1 & \frac{1}{2} \\ \frac{1}{5} & \frac{1}{4} & 2 & 1 \end{Bmatrix}, \quad C3 = \begin{Bmatrix} 1 & 3 & 5 & 6 \\ \frac{1}{3} & 1 & \frac{1}{2} & 4 \\ \frac{1}{5} & 2 & 1 & 3 \\ \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & 1 \end{Bmatrix},$$

$$C4 = \begin{Bmatrix} 1 & 2 & 4 & 4 \\ \frac{1}{2} & 1 & 2 & 3 \\ \frac{1}{4} & \frac{1}{2} & 1 & 2 \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & 1 \end{Bmatrix}, \quad C5 = \begin{Bmatrix} 1 & 3 & 3 & \frac{1}{4} \\ \frac{1}{3} & 1 & \frac{1}{3} & 4 \\ \frac{1}{3} & 3 & 1 & 3 \\ 4 & \frac{1}{4} & 3 & 1 \end{Bmatrix}$$

Test the rationality of the weight by using the consistency index:

$$\mathrm{CI} = \frac{\lambda \max - n}{n-1}, \quad \mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}}$$

The values of RI are demonstrated in Table 4 as follows.
The judgment matrix G is obtained, $\lambda^{(0)}$ max = 4.15, RI = 0.9.

$$\mathrm{CI} = \frac{4.15 - 4}{4 - 1} = 0.24$$
$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} = \frac{0.024}{0.90} = 0.027 < 0.1$$

Table 4 Values of RI

n	1	2	3	4	5	6	7	8	9	10	11
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

Since the result is less than 0.1 and it changes within the allowable range, it means that consistent test results are reasonable. The individual instrumentation from the G matrix can be considered as a weighted vector by using consistency check on the judgment matrices, we get the result from the target class to its schema class.

The results are depicted as follows:

$$i(1) = (i(1)1, i(1)2, i(1)3, i(1)4, i(1)5)$$

$$\begin{Bmatrix} 0.162 & 0.204 & 0.225 & 0.292 & 0.219 \\ 0.396 & 0.301 & 0.231 & 0.222 & 0.334 \\ 0.345 & 0.312 & 0.323 & 0.211 & 0.221 \\ 0.187 & 0.193 & 0.221 & 0.277 & 0.223 \end{Bmatrix}$$

$$i = i(1), i\ (0)$$

$$\begin{Bmatrix} 0.162 & 0.204 & 0.225 & 0.292 & 0.219 \\ 0.396 & 0.301 & 0.231 & 0.222 & 0.334 \\ 0.345 & 0.312 & 0.323 & 0.211 & 0.221 \\ 0.187 & 0.193 & 0.221 & 0.277 & 0.223 \end{Bmatrix} \begin{Bmatrix} 0.254 \\ 0.215 \\ 0.243 \\ 0.149 \\ 0.140 \end{Bmatrix} = \begin{Bmatrix} 0.303 \\ 0.234 \\ 0.239 \\ 0.224 \end{Bmatrix}$$

By establishing and analyzing a quantitative group index hierarchy structure model of strategy for management college teachers in India to stimulate students' learning motivation, we find that the relationship between the factors that affect students' independent learning and strategy making is helpful to evaluate the practical value of strategy for teachers in stimulating students' learning motivation. According to the inductive statistics, we obtain the main factors that affect the learning attitude, i.e., teachers teaching methods, multimedia information courseware, teaching case-based studies and simulation 0.254, 0.215, 0.243, 0.149, and 0.140.

Based on the weighted analysis, it is not very difficult to find the strategies for stimulating students' learning motivation, i.e., stratified teaching according to students' differences, the use of assistive technology to establish a diversified teaching form, establish a good learning environment and the relationship between teachers and students to cultivate students' learning interest, and strengthen students' practical application, with the weighted values being, respectively, 0.303, 0.234, 0.239, and 0.224.

In this paper, we test the 34 MBA students in a management institute to evaluate the effectiveness of the motivation excitation strategy. We select 314 students admitted to first year MBA program whose levels are in basic balance, and divide them into A (Slow learners) and B (Advance Learners) groups to learning under the same course. Teachers in the Slow Learners' (A) group train the students through motivational strategies, while the Advance Learners (B) group teachers train students according to a daily curriculum. Before doing this experiment, the average score of students in two groups was checked by T-test, the results are shown in Table 5.

The results showed that the mean score of group A was 2.5767, the average score of group A was 2.5767, group B is 2.3160, and T-test result is 2.334. $P = 0.480 >$

Table 5 The average score of the two groups students tested by *t*-test

	X	SD	T	P
Group A	2.5767	1.9930	2.334	0.480
Group B	2.3160	19,609		

Table 6 Average score of the two groups tested by *t*-test after the training

	X	SD	T	P
Group A	7.7582	1.3586	2.579	0.030
Group B	6.8051	1.7270		

0.05 shows that students in group A and group B were essentially the same in skills before the test. After 12 weeks of training, we re-checked the mean scores of the two groups with the T-test, the results are shown in Table 6 as follows.

Table 6 shows that the mean scores of the two groups after 12 weeks of training are: the Slow Learners (A) group is 7.7482, the Advance Learners (B) group is 6.8051, and the t-test results are 2.579. $P = 0.030 < 0.05$ means that the gap begins to appear between the Slow Learners (A) group and the Advance Leaners (B) group in the formation of. According to the deviation of the mean score, it can be said that the progress of group A students is higher than that of group B students, that is, the effectiveness of the strategy of adjusting students' independent learning attitude is relatively strong.

4 Conclusion

By dividing students into two groups for empirical research, we discuss students' performance in different periods, then observe the effectiveness of the strategies, i.e., the effectiveness of the control strategies to adjust students' independent learning attitude which is relatively strong.

Information technology and the integration of classroom teaching with computers; as a tool of the network; as a resource for activities; as a vector of the main channel of language communication [8, 9]. Effective interaction with computers arouses students' interest in learning so that students are eager to learn. The formation of learning motivation. In this learning environment, students can, according to their interests, choose learning content suitable for their level [10]. Students in this environment have the opportunity to actively participate, promote initiative with enthusiasm and creativity. It can be said that subjects in the information technology environment are the boon for the quality of education and create talents that are cultivated in the most effective way.

References

1. Gu SM (2013) Study on the factors of promoting autonomous learning in college. Shanghai International Studies University
2. Jiang BQ (2013) Computer teaching in classroom life. Mod Read 01:162–163
3. Qi C, Chen XY, Yu AZ (2014) An experimental study on the training of vocabulary learning strategies for college learners. Chengdu Normal Univ J 04
4. Wang ZM (2014) The effectiveness of strategy for college teachers to stimulate students' learning motivation. Shanghai International Studies University
5. He ZL, Fu Y, Fang FM, Min SC (2011) A study on the cultivation of autonomous learning ability of non-major college students in China. Foreign Lang China 05
6. Xu Q (2011) Research on the influencing factors of college teachers' beliefs. Huazhong Normal University
7. Xiong HF (2014) A study on the influence of reading teaching mode on students' learning motivation and achievement in senior high school cooperative learning. Huazhong Normal University
8. Wang ZM (2012) The research and practice of motivation strategies in learning. Educ Vocat 33
9. Fang XQ (2012) An overview of the study on two language learning motivation in foreign countries, vol 09. Neijiang Baike Sci. & Tech. Co
10. Gao Y (2012) Review and prospect on the study of two language motivation strategies at home and abroad. Shandong For Lang Teach J 02

IoT Based Plant Watering Mobile Robot

Jangam Akhila, Pisati Anusha, Boyapally Swathi, and Shakti Raj Chopra

Abstract Among other present natural resources, water is vital. As the population increases, there is a rising regular demand for water resources. A number of suggestions were made in order to reduce the amount of water used for irrigation. Thus, a robot-based plant spraying system has been suggested to enhance the system's performance. A system of automated irrigation that supplies water to the field in acres has been created to reduce the amount of water used by plants. By incorporating smart technologies and devices, the automation method is combined to drive devices to function autonomously and communicate, enabling them to carry out a variety of duties with no assistance from a human. With a single tap on our mobile phones, a plant watering mobile robot system can adapt to any environment and take responsibility of the plants' watering requirements, utilising the temperature and humidity sensing module (Bellingham in the role of soil moisture on our climate, weather, and global warming [1]). Wireless communication is used by the mobile robot that waters plants to communicate with the sensing module that is positioned within the plant's pot. This paper's goal is to suggest an IoT product that can water plants regularly whenever the relative humidity in the pot falls below a predetermined level.

Keywords IoT · Sensors · RFID · Wi-Fi module · Micro controller

1 Introduction

According to overall advancement of innovation throughout the recent century, such included both creation in electronics or the usage on robotics, overall quality of living for people had improved significantly. Machines have gradually taken the

J. Akhila (✉) · P. Anusha · B. Swathi · S. R. Chopra
School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, India
e-mail: akhilareddy010600@gmail.com

S. R. Chopra
e-mail: Shakti.chopra@lpu.co.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_28

position of humans in monotonous industrial tasks. A growing tendency over the past ten years has been the deployment of automated robotics to conduct home tasks. Gardening upkeep is among those most arduous duties that requires a great care and effort whenever it comes to housekeeping chores. Although moving, watering, and fertilising were all three main responsibilities which must been done, their cost and labour-intensive nature of those operations interfere with any plant owner's everyday activities. A lot of garbage contamination had also been produced as a result of inappropriate procedures.

Whenever it comes to overall landscape upkeep, irrigation is that activity which should be done more frequently in one region where a majority of waste is occurring [2].

Modern survival depends on plants in many ways. They offer a people with a variety of resources while also preserving the natural equilibrium. One's life is heavily impacted by the desire for vegetation conservation. This is a danger that the plant will lose ground richness and produce waste water if the user doesn't plant the water on a regular basis. Moreover, overwatering harms the soil.

The intelligent plant watering system is required for monitoring and control. This system incorporates a cellphone device with high and down values to regulate the water motor or automatically supplies its vegetation depending upon sensors readings. Here was a low-cost suggestion for a system that automatically waters plants while monitoring the environment's heat and the humidity content inside a plant. Because of replacing the air within that soil's permeable openings with water, excessive spraying may lead your root stems to drown. In this result, its grass will have a relatively shallow root system, which will make it vulnerable to infection or pest damage. Heavy irrigation also contributes to waste water and nitrogen pollution. Conversely, if a plant is not given enough water, the grass will become dry and yellow. As a result, with the system, watering the plant in accordance with its substrate humidity levels is essential. Hence, an irrigation solution that can effectively serve that entire area with minimal water waste is needed to increase the operation effectiveness of crop watering. This stable spray technology is typically utilised alongside ground water level monitoring capability, although this raises the system's initial cost. Bringing a freshwater container and filling it up when it's distributed will need several refill trips, and enlarging the vessel will result in unnecessarily big systems. There is a need for a more optimised solution as a result. As the crop can expand quickly, it is crucial to check their health. Many frequently neglect that drink house plant in today's busy society, which results in poor plant development and health. Establishing ideal environmental conditions for plant growth is crucial to ensure the full development of plants [3]. Almost everybody has the opportunity to provide their plant the best care possible. So, you shall automate your vegetation irrigation mechanism within your project. This pump will start if there is a liquid scarcity in the soil and continues running until there is enough moisture there.

The plant's path must be dark on a light background because LDR sensors must follow the path to the desired location. There is an RFID tag attached to each plant. It is set-up for two wireless X bee modules to communicate with one another and form a network [4].

2 Design and Development

2.1 Proposal Module

This is the proposal module using an IR sensor (Fig. 1).

3 Components

3.1 RFID Sensor

Radio frequency identification, or RFID, refers to a technology in which a reader reads digitised data that has been contained in radiofrequency chips or intelligent cards using electromagnetic waves. By allowing tags can get detected beyond direct contact on vision, and based upon its type of RFID, providing a scanning distance within a several centimetres and around 20+ m, RFID pushes auto-ID technique to a new level (Fig. 2).

3.2 IR Sensor

Its motion identification device and the infrared sensor were identical. Due to the presence of an IR LED and an IR photo diode within such sensor, a photocoupler or optocoupler could be created by fusing the two components. Stephan Boltzmann and Wein's displacement and plank emission were two physical characteristics utilised in this sensor.

According to Planck's radiation law, any item that has a temperature that is near zero releases radiation.

Law of Stephan Boltzmann: A dark body's total energy emission across all wavelengths is correlated with its absolute temperature [5].

According to Wein's displacement law, the peak wavelengths of the emission spectra from materials of differing temperatures are inversely proportional to temperature (Fig. 3).

3.3 Temperature Sensor

With ordering either record, monitor, or communicating environmental changes, a temperature sensor includes an electronic instrument which monitors overall warmth at surroundings (Fig. 4).

Fig. 1 Proposal module

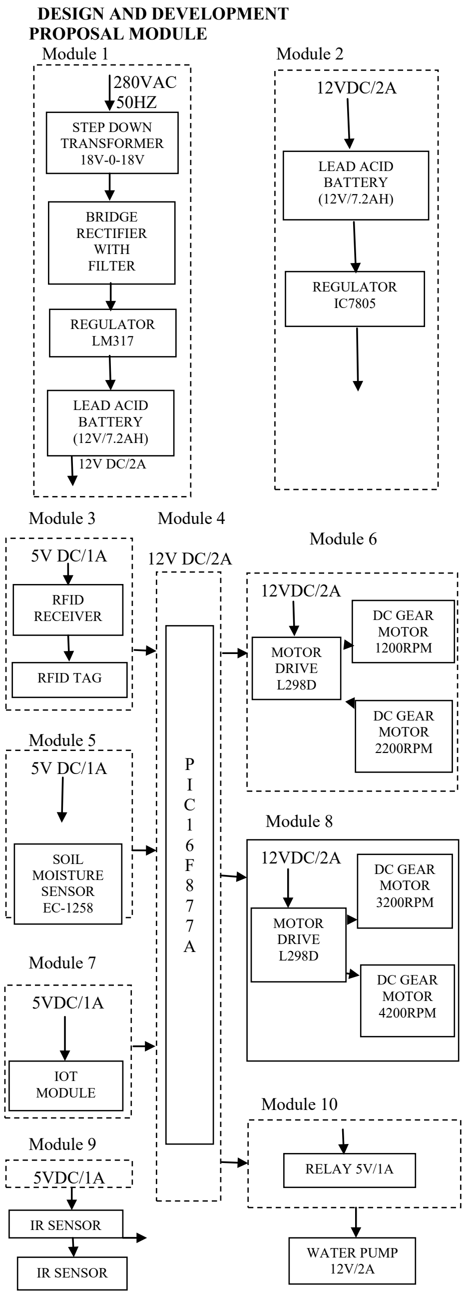


Fig. 2 RFID sensor

Fig. 3 IR sensor

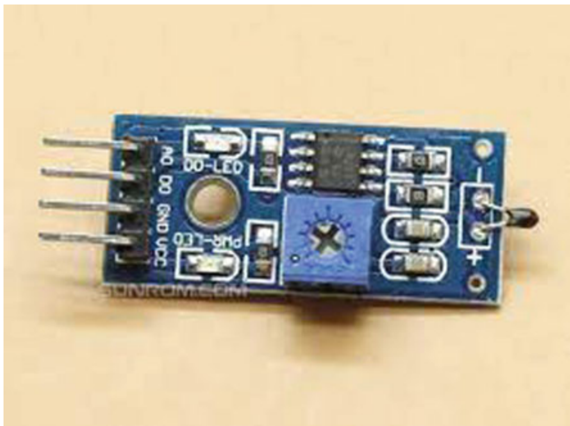


Fig. 4 Temperature sensor

Fig. 5 Soil moisture sensor

3.4 Soil Moisture Sensor

As gauge the amount of humidity within underlying soil, a ground water monitor is employed. A low level (0 V) would represent its digital input while that ground humidity measurement detected from that detector was over monitoring criterion value, while a medium range (5 V) would become an information result once this falls beneath this cutoff level. To determine whether the present topsoil water level was over a limit or not, an electronic pin is utilised to immediately detect that value (Fig. 5).

3.5 IoT Module

As such ecosystem known as the Internet of Things (IoT), things, animals, or individuals were given distinct identities and the capacity can send information across a network eliminating the need for human-to-human or human-to-computer interaction. The Internet and wireless technology (MEMS) came together to form the Internet of Things. The phrase "Internet of Things" is another name for the idea [6]. A person having a heart monitor implant, an agricultural creature having a biological tag transponder, or an automotive with built-in sensors to warn the driver when tyre pressure is low are all examples of things in the Internet of Things.

3.6 PIC16F877A

An 8-bit CMOS flash-based microcontroller is device PIC16F877A-I/P. The device is incredibly easy to use, and designing or coding it is also much simpler. Its ability to have write-erased as many instances as necessary due to the usage of flashing

Fig. 6 PIC16F877A

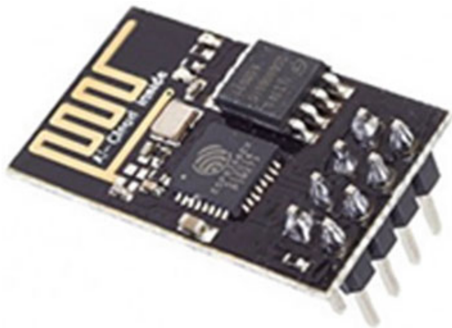


Fig. 7 ESP8266 Wi-Fi module

storage architecture is part of its primary features. They comprise of 33 input and output connections consisting of a maximum of 40 pins on it (Fig. 6).

3.7 ESP8266 Wi-Fi Module

Any micro controller can access your Wi-Fi network with the help of the ESP8266 Wi-Fi module, a self-contained system on a chip (SOS) with an integrated transmission control protocol/Internet protocol (TCP/IP) stack. The ESP8266 is capable of offloading all enabled devices connectivity tasks from the other application processor or hosting an application [7] (Fig. 7).

4 Isometric View

See Fig. 8.

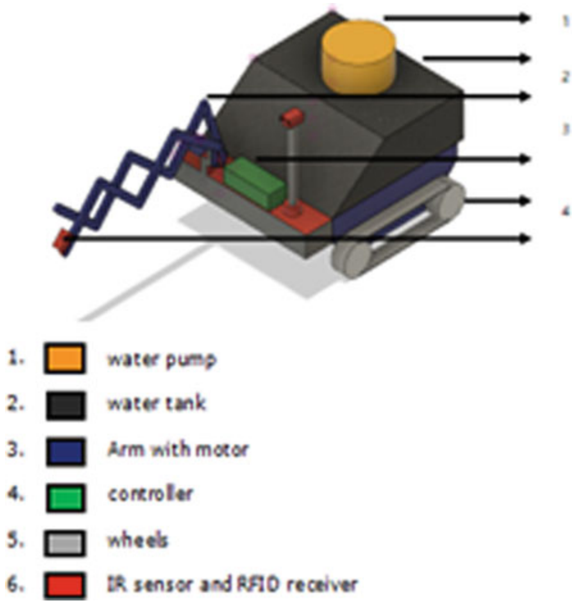


Fig. 8 Isometric view of project

5 Real View

See Fig. 9.

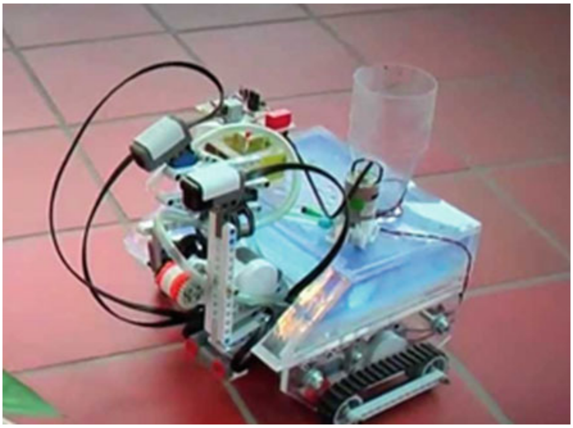


Fig. 9 Real view of project

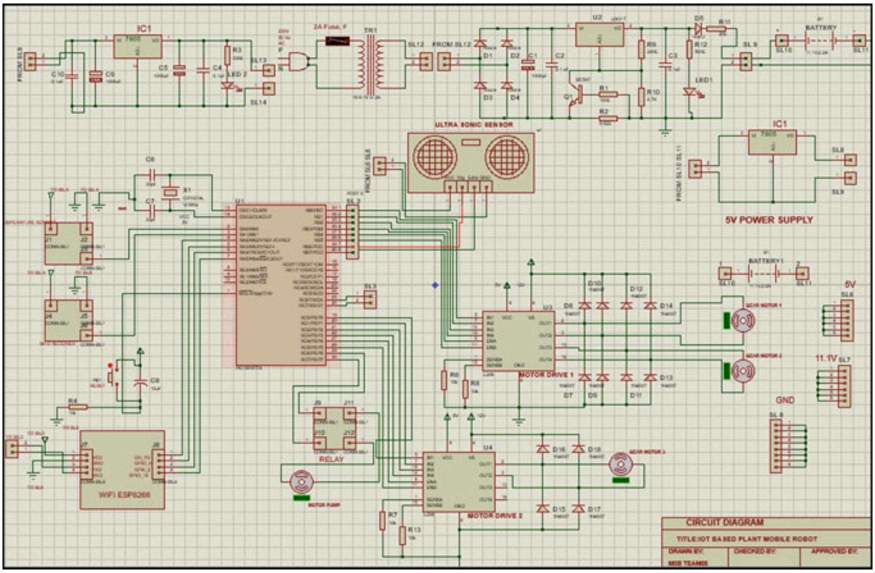


Fig. 10 Circuit diagram of hardware

6 Circuit Diagram

See Fig. 10.

7 Opportunities

It is beneficial for those who have hectic lives and neglect to maintain on their beloved plants.

Also, it is beneficial for homeowners who have lived there for a while.

It is a cheap device that is functional in daily life [8].

8 Result

As a result, the system is helpful for keeping an eye on plant characteristics including temperature, humidity, and plant development while also supplying water to that same motor pump through the IoT module. The system decreases manual labour and workforce. A microcontroller, a soil moisture sensor, an infrared sensor, and an IoT module were used to implement this arrangement. This might say page can be created to allow system control via a mobile device. An infrared sensor is included into the system to track the condition of the plants, and anyone may check on their plants online at any time and from any place. Users may obtain information about plants from IoT sensors. The mobile application platform was put to the test to see if the developed software and watering technique would work, as well as to see how well the selected sensor systems would work together to sense the surroundings for precise navigation and spraying. The plant patch was used for the tests. Based on the results of the testing, it was determined that the programme was successfully implemented as well as the system was efficiently used to water plants [9].

Acknowledgements The department of electronics and communication engineering helped fund some of this work (ECE). We take this opportunity to thank our mentor, Prof. Dr. Shakti Raj Chopra, for his insightful contributions and capable direction throughout the effort.

References

1. Bellingham K, The role of soil moisture on our climate, weather and global warming [Online]. Available FTP http://www.stevenswater.com/articles/soilandclimate.aspx
2. Johnson S, The impact of global warming on soil moisture [Online]. Available FTP; http://science.opposingviews.com/
3. Dursun M, Ozden S, A wireless application of drip irrigation automation supported by soil moisture sensors. Sci Res Essays 6(7):1573–1582
4. Chaudhary D, Nay S, Waghmare L (2011) Application of wireless sensor networks for greenhouse parameter control in precision agriculture. Int J Wirel Mob Netw (IJWMN) 3(1):140–149
5. Suba GM, Jagadeesh YM, Karthik S, Sampath ER (2015) Smart irrigation system through wireless sensor networks. ARPN J Eng Appl Sci 10(17):1
6. Darshana S, Sangavi T, Mohan S, Soundharya A, Desikan S (2015) Smart irrigation system. IOSR J Electron Commun Eng (IOSR-JECE) 10(3):32–36 (Ver. II)
7. Su SL, Singh DN, Baghini MS (2014) A critical review of soil moisture measurement. Measurement 54:92–105
8. Monitoring soil moisture for optimal crop growth. Help Desk [Online]. Available: http://observant.zendesk.com/hc/en-us/articles/208067926-Monitoring-Soil-Moisture-for-Optimal-Crop-Growth. Accessed: 13 Jun 2017
9. Widtsoe JA, McLaughlin WW (1912) The movement of water in irrigated soils. Utah Agricultural College Experiment Station, Logan

Design and Surveillance System Using Wireless Technologies

Annam Rohan Venkata Sai, Yalla Ebenezer Sunith, Annam Rohith Naga Sai, Mekala Pranav, and Shakti Raj Chopra

Abstract A huge crowd of individuals visit public places seasonally. Crowd management is vital in these things so as to avoid in order to manage these crowds to reduce the speed of pandemic and avoid spreading. This frame work would be good and efficient. The framework is comprised of three layers: detector, management, and interface layers. The sensor layer is to blame for gathering the info of range of people stepping into a area. The management section is responsible for checking whether or not the restricted count of a specific room is earned or not. The interface layer is responsible for alerting the protection officials with a buzzer to convey that the restricted count has been reached and no additional civilians are allowed within the room. This setup is applicable in closed space sort of a room, shop, room, lift, and so on. In wide area, we have a tendency to implement OpenCV ideas victimization Python to live the space between every individual and heat them if they exceed the distance limit among the crowd. The assumed distance between 2 people is calculated supported the component position of a specific person within the scene. If the limit among the individuals exceeds, then it'll show warning in browse color, otherwise they're smart to go.

Keywords Camera · Biometric · Face recognition · Ultrasonic sensors · Arduino

1 Introduction

Crowding is a common occurrence at large events such as concerts, festivals, sporting events, games, and entertainment. The analysis of crowd behavior is one of the most fascinating and active research topics in computer vision. A crowd is a group of people who have gathered in one place. Crowds differ in many ways, such as the crowd in a temple versus the crowd in a search area. A crowd is a group of people who share a common physical location. Currently, the increase in human population tends

A. R. V. Sai (✉) · Y. E. Sunith · A. R. N. Sai · M. Pranav · S. R. Chopra
School of Electronics and Electrical Engineering, Lovely Professional University Phagwara, Phagwara, India
e-mail: rohanvsai@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_29

to crowd public areas [1]. As a result, it is necessary to investigate the surveillance system with many circuit tv is used to observe the group. The human eye cannot see all of the cameras at the same time. As a result, an automatic technique should be used be used to keep an eye on the crowd indefinitely. The difficulties in mechanically detecting the required events are the simultaneous occurrence of several events, the large amount of information that must be processed, occlusions, and real-time detection. The proposed method is limited to a small set of objects. The Internet of things is a powerful industrial system made up of radio frequency identification and wireless devices that can transfer data over a network without requiring human interaction. The analysis of 13 crowd behaviors using surveillance videos is an important issue for the general public. Safety because it allows for the detection of potentially dangerous crowds and where they are headed. People counting, people tracking, and crowd behavior analysis are the three types of crowd analysis algorithms based on computer vision [2–4]. The Internet of Things is made up of three layers: the sensing layer, which collects data from the real world using existing hardware such as sensors, the network layer, which transfers the data collected over a wired or wireless network, and the application layer. The application layer is in charge of the user's and the system's two-way communication. In a variety of fields, IoT applications are rapidly evolving and expanding. To alleviate overcrowding in enclosed spaces such as rooms, conference halls, and so on, we used IOT to detect crowds, whereas computer vision technology was used to identify overcrowded areas in open spaces such as malls, parks, and public areas. The goal of this project is to detect and respond to overcrowding issues as early as possible [5]. This will allow us to assess how effective our crowd safety precautions are, and the considerations that are taken into account are as follows: monitoring the overall number of people to ensure the venue's safety, monitoring the People are distributed to help to prevent local overpopulation, and potential crowd problems are identified to prevent public outrage from escalating. With a growing population and a wide range of crowds connexon at common places for meetings, gatherings, and reunions, there is an increased demand for these services. Moreover, during this pandemic global crisis, the importance of these types of projects has created a necessary desire among governments from various countries to watch the general public crowd [6, 7].

1.1 Literature Survey

Smartphones with sensors are now being used as tags by researchers for extensive human sensing. More individuals will have the ability to be recorded in the future without supplying any tags as a result of the widespread usage of smartphones. A client's smartphone must have an application installed in order to make use of certain smartphone-based location tracking systems. Using the GPS sensor on the smartphone, the installed application transmits the location to a distant server over an Internet connection. Nonetheless, it is improbable that consumers in crowded areas and far-off places will always have access to the Internet. The variety of operating systems for smartphones and models makes it challenging to create and distribute applications [8, 9]. The majority of current systems use a two-stage procedure to intercept the MAC id; the very first stage is to capture the probes, and the second is to analyze them. As a result, a more flexible, workable, and effective approach for leveraging wireless technology to monitor a large number of people is required. It detects MAC identifier from probe queries sent by wireless connections made by smartphones used by the user. The person-tracking smartphone sensor does not require the installation of whatever additional equipment infrastructure or changes to the device's architecture [5]. There are different sorts of crowd counting in video sensing:

LOI counts can be done through feature tracking methods, where the features are either turned into paths and their trajectories are combined to form object tracks, or it can be done by counting the number of people in each video temporal frame [10].

One of the intriguing and active fields of computer vision research at the moment is the study of crowd behavior. A group is a collection of people who have collected together in one location. Crowds vary based on the circumstance; for instance, a large crowd at a mall will be different from one in a temple. A public is a group of people who congregate in the same area. Currently, crowding in public spaces increases as the human population rises. In order to view the crowd, it is important to analyze the surveillance system's extensive closed-circuit television (CCTV) camera array. The cameras cannot all be seen at once by the human eye. As a result, it is necessary to continuously watch the audience for a long time using an automatically generated technology [11]. Throughout the past ten years, there has been a significant rise in cell phone usage. The percentage of cellular telephone subscriptions was 97%, or 7084 million subscribers globally, according to the 2015 ITU World Telecommunications report. The technological advantages offered by mobile devices also reflect this rise in subscribers. Modern mobile devices also feature a lot of embedded gadgets, several communication methods, and great processing power (GPS, gyroscope, accelerometer, microphone, and webcam, among others). Due to this improvement in technology and the rising number of subscribers, several different apps using cellphones as sensors have been developed by the developer and research communities [6]. It has been a fascinating field of study due of the challenges and importance of numerous trackers. That is the main objective of the surveillance system and video understanding. Deep learning's exceptional abilities in object detection have led directly to the recent emergence of tracking through detection systems. Although many people or pedestrian tracking are benefited by this development, it is still a difficult task because not all issues have been overcome. Occlusions, in addition to changes in posture, size, and frame, are instances of these challenges. Results from earlier research utilizing fully convolutional network (CNN) features for multiple person monitoring were encouraging. Leal-Taixe, Ferrer, and Schindler1 employed Siamese CNN, a modified numerical approach, and a slope boosting classifier on the MOT Contest dataset2. Zhu, Li3, and Porikli employed a PETS5 dataset based on regions, and CNN uses the SSVM and RCNN4 classification techniques that are more rapid. In the past, faster-RCNN features were thought to be the most effective for recognizing objects because they were the main technique employed by the highest-scoring object detector in the Large Scale For Challenge 6 (ILSVRC) 2016 [2]. In Wuhan, China, a pandemic of respiratory distress sickness was discovered to be caused by a brand-new coronavirus called as SARS-CoV-2 in late 2019. The condition was referred to as COVID-19, which was actually short for coronavirus illness 2019, by the World Health Organization in February 2020. Acute respiratory distress syndrome, sepsis, and multi-organ failure are all possible outcomes of 2019-nCoV infection, all of which can be fatal [1]. The WHO first referred to the COVID-19 infection as a public health issue of international concern on January 30, 2020, and then, in March 2020, it started using the term “disease outbreak” to emphasize the seriousness of the situation and encourage all nations to take measures to detect infection [12]. The use and appeal of digital technologies has been a striking feature of the COVID-19 deadly virus, fusing what these advancements may possibly offer in terms of effective communication and public wellbeing reactions to support individuals and communities cope with and comprise the crisis, on the one hand, and expanding sources for social events, presence, having significance, going to persist with one another and reconfiguring rituals, and summoning with the deep affect. Due to the increased usage of Internet, smartphones, social networks, information, and artificial intelligence (AI). As the pandemic began, corresponding technology already were widely, if unevenly, dispersed around the world, particularly in middle-income and high-income states, resulting to a significant reliance on electronic technology—in terms of the amount of digitalization allowed. Governments in a number of countries also used the occasion to call for the progress of digitization, particularly among groups and demographics where uptake and inclusion have been low. This was due to technology, literacy and education, information, affordability, and other considerations [7]. In December 2019, symptoms of a sickness mimicking pneumonia first appeared in Wuhan Town, the capital of Hubei Province, China. The research revealed that the instances were brought on by a form of coronavirus that had not yet been recognized. This virus strain is also known as COVID-19, or Coronavirus 2019, since it first surfaced in 2019. The virus is thought to have originated in the Huanan seafood sector in Wuhan, China. Although the molecular mechanisms of COVID-19 transmission from humans to humans is still unknown, the general concept of pulmonary illness transmission is similar. Droplet dispersion is a method of respiratory disease transmission. Sneezing or coughing when unwell can cause the microorganism to spread to others. As an alternative, environmental variables greatly aid in the propagation of this illness [1]. Given that one person might be hidden in a variety of ways, people detection and identification is an essential field of research in terms of identification, monitoring, and activity recognition. The segmentation of body components in a large analysis (for use in gesture or computer interaction, for example). People can be addressed with group behavior once they can be a component of complex organizations like groups or crowds [9].

2 Methodology

Module 1: Developing an IOT device for monitoring the crowd in closed space. This is a hardware kind of approach where we have used an Arduino as main board and sensors to implement it; it also contains a buzzer to alarm indicating crowd violation.

Module 2: Developing an image processing model for monitoring crowd in open space. This is a software kind of approach where we have used computer vision to identity crowds; it requires a minimum specification computer to run the proposed model; the description and required specification for the computer are already discussed in the hardware component section [3].

2.1 IOT Method

This method is a hardware-oriented approach that is used in closed spaces such as rooms, conference halls, and others. We used ultrasonic sensors to count the total number of people entering the room, and the total is sent to Arduino to process where it checks for the desired count as and when people enter the room. If the restricted the percentage of individuals in a given room is reached [11], a buzzer alerts that the total admissible count of the specific room has been reached and no more people should be allowed inside the room to avoid overcrowding (Fig. 1). The restricted count is set up using the construction measurements of the room. The Internet of Things' most recent development effects has governed a variety of surveillance operations, particularly crowd monitoring. Over time, the management and coverage of crowds at public gatherings is made possible by surveillance systems with Internet of things bias and detectors. Since it is used so frequently in security and monitoring operations, group surveillance has recently lost some of its appeal in smart cities. In the disciplines of deep literacy, machine literacy, and computer vision, it has gained significance. There have been many different methods developed for automatically observing people in a variety of settings, including malls, train stations, airports, interior spaces, and other public places. It behaves on a huge scale, much like crowd monitoring, group evaluation, irregular exhaustion detection, door security systems, hand workout having to cover, and other similar things. The same goes for a people's length, measure, body articulation, and disguise, these operations are considered a difficult task for experimenters (Figs. 2, 3, 4 and 5).

2.2 Image Processing Method

This method is a software-oriented approach that is used in open spaces such as malls, public areas, and others. We used computer vision technology to detect people with the help of cameras in public areas, and the input to our model is directly taken from

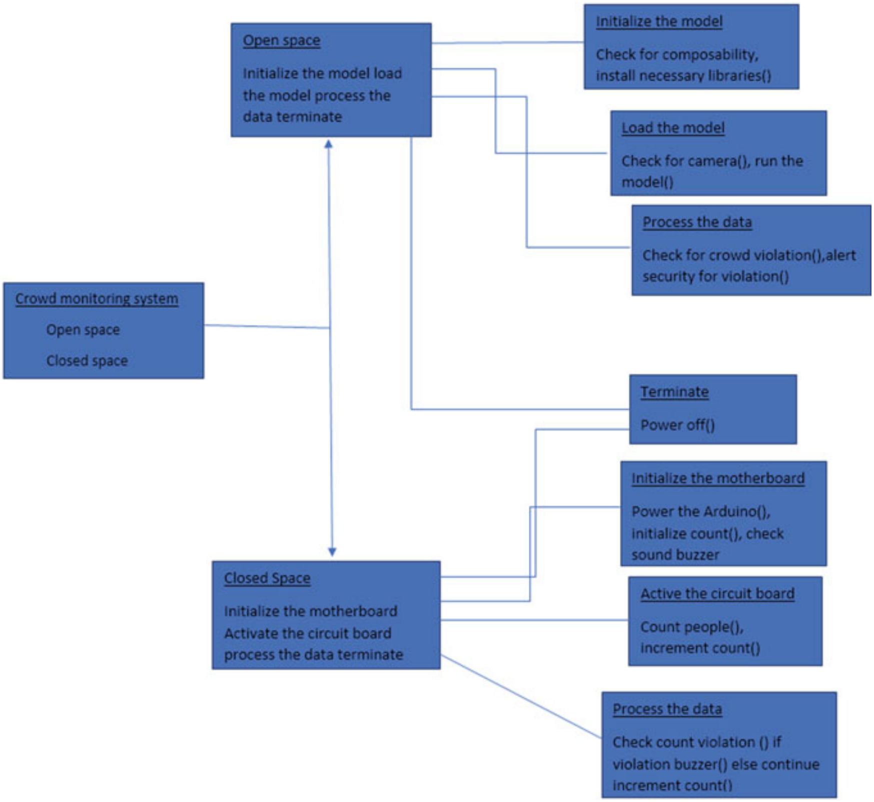


Fig. 1 Architecture diagram

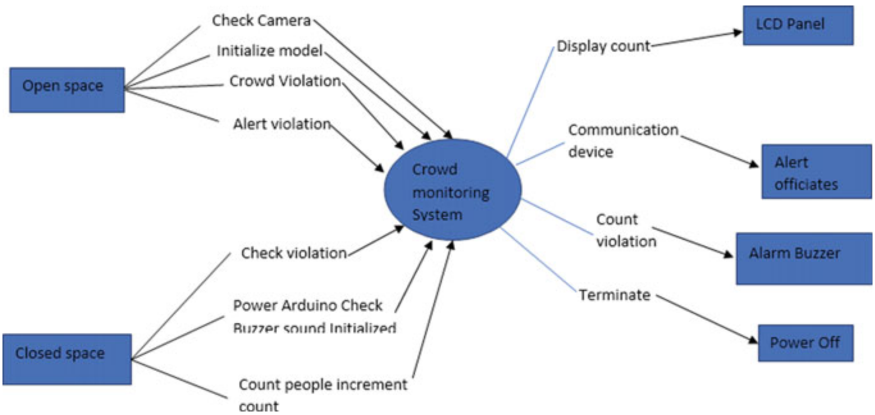


Fig. 2 Data flow diagram

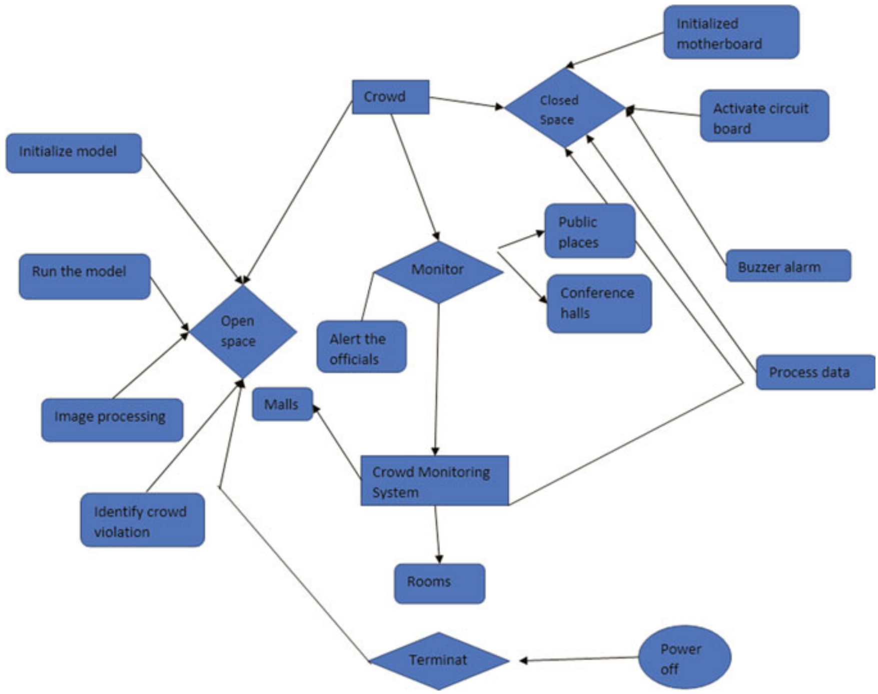


Fig. 3 Er diagram

the real-time monitoring cameras installed in public areas where the model will be installed in the system, and the camera v positions of the people in the video (Figs. 5 and 6).

2.2.1 Methods Based on Detection

In this instance, a moving glass door detector is used to recognize and the number of individuals in an image. The detection techniques used demand the use of trained classifiers that can extract basic features. The major part of the objects in crowded images are hidden by the background, despite the fact that these methods are good at detecting faces.

2.2.2 Regression-Based Techniques

We were not able to extract reduced features using the method outlined above. Here, regression-based techniques excel. We first obtain the low-level information for each patch before cropping patches from the original image [13].

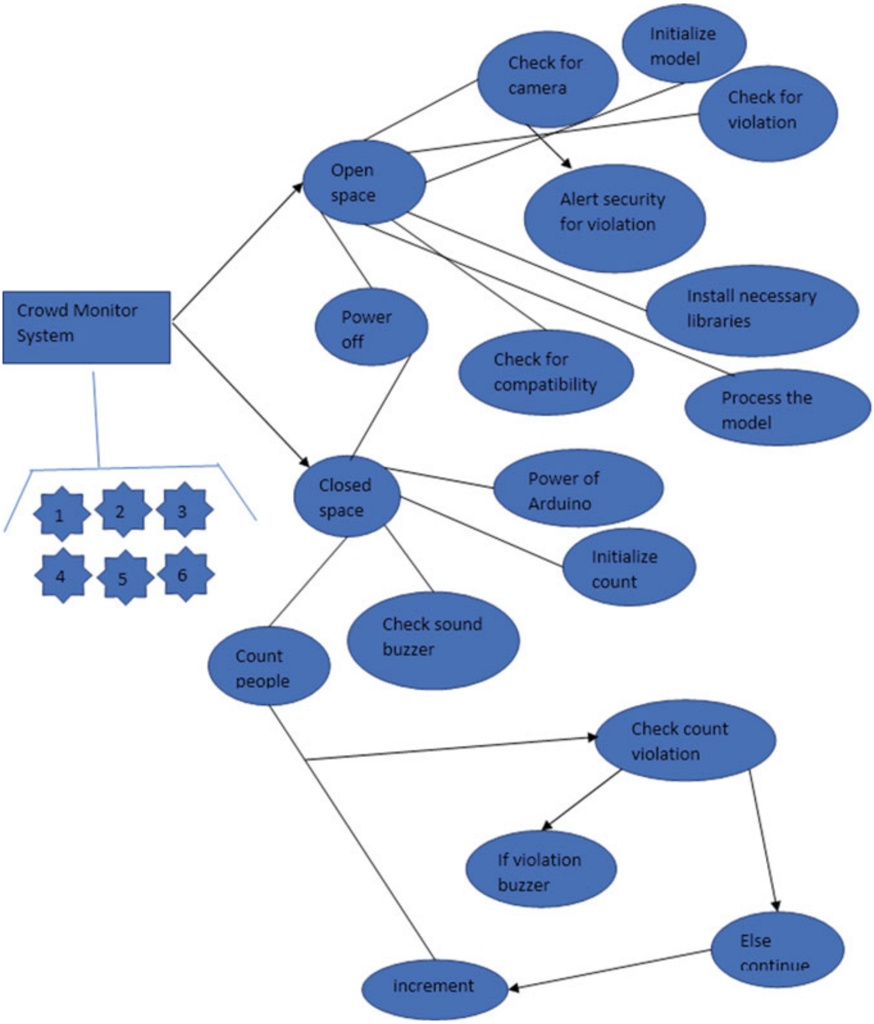


Fig. 4 Use case diagram

2.2.3 Density-Based Estimation Techniques

Making a mapping population of the objects is the first step. After that, the algorithm discovers a linear mapping among extracted Object density maps that correspond to features. Nonlinear mapping can also be learned using random forest regression (Fig. 2).

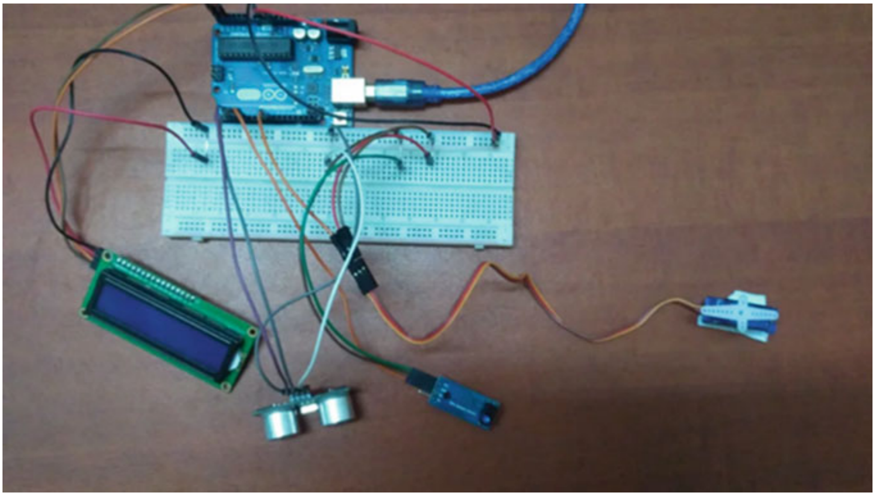


Fig. 5 Input and output

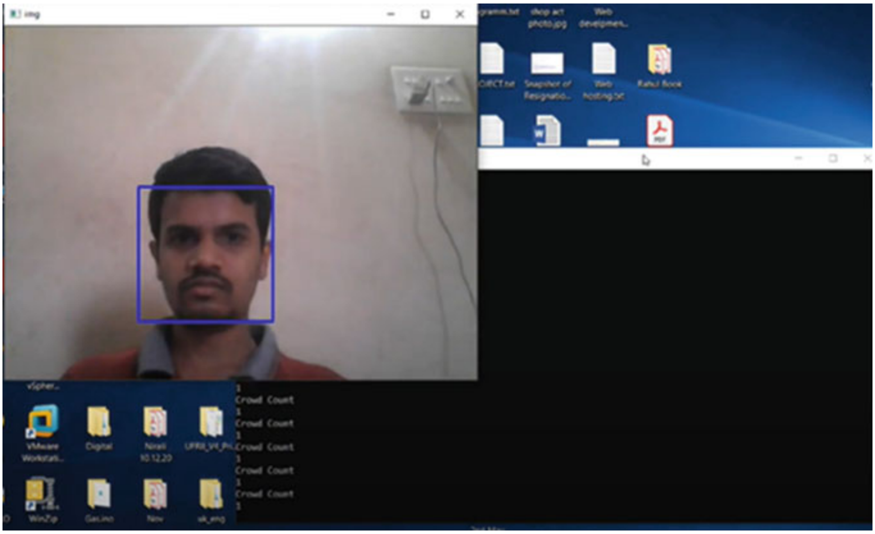


Fig. 6 Input and output

2.3 Algorithm

Step 1: Install the most recent version of the Arduino IDE.

Step 2: Write the code, extract it, and install it in Arduino Uno.

Step 3: Connect the (male/female) wires of the ultrasonic sensor, LCD display, servo motor, and IR tracking sensor.

Step 4: Now run the program that we installed in Arduino Uno first.

Step 5: There will be a one-fold increase in the count variable's value. The count variable indicates how many people are in a building or car at any given time.

Step 6: The LCD display module suggests that no one is allowed to enter and that the room will remain closed until someone exits the building when the cost in the count variable exceeds the maximum occupancy value.

3 Testing

The unit testing is done on major important components of the proposed model in our IOT design, working of it is checked and verified then the sensors where its working is tested and verified then the battery to power up Arduino and sensor is tested and verified, now the components related to software working of the model are checked and verified the camera used is tested for its working and verified then here to run the Python code for crowd detection we have used Python which is executed in the local machine to which the CCTV camera is connected so its setup is verified.

The functioning for which the components are being used is tested in this phase whether the desired result is produced by the components. The sensor is tested along with the Arduino whether it counts the people correctly and buzzers the alarm when the restricted count is reached. The Python code is executed on the input video from the CCTV camera to check whether it identifies crowd in public places and distinguishes them with green and blue rectangular boxes.

The combined working of the components is tested and verified in this testing. The Arduino along with the sensor and battery is tested for its working collaboratively with each other getting attached. The camera and the code are tested along together whether the image from camera is used correctly for detecting crowd.

For checking whether the data is available in the desired folder we need to specify the correct path for it to detect the crowd from the CCTV camera available in public places and malls for that we need to specify the inputs for it assume the expected results you need the program to produce should be done beforehand for better results. This testing is done without the need of developer, the person who doesn't know anything about developing he just checks whether the expected outputs are produced for the given inputs. For this test give a few minutes recorded video as input and test whether the program identifies the crowds in the video and alerts about it.

4 Conclusion

The suggested solution helps to manage crowds in both public and private spaces, which has significant social benefits. Because high-quality camera sensors are affordable, people can take advantage of this transformation at a very low cost. The literature review for our system made note of the industrial limitations as well. The findings demonstrate that the suggested system has a favorable marginal impact because it can be used on any platform, producing a favorable increase in entrepreneurial growth.

References

1. Atalan A (2020) Is the lockdown important to prevent the COVID-19 pandemic? Effects on psychology, environment and economy-perspective. Ann Med Surg 56:38–42
2. Chahyati D, Fanany MI, Arymurthy AM (2017) Tracking people by detection using CNN features. Procedia Comput Sci 124:167–172
3. Chopra SR, Teja T, Gowtham R, Kumar GD, Vivek TS, Venkateswarlu K (2023) Moving object detection using satellite navigation system. In: Artificial intelligence and machine learning in satellite data processing and services: Proceedings of the international conference on small satellites, ICSS 2022, Springer Nature Singapore, Singapore, Jan 2023, pp 99–108
4. Chopra SR., Kumar S, Kumar N, Balakrishna S, Jha A (2022) Smart highway navigation system with toll collector using Li-Fi. In: AIP conference proceedings. AIP Publishing LLC 2421(1):050004, Jan 2022
5. Duives DC, van Oijen T, Serge P (2020) Enhancing crowd monitoring system functionality through data fusion: Estimating flow rate from wi-fi traces and automated counting system data. Sensors 20(21):6032
6. Facts ICT (2015) Figures-the world in 2015. The International Telecommunication Union (ITU), Geneva
7. Goggin G (2020) <? covid19?> COVID-19 apps in Singapore and Australia: Reimagining healthy nations with digital technology. Media Int Aust 177(1):61–75
8. Hashish S, Ahmed M (2015) Efficient wireless sensor network rings overlay for crowd management in Arafat area of Makkah. In: 2015 IEEE international conference on signal processing, informatics, communication and energy systems (SPICES), IEEE, pp 1–6
9. Junior JCSJ, Musse SR, Jung CR (2010) Crowd analysis using computer vision techniques. IEEE Signal Process Mag 27(5):66–77
10. Kowcika A, Sridhar S (2016) A literature study on crowd (people) counting with the help of surveillance videos. Int J Innov Technol Res, pp 2353–2361
11. Sowri VO, Chandra BE, MoniSwahith K, Sarvani PV, Chopra SR (2023)Detection of automobile accidents through satellite navigation system. In: Artificial intelligence and machine learning in satellite data processing and services: Proceedings of the international conference on small satellites, ICSS 2022, Springer Nature Singapore, Singapore, pp 31–42
12. Wei WE, Li Z, Chiew CJ, Yong SE, Toh MP, Lee VJ (2020) Presymptomatic transmission of SARS-CoV-2—Singapore, January 23–March 16, 2020. Morb Mortal Wkly Rep 69(14):411
13. Chopra SR, Teja T, Gowtham R, Kumar GD, Vivek TS, Venkateswarlu K (2023) Moving object detection using satellite navigation system. In: Artificial intelligence and machine learning in satellite data processing and services: Proceedings of the international conference on small satellites, ICSS 2022, Springer Nature Singapore, Singapore, pp 99–108

Classification of Minor Roads Inside Dense Informal Settlements of Mumbai Using Machine Learning

Vaibhav Kumar and Vivek Kumar Singh

Abstract The paper discusses the issue of haphazard development in Indian cities, leading to the rise of unplanned informal settlements and narrow roads, which makes it difficult for essential services to be delivered to residents. The lack of accurate network information on minor roads inside informal settlements further exacerbates this problem. The paper proposes the use of a deep neural network (DNN) model to classify urban features, including minor roads inside dense informal settlements of Mumbai city, using high-resolution satellite imagery. The proposed model achieved an accuracy of around 95% in classifying urban features and 85% in classifying minor roads inside informal settlements. The outcomes presented in the paper have significant implications for sustainable development goals, as they can support effective resource planning to cater to marginalized communities living in slums. Accurately mapping minor roads and urban features can help improve the delivery of essential services such as emergency support, healthcare, and sanitation to these communities. The proposed approach can be replicated in other cities facing similar challenges, contributing to marginalized communities' overall development and well-being.

Keywords Urbanization · Machine learning · GIS · Minor road classification · Informal settlements

1 Introduction

The rapid urbanization in Indian cities has led to the growth of informal settlements or slums that are characterized by highly dense and narrow roads. These roads play a critical role in the delivery of essential services and in connecting the slums with the

V. Kumar · V. K. Singh (✉)
Department of Data Science and Engineering, Indian Institute of Science Education and Research, Bhopal 462066, India
e-mail: december.keviv@gmail.com

V. K. Singh
Center of Environmental Sensing and Modeling, Singapore-MIT Alliance for Research and Technology, Singapore 138602, Singapore

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_30

surrounding neighborhoods [1]. However, developing a high-resolution geodatabase of these roads and their attributes is a resource-intensive task that is typically carried out through on-ground physical surveys.

Geospatial artificial intelligence (GeoAI) techniques can help in the development of such a database by applying classification algorithms to high-resolution Earth observation (EO) imageries. However, extracting or classifying aerial or satellite image pixels into semantic objects of the dense urban environment, such as minor roads in informal settlements, is a challenging research problem due to the heterogeneous regions with significant class variations and similar spectral signatures. Figure 1 illustrates the complex nature of informal settlements with features such as narrow roads, slum buildings, and vegetation. Overcoming these challenges and developing accurate geodatabases can significantly aid in better planning and service delivery to these informal settlements.

In recent years, a plethora of research has focused on using very high-resolution satellite imagery to identify slum areas. These studies have proposed various machine learning (ML) and deep learning (DL) models to classify urban objects accurately. Using machine learning techniques, researchers have implemented these models for identifying residential neighborhoods, segmenting informal settlements, and mapping utility services such as toilets and restrooms. The deep learning approach consists of using convolutional neural networks (CNN) such as UNet++, regular residual networks [2–5], and fully convolutional neural networks (FCN) [6]. Despite existing research on road extraction, minor road extraction inside informal settlements remains a research gap that has not been attempted.

The current study aims to fill this research gap by classifying minor roads in informal settlements and other urban objects in different regions of Mumbai city. The study proposes the following contribution:

- The study proposes developing a deep neural network (DNN) model that uses high-resolution satellite data to classify minor roads in urban areas. The aim is to fill the research void in minor road extraction inside informal settlements, which has not been attempted before.
- The study also intends to create a high-resolution land use land cover (LULC) map with a minor road class for both formal and informal settlement areas of Mumbai

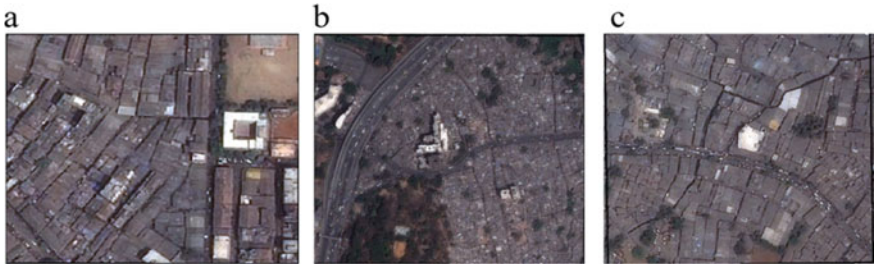


Fig. 1 Dense informal settlements with narrow lanes and roads

city. The outcomes of the study can be helpful in governance and decision-making processes in transport planning and urban settlement policies.

- The findings can also provide valuable insights into the spatial distribution and characteristics of minor roads in urban areas, which can be used to inform urban planning and development policies.

2 Methodology

2.1 Study Area

Mumbai is one of the metro cities in the western coastal ghat of India. It was estimated that more than 40% of the population of Mumbai city lives in the slums. The city is located between 18° 53′ N latitude to 19° 160′ N latitude and 72° E longitude to 72° 59′ E longitude [7].

As a result of the city's high population density, there has been a significant change in land use over the years. According to Sahana et al. [7], the built-up density has increased from 55 to 71% between 1990 and 2015, while green spaces, such as vegetation and urban forests, have decreased from 14 to 5%. To classify minor roads, this study utilized a Pleiades-1A satellite with a spatial resolution of 0.5 m to generate training and testing datasets and implement a DNN model in six different image patches, as depicted in Fig. 2.

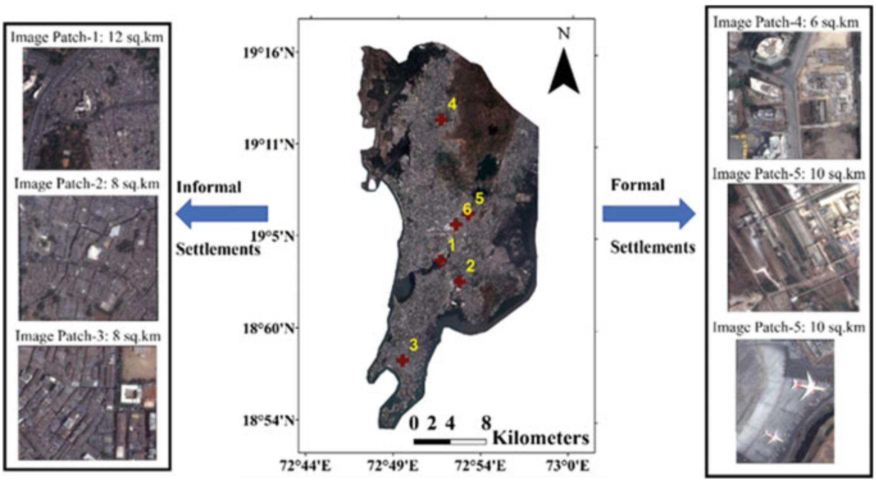


Fig. 2 Mumbai city and selected image patches with area to demonstrate model performance

2.2 Training Dataset Creation

The process of creating training samples involved manually annotating a total of 17,919 ground truth locations by zooming in on different point locations using geographic information systems (GIS). These samples were categorized into seven classes: water, barren, built-up, main road, vegetation, minor road, and other (including vehicle and aircraft). Figure 3 illustrates the process of preparing training data. The annotated points were processed to obtain their coordinates, and a 6 × 6 window was applied to select neighboring points of the same class, resulting in a total of 645,095 annotated samples. A similar approach was used to generate a total of 129,019 test samples. Table 1 provides an overview of the total number of training and test dataset samples collected for each class.

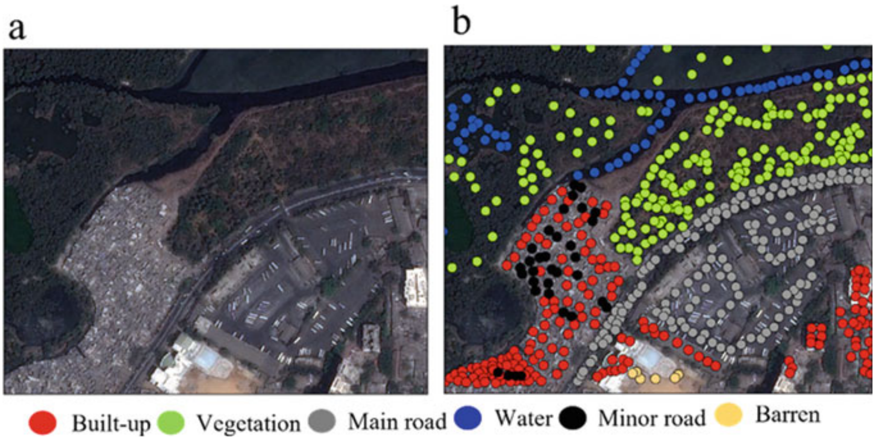


Fig. 3 **a** Image **b** collected training data samples

Table 1 Number of training and test data

LULC category	Class	Training samples	Test samples
1	Water	75,560	15,112
2	Barren	89,215	17,843
3	Built-up	155,875	31,175
4	Main road	120,445	24,089
5	Vegetation	86,530	17,306
6	Minor road	110,915	22,183
7	Other	6555	1311
	Total	645,095	129,019

2.3 Model Training and Predictions

The DNN comprises the input layer, hidden layers, and output layer. The term "deep" indicates a series of layers that transform the data at different levels and generate a multilayered representation of data patterns for detection and classification. The input variables are provided as binary signals and processed by individual neurons in the first layer of the input layer. The second set of layers is known as the hidden layers, and the third and final layer is the output layer [8]. Therefore, several hidden layers are defined through a trial-and-error process, along with some prior knowledge of the dataset. The model architecture includes perceptrons with activation functions like relu and softmax applied in various layers to normalize the output. The back-propagation algorithm is used to train the model by minimizing the error function using gradient descent. The dataset is randomly divided into training and validation sets, and hyperparameters are selected to optimize model performance. Complete detail of the model configuration is mentioned in Table 2.

After training the model, it is tested on unseen test data using various performance metrics such as overall accuracy (OA), per class accuracy (PCA), Kappa index (KI), and precision index (PI). These metrics help to evaluate the model's performance and determine if it is accurate and reliable for predicting outcomes.

Table 2 Model configuration

Parameter	Proposed model
Batch size	200
No. of input layers	1
No. of neurons (n) in input layer	3
No. of hidden layers	6
No. of neurons (n) in hidden layer	Hidden layer 1: 256 → hidden layer 2: 224 → hidden layer 3: 128 → hidden layer 4:128 → hidden layer 5: 64 → hidden layer 6: 32
No. of neurons (n) in output layer	6
Activation function	Relu: for hidden layers, Softmax: for output layer
Loss method	Sparse_crossentropy
Optimizer	Adam
Regularization	0.01
Dropout	0.3

3 Results

The DNN model was trained on a personal computer with an Intel(R) Core(TM) i7-8750H CPU, 12 GB of RAM, and an x64-based processor. The model achieved an overall accuracy of approximately 95% after 200 training epochs for both the training and validation sets, as depicted in Fig. 4. This high accuracy indicates that the model is neither overfit nor underfit and performs well in dense regions. This outcome is particularly notable as it has not been reported in prior research for developing countries cities.

After validating the model, it was applied to the test dataset, and the results are presented in Table 3, which shows the confusion matrix for the model implementation. The model correctly classified 28,285 dataset points as urban and built-up cover. However, there were some misclassifications, where 1611 urban points were classified as main roads, and 841 urban points were classified as minor roads. The similarity in spectral signature between urban areas and roads and the presence of vegetation near buildings and roads may have contributed to these misclassifications. Similarly, vegetation near water bodies may have led to misclassification of water bodies as vegetation. Despite these misclassifications, the large dataset used for the classification allowed the model to perform well overall, achieving an OA of 91.88% and KI of 0.90. The PI was also found to be 0.96, indicating that the model is reliable for predicting land use and land cover classes.

The developed model was applied to different regions of Mumbai to classify features in informal settlements. The results of the model's prediction are shown in Fig. 5a–c. It appears that the model accurately classified features such as minor roads, vegetation, vehicles, and slum buildings. This is a significant contribution to the field as such high accuracy in land use and land cover classification has not been reported previously. The classified map of informal settlements can be used to develop an automated geodatabase, which can save resources in terms of cost, money, and time and can be a valuable input in decision-making systems. Furthermore, the model's performance in classifying features in formal settlements was analyzed (as shown

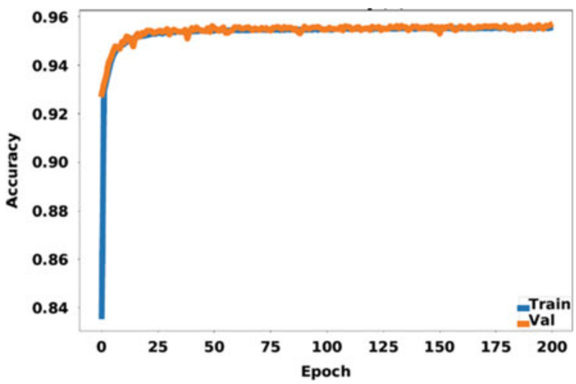


Fig. 4 Model accuracy with epoch for training dataset

Table 3 Confusion matrix and accuracy statistics for test dataset

LULC	Water	Barren	Urban and built-up land	Main road	Vegetation	Minor road	Other	User accuracy (%)
Water	14,948	0	0	54	0	110	0	98.91
Barren	0	17,754	79	0	10	0	0	99.50
Built-up	114	134	28,285	1611	80	841	111	90.73
Main road	183	0	1564	22,057	151	134	0	91.56
Vegetation	35	0	95	478	16,468	230	0	95.16
Minor road	123	0	2374	223	431	19,032	0	85.80
Other	0	0	111	0	0	0	1200	91.53
Producer accuracy (%)	97.05	99.25	87.31	90.31	96.08	93.54	91.60	
OA	91.88%							
KI	0.90							
PI	0.96							

in Fig. 2), and the results are presented in Fig. 6a–c. The model was observed to be highly efficient in classifying features in the selected formal region, which contained residential, commercial, industrial, utilities, and mixed built-up areas.

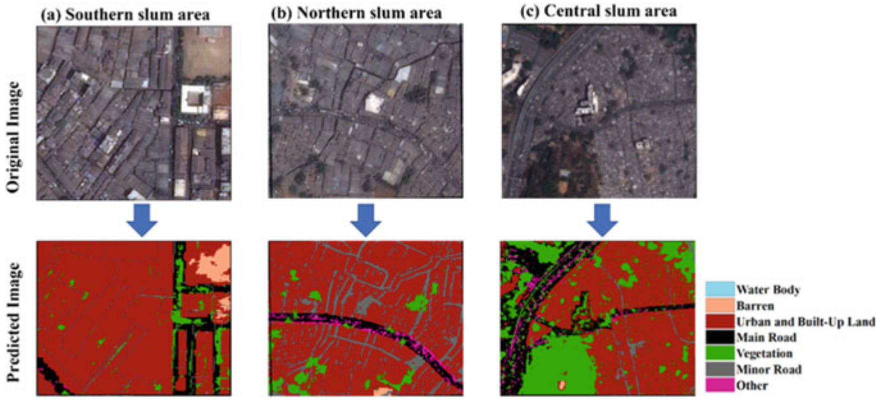


Fig. 5 Different areas of the informal settlement in the Mumbai area

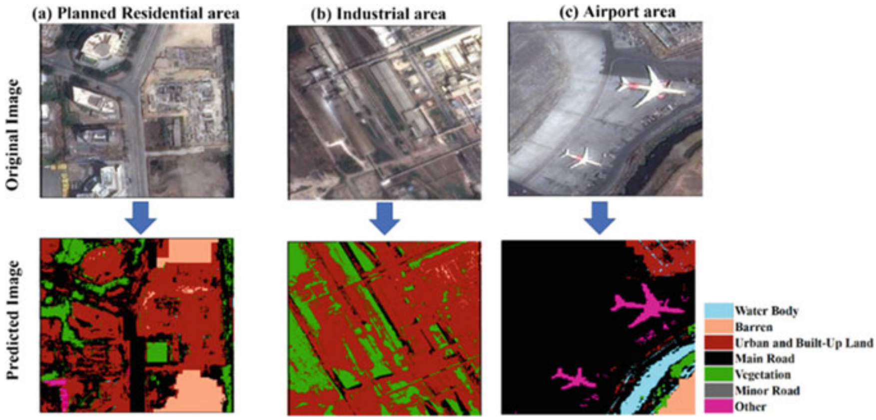


Fig. 6 Different areas of the formal settlement in the Mumbai area

4 Discussions

The study's primary objective was to classify minor roads in Mumbai using machine learning techniques to automate the task of creating a high-resolution geodatabase of the roads. The study utilized the DNN model to classify various urban objects, including water, barren land, urban and built-up land, main road, vegetation, and minor road, in both formal and informal settlements of Mumbai. The study achieved an overall accuracy of 91.88%, and the accuracy of classifying minor roads, which was the focus of the research, was about 85%. The model could classify minor roads accurately, even if dense buildings and vegetation surrounded them. The study suggests that incorporating high-resolution elevation details can further improve the accuracy of the classification process and reduce the intermixing of roads and building classes. The generated outcomes of the study can be used to derive various policy instruments such as land-use planning, transportation planning, infrastructure development, and disaster management as follows:

- *Quantification of vehicle accessibility in dense regions*: The study's outcomes can serve as a useful decision-making tool for evaluating vehicle accessibility in dense regions, such as assessing accessibility to services like electricity, street lighting, emergency services, and healthcare.
- *For planning commercial and recreational activities*: A high-resolution LULC map, combined with a detailed geodatabase of roads, can assist relevant agencies in identifying suitable locations to serve potential customers in informal settlements. These settlements house various economic activities, including small-scale manufacturing, hawking, vending, schooling, waste recycling, and cultural activities like processions, gatherings, and performances on public streets [9].

- *Accessibility-driven development*: The development policies should prioritize utility service accessibility for every citizen, making it a metric in land use planning. To achieve this, built-up areas should be designed to have broader roads. A fire-driven land-use planning framework proposed by Kumar et al. [10] optimizes the built-up compositions at the microscopic level, catering to twice the population while reducing fire susceptibility. In resource-constrained urban areas such as slums, the maximum coverage variability of any vehicle is only 9%, as explained in a study by Kumar et al. [11]. Therefore, extracting minor roads in slum areas becomes crucial for creating a geodatabase, which is eventually employed to quantify various service deliveries in those areas.
- *Integration with urban development plans*: The integration of high-resolution LULC maps and geodatabases of minor roads can also be used in urban development plans to effectively map the presence and expansion of minor roads and their empirical associations with sustainability drivers such as population density, health, emergency services, and more.

5 Conclusion

The study demonstrated that DNN could effectively classify minor roads in dense urban settlements with an accuracy of 85%. Various essential service providers can utilize the generated geodatabase for planning and optimizing the accessibility of their services. The study emphasizes the need for urban development policies that consider the accessibility of essential services to every citizen, which can be achieved through broader roads and redevelopment of built-up areas. Overall, the study provides valuable insights for urban planners and policymakers to improve the quality of life for people residing in dense urban settlements.

The study demonstrates the potential of using machine learning techniques, specifically DNN, to accurately classify minor roads in dense informal settlements using high-resolution satellite imagery. The model's high accuracy can help generate neighborhood information on the roads and overall city objects, and it can be implemented in any city across the globe for decision-making tools. However, limitations such as the inability to differentiate between roads and slum buildings and missing details such as hanging objects on the roads can be addressed by incorporating high-resolution height information from remote sensing technologies such as LiDAR.

Future Scope

- The study's outcomes have several practical implications. Firstly, the dedicated classification of informal and formal settlements can be used to correlate with socioeconomic factors such as energy consumption to develop policy instruments.
- Secondly, the model's results can be used to appropriately map and assess utility services such as electricity, toilets, water supply, and more.

- Thirdly, the model can be implemented in other dense cities of India, such as Delhi, Chennai, and Kolkata, where over 50% of the population resides in slum areas facing various socioeconomic and environmental issues.
- Finally, a high-resolution digital surface model (DSM) has been included to minimize the intermixing of building and road class pixels, improving the classification accuracy.

References

1. Mandeli K (2019) Public space and the challenge of urban transformation in cities of emerging economies: Jeddah case study. Cities 95:102409
2. Wang W (2016) A review of road extraction from remote sensing images. J Traffic Transp Eng Engl Ed 3:271–282
3. Cheng G (2017) Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network. IEEE Trans Geosci Remote Sens 55:3322–3337
4. Fu G, Liu C, Zhou R, Sun T, Zhang Q (2017) Classification for high resolution remote sensing imagery using a fully convolutional network. Remote Sens 9:498
5. Zhong Z, Li J, Cui W, Jiang H (2016) Fully convolutional networks for building and road extraction: preliminary results. In: 2016 IEEE international geoscience and remote sensing symposium (IGARSS) (IEEE, 2016), pp 1591–1594. https://doi.org/10.1109/IGARSS.2016.7729406
6. Yerram V (2022) Extraction and calculation of roadway area from satellite images using improved deep learning model and post-processing. J Imaging 8:124
7. Sahana M, Dutta S, Sajjad H (2019) Assessing land transformation and its relation with land surface temperature in Mumbai city, India using geospatial techniques. Int J Urban Sci 23:205–225
8. Alzubaidi L (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8:53
9. Madapur ABS (2018) Inclusive streets for fostering the notion of public space. In: Proceedings of the international conference on applied research in engineering, science and technology (ACAVENT, 2018). https://doi.org/10.33422/icarest.2018.09.43
10. Kumar V, Bandhyopadhyay S, Ramamritham K, Jana A (2020) Optimizing the redevelopment cost of urban areas to minimize the fire susceptibility of heterogeneous urban settings in developing nations: a case from Mumbai, India. Process Integr Optim Sustain 4:361–378
11. Kumar V, Jana A, Ramamritham K (2020) A decision framework to assess urban fire vulnerability in cities of developing nations: empirical evidence from Mumbai. Geo-carto Int 1–17. https://doi.org/10.1080/10106049.2020.1723718

Taxonomy of Structure-Based Routing Protocols

Abhishek Rathi, Harmandeep Singh, Tarun Kumar Rathi, and Anamika Chauhan

Abstract The utilization of Internet of Things (IoT) in daily life and networking has been on the rise. Smartwatches, computers, smartphones, and other electronic devices utilize IoT for communication infrastructure. The primary objective of IoT is to provide a communication mechanism between endpoints, which face a range of issues such as data security, availability, mobility, performance, congestion control, and load balancing. To facilitate this, various routing protocols are employed in IoT, including those based on network structure. However, there are several routing challenges and problems in IoT that require further investigation, as discussed in this study.

Keywords Flat-based · Hierarchical-based · Location-based · IoT · Routing protocols · Energy efficiency · Quality of service

1 Introduction

In 1999, Kevin Ashton proposed the Internet of Things at a conference, who defined it as "uniquely connected devices" utilizing radio-frequency identification. While thinking about the development of the Internet of Things, many technologies including RFID, wireless sensor networks, and mobile computing are taken into consideration. Two billion individuals were using the Internet on a daily basis for a variety of purposes worldwide. The digital and networked ecosystem involves the sharing of packets between nodes, which can be any type of device such as end devices, routers, or data generators. Routers and switches are utilized as equipment to implement the programmed routing protocols and algorithms.

A. Rathi (✉) · H. Singh · T. K. Rathi · A. Chauhan
Department of Information Technology, Delhi Technological University, New Delhi, India
e-mail: abhishekrathi573@gmail.com

A. Chauhan
e-mail: anamika@dce.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_31

Routing is a significant aspect of IoT systems due to its unique properties, making it a complex component. A routing protocol is also known as a routing policy, which outlines how routing devices link up in the network and share information to determine the best route between two nodes among various options.

2 Internet of Things

The Internet of Things (IoT) has garnered significant attention in recent times for its ability to connect physical objects to the Internet, allowing them to collect and exchange data. Although the concept of IoT can be traced back to the early days of the Internet, it has only become feasible in recent times, thanks to technological advancements such as small, low-cost sensors and wireless communication devices that can be incorporated into everyday objects [1]. Key benefits of IoT are collecting data from a wide range of sources to monitor environmental conditions and user behavior. As technology continues to advance, we can anticipate even more innovative applications of IoT in the coming years.

Characteristics of IoT are

1. Intelligence: What makes IoT intelligent is the combined usage of computing, software and hardware. This intelligence capability of IoT enables the objects to respond adapt to their surroundings help perform certain tasks.
2. Connectivity: Communication and connectivity of sensors and objects is important characteristic of and IoT infrastructure.
3. Enormous Scale: In an IoT system, the number of devices that must be controlled and coordinated are increasing day by day as the new use cases emerge.
4. Sensing: Sensors enable IoT to have a grasp over the physical world.

Applications of IoT are

Below are a few illustrations of clever IoT applications:

- Automated Smart Homes: Lighting, temperature control, security systems, and entertainment systems are just a few of the things that can be automated in a home with IoT devices.
- Agriculture: The Internet of Things (IoT) has multiple applications in the agricultural industry, such as precision farming, soil analysis, and weather monitoring.
- Transportation: The transportation industry utilizes IoT technology to achieve tasks such as keeping track of vehicles, optimizing routes, and monitoring the health of vehicles. Figure 1 shows the different applications of Internet of Things.

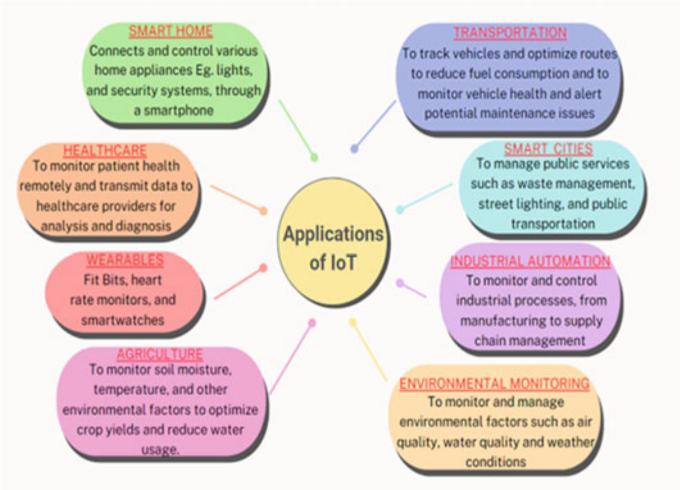


Fig. 1 Applications of Internet of thing

3 Routing Protocols in IOT

- Routing protocols are essential to transmit data in IoT networks efficiently.
- The goal of routing is to maintain a path between devices by selecting the shortest or easiest path to the destination [2].
- Based on wireless communications, there are three types of routing protocols in the Internet of Things [3].

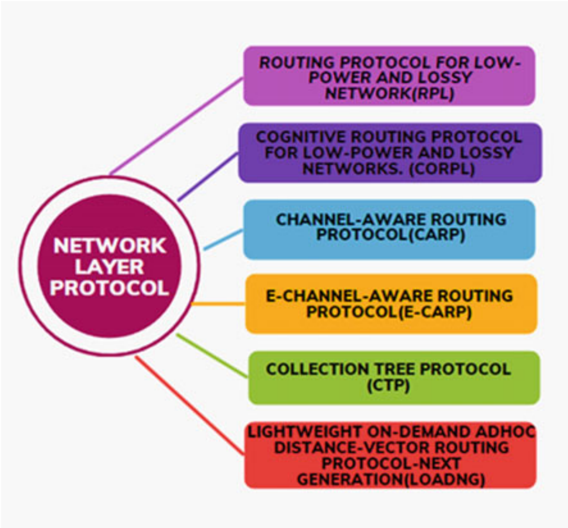
Networking Layer Protocol:

- Security: Security is a critical consideration in IoT networks. The networking layer protocol used in an IoT network should be secure and provide protection against unauthorized access and data breaches.
- Scalability: IoT networks are expected to grow rapidly in the coming years. The networking layer protocol used in an IoT network should be scalable to support large numbers of devices.
- Power Consumption: With IoT networks, power consumption is a crucial factor. An IoT network's networking layer protocol should be developed to reduce power consumption and increase battery life for networked devices.

Types (Fig. 2 is visual representation of Types of Network Layer Protocol):

1. RPL: RPL supports multiple routing metrics, allowing for customization and optimization of network performance [4].
2. CTP: CTP uses low-power listening to conserve energy and extend node lifetime, making it suitable for long-term IoT deployment.

Fig. 2 Types of network layer protocol



3. LOADng: Energy-efficient, designed for low-power consumption, ideal for IoT devices with limited power sources.
4. CORPL: CORPL is scalable, allowing for easy implementation and deployment in large-scale IoT networks with diverse network topologies [4].
5. CARP: CARP protocol helps avoid packet collisions in IoT networks using a deterministic algorithm.
6. E-CARP: E-CARP protocol improves energy efficiency in IoT devices by allowing them to sleep during periods of inactivity and adjusting transmission power based on distance.

Network Organization:

Network organization is essential for the routing protocol to function in the Internet of Things. This networking structure gives a succinct rundown of a network's features. The following is a list of the numerous protocols engaged in a core network:

i. Flat-Based:

- Another name for this is horizontal routing. Often, networks with a horizontal or flat topology employ this kind of protocol. Coordination of the network or its traffic is not required in this scenario. Contention-based planning is flat-based routing.
- Below are the different types of Flat-Based Routing Protocols (Fig. 3).

(a) Sensor Protocol for Information via Negotiation (SPIN):

  - Negotiation-based communication: The SPIN protocol uses a negotiation-based approach for communication between sensors and a base station in IoT networks [5].

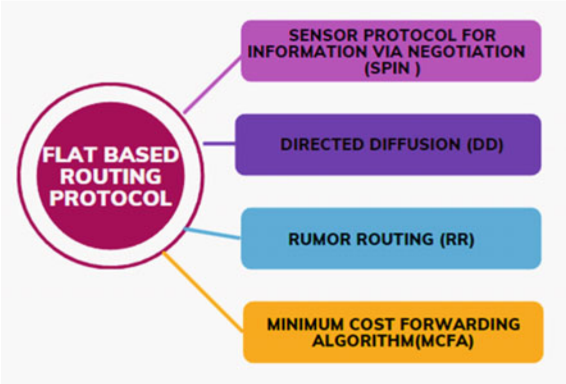


Fig. 3 Various flat-based routing protocol

- Lightweight and flexible: SPIN is a lightweight protocol that is designed to be flexible and easily customizable to different IoT applications.
- Scalable: This flat-based structure allows for efficient communication and management of large numbers of sensors in IoT networks.

(b) Directed Diffusion (DD):

- DD protocol is a lightweight protocol that provides a common language for IoT devices to communicate with each other [6].
- It uses a standardized format for data descriptions, allowing devices to quickly and easily understand the data being transmitted.
- DD protocol supports a wide range of data types, including integers, strings, and binary data, as well as more complex data structures such as arrays and objects.

(c) Rumor Routing (RR):

- Decentralized Communication: The RR protocol allows devices to exchange information with each other without the need for a centralized communication entity such as a server [7].
- Energy-Efficient: The RR protocol is designed to be energy-efficient, which is important for IoT devices that often have limited battery life.
- Robustness: The RR protocol is also designed to be robust to network failures and changes.

(d) MCFA:

- Minimum Cost Forwarding algorithm is a routing protocol that is designed specifically for IoT networks, which are characterized by a large number of low-power, resource-constrained devices [8].
- MCFA uses a distributed algorithm to update the cost map, which allows for efficient adaptation to changes in the network topology.

- MCFA is well-suited for IoT applications that require reliable and efficient communication over a large area. Comparative analysis of Flat-Based Routing Protocols in given in Table 1.

i. Hierarchical-Based:

One approach to routing in IoT networks is hierarchical routing. In this approach, the network is divided into multiple levels of hierarchy, with each level responsible for routing packets between nodes within its own level and between levels.

(a) Cluster-based routing protocol:

In cluster-based routing protocols, the network is divided into clusters, each of which is headed by a cluster head (CH) responsible for routing packets within the cluster and between clusters. Examples of cluster-based routing protocols include LEACH, (HCEL) [9], Advanced Multi-hop LEACH [10], (SEP) [11], (DEEC) [6], (SOC-M2M) [12].

Types of Hierarchical Routing Protocols (Refer Fig. 4)

Tree-based routing protocol: Tree-based routing protocols are particularly useful in IoT networks that have a fixed infrastructure, such as a wireless sensor network deployed in a building or factory. Examples of tree-based routing protocols include LBT [13], ETSP [13], and WSTDO [14].

Mesh-based routing protocol: Mesh-based routing protocols are designed to create a mesh structure in which every node can communicate with several other nodes in the network. Ad-hoc On-demand Distance Vector (AODV) and Dynamic Source Routing (DSR) are two examples of mesh-based routing protocols. Comparative analysis of different Hierarchical-Based Routing Protocols in given in Table 2.

ii. Location-Based:

- Location-Based Routing enables IoT devices to communicate with each other based on their proximity to each other. This is achieved by using location-aware technologies such as GPS, Wi-Fi, and Bluetooth to determine

Table 1 Characterization of flat structure routing protocols with its comparative analysis

Various routing protocol	SPIN	DD	RR	MCFA
Control overhead	Lower	Lower	Lower	Lower
Based on query	Yes	Yes	Yes	No
Adaptability strength	Feasible	Restricted	Restricted	No
Based on scalability	Restricted	Restricted	Perfect	No
Support multipath routing	Yes	Yes	No	No

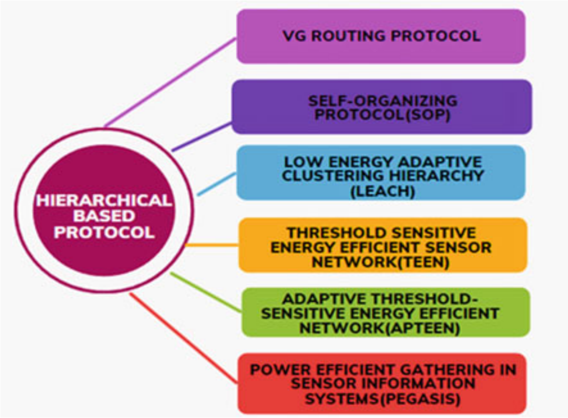


Fig. 4 Various hierarchical routing protocol

Table 2 Characterization of hierarchical structure routing protocol with its comparative analysis

Various routing protocol	VGR	SOP	LEACH	PEGASIS	TEEN
Control overhead	Higher	Higher	Higher	Higher	Higher
Based on query	No	No	No	No	No
Adaptability strength	No	No	Restored	Restored	Restored
Based on scalability	Perfect	Restricted	Perfect	Perfect	Perfect
Support multipath routing	Yes	No	No	No	No

the physical location of each device. There are several benefits of location-based routing in IoT, such as improving network efficiency and reducing latency. Additionally, location-based routing can help to reduce network congestion by directing data packets away from heavily trafficked areas.

- Accuracy of Location Data: Location-based routing relies heavily on the accuracy of location data obtained from sensors and other devices.
- Scalability: The network has to be scalable enough to accommodate the growing data traffic as the number of IoT devices rises.
- Security: Due to the fact that location-based routing includes sending sensitive data, it's crucial to make sure the network is safe and guarded against unwanted access.

Several routing algorithms rely on location information to optimize packet forwarding. These algorithms include Geographic Adaptive Fidelity (GAF), SPAN, Greedy Other Adaptive Face Routing (GOAFR), and GEAR (see Fig. 5). Comparative analysis of different location-based routing protocols is given in Table 2.

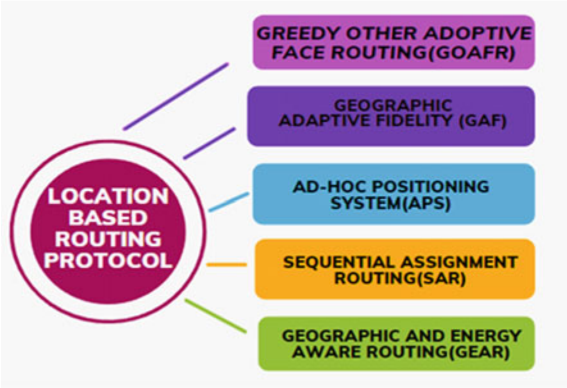


Fig. 5 Various location-based routing protocol

4 The Brief Comparison Between Flat, Hierarchical, and Location-Based Routing Protocols

Flat-Based Routing Protocols:

1. **Scheduling basis**: Based on content scheduling.
2. **Collision**: Collisions occur frequently.
3. **Duty cycle**: The nodes in flat-based routing are sleeping while not in use, and the duty cycle varies.
4. **Efficiency**: Optimal routing.
5. **Fairness**: Fairness is not assured.
6. **Data aggregation**: Neighbors are responsible for data aggregation.
7. **Synchronization**: This protocol does not require synchronization.

Hierarchical-Based Routing Protocols:

1. **Scheduling basis:** Based on reservation scheduling.
2. **Collision:** Collisions are less frequent and prevented.
3. **Duty cycle:** By putting the nodes to sleep, the duty cycle of this routing protocol can be decreased.
4. **Efficiency:** Inefficient routing.
5. **Fairness:** Fairness is assured.
6. **Data aggregation:** The cluster leader is in charge of gathering data.
7. **Synchronization:** This protocol requires both local and global synchronization (Table 3).

Location-Based Routing Protocols:

1. **Scheduling basis**: Based on content scheduling.
2. **Collision**: Collisions are less frequent and prevented.
3. **Duty cycle**: When not in use, the nodes are asleep in location-based routing.
4. **Efficiency**: Optimal routing.

Table 3 Characterization of location structure routing protocol with its comparative analysis

Various routing protocol	GOAFR	GAF	SAR	GEAR
Control overhead	Higher	Medium	Higher	Medium
Based on query	No	No	Yes	No
Adaptability strength	No	Restricted	No	Restricted
Based on scalability	No	Perfect	Restricted	Restricted
Support multipath routing	No	No	No	No

5. **Fairness**: Fairness may be assured.
6. **Data aggregation**: Clustering is used for data aggregation.
7. **Synchronization**: This protocol requires synchronization.

5 Conclusion

In conclusion, taxonomy on structure-based routing protocols in IoT has been discussed, with a focus on the three main types: flat, hierarchical, and location-based. Each type has its own advantages and limitations, depending on the application and the network structure. Flat routing protocols are simple and easy to implement but lack scalability, while hierarchical routing protocols provide better scalability but have higher overhead. Location-based routing protocols are suitable for applications that require location information but may be challenging to implement in practice. Therefore, the choice of the routing protocol should be made based on the specific requirements of the IoT application. Overall, this discussion provides a useful framework for understanding the different types of routing protocols and their characteristics in IoT networks. Researchers and practitioners can use the taxonomy as a guide to select the best procedure for their unique application requirements. Moreover, future research can focus on developing hybrid protocols that can provide a balance between network overhead and routing delay while considering the network's structure and topology.

References

1. Agrawal SK, Sharma K (2016) 5G millimeter wave (mmWave) communications. In: 2016 3rd international conference on computing for sustainable global development (INDIACom). IEEE, pp 3630–3634
2. Intanagonwiwat C, Govindan R, Estrin D (2000) Directed diffusion: a scalable and robust communication paradigm for sensor networks. In: Proceedings of the 6th annual international conference on mobile computing and networking, pp 56–67
3. Kharkongor C, Chithralekha T, Varghese R (2016) A SDN controller with energy efficient routing in the internet of things (IoT). Proc Comput Sci 89:218–227

4. Alahari HP, Yalavarthi SB (2017) A survey on network routing protocols in internet of things (IOT). Int J Comput Appl 160(2):18–22
5. Heinzelman WR, Kulik J, Balakrishnan H (1999) Adaptive protocols for information dissemination in wireless sensor networks. In: Proceedings of the 5th annual ACM/IEEE international conference on mobile computing and networking, pp 174–185
6. Jacquet P, Muhlethaler P, Clausen T, Laouiti A, Qayyum A, Viennot L (2001) Optimized link state routing protocol for ad hoc networks. In: Proceedings. IEEE international multi topic conference, 2001. IEEE INMIC 2001. Technology for the 21st century. IEEE, pp 62–68
7. Braginsky D, Estrin D (2002) Rumor routing algorithm for sensor networks. In: Proceedings of the 1st ACM international workshop on wireless sensor networks and applications, pp 22–31
8. Ye F, Chen A, Lu S, Zhang L (2001) A scalable solution to minimum cost forwarding in large sensor networks. Comput Commun Netw 304–309
9. Al-Kaseem BR, Nyanteh AO, Al-Raweshidy HS (2016) Self-organized clustering technique based on sink mobility in heterogeneous M2M sensor networks. In: International conference for students on applied engineering (ISCAE), pp 431–436
10. Perkins CE, Bhagwat P (1994) Highly dynamic destination sequenced distance-vector routing (DSDV) for mobile computers. ACM SIGCOMM Comput Commun Rev 24(4):234–244
11. Boukerche A, Das SK, Fabbri A (2001) Analysis of a randomized congestion control scheme with DSDV routing in ad hoc wireless networks. J Parall Distrib Comput 61(7):967–995
12. Villasenor-Gonzalez L, Ge Y, Lament L (2005) HOLSR: a hierarchical proactive routing mechanism for mobile ad hoc networks. IEEE Commun Mag 43(7):118–125
13. Malluh AA, Elleithy KM, Qawaqneh Z, Mstafa RJ, Alanazi A (2014) EM-SEP: an efficient modified stable election protocol. American Society for Engineering Education (ASEE Zone 1), pp 1–7
14. Sharma V, Saini DS (2015) Performance investigation of advanced multi-hop and single-hop energy efficient LEACH protocol with heterogeneous nodes in wireless sensor networks. In: Advances in computing and communication engineering (ICACCE), pp 192–197

Contactless Gesture Recognition Using Air Canvas

Janhavi Gosavi, Neerajaa Kadam, Ameya Shetty, Anant Verekar, and Pinki Vishwakarma

Abstract The proposed work is a real-time implementation in Python for air-writing using algorithms like media pipe hands and easy OCR. Air writing has emerged as one of the most difficult study fields in image processing and pattern recognition in recent years. Air writing is the act of using one's hands or fingers to scribble letters or words in a clear area. The proposed work aims at the implementation of a contactless gesture recognition software which uses an air canvas which sums up as providing an image input to the machine via a camera device, while the machine processes it and gives a delineated digital output of the work in real time, features such as optical character recognition, screenshot, and duplication will be an added attempt at closing the gap of limitations.

Keywords Air canvas · Media pipe · OCR · Python · OpenCV · NumPy · PyAutoGUI · PyGettWindow · Image processing · Computer vision · Word document automation

J. Gosavi (✉) · N. Kadam · A. Shetty · A. Verekar · P. Vishwakarma
Department of Computer Engineering, Shah and Anchor Kutchhi College, Mumbai, India
e-mail: janhavi.gosavi_19@sakec.ac.in

N. Kadam
e-mail: neerajaa.kadam_19@sakec.ac.in

A. Shetty
e-mail: ameya.shetty_19@sakec.ac.in

A. Verekar
e-mail: anant.verekar_19@sakec.ac.in

P. Vishwakarma
e-mail: pinki.vishwakarma@sakec.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_32

1 Introduction

The proliferation of technologies has been immense in these past hundred years with the field of machine learning and artificial intelligence having an immeasurable expanse of utility and popularity. Among these, writing on an air canvas is one of the top trending technologies that has a high practicality and several benefits. Contactless gesture recognition using air canvas is a project that mainly focuses on the concept of air writing, which simply put is a technology wherein content written or drawn in space is identified and presented as a digital output [1]. This technology is a brilliant amalgamation of domains like image processing and computer vision. The primary idea is to recognise the fingertip or instrument acting as a pointer using a camera lens and trace its path such that the content written or drawn by the end user is recorded by the computing machine, followed by the processing of this input to provide a digital output on the machine screen in real time. The field of computer vision in artificial intelligence (AI) enables systems and computers to extract usable information from digital pictures, videos, and other visual inputs and to perform actions or make suggestions in response to that information. If AI allows computers to think, computer vision allows them to see, watch, and comprehend.

Computer vision research focuses on recreating some of the complexity of the human visual system so that computers can recognise and analyse elements in photos and videos in the same way that humans do. Computer vision has just lately been completely functional. Human vision has a historical edge over computer vision in that it has been around for a longer period of time. Human sight has the advantage of learning how to discern between items, measure their distance from the observer, decide whether they are moving, and assess whether an image is right over the course of a lifetime. In recent years, the area has made considerable strides and been able to beat humans in various tasks relating to object detection and object labelling. This has been made possible by developments in artificial intelligence, deep learning, and neural networks. The quantity of data we produce today, which is subsequently utilised to train and improve computer vision, is one of the key elements influencing the development of this technology. Computer vision trains computers to perform identical tasks in significantly less time by using cameras, data, and algorithms instead of retinas, optic nerves, and a visual brain. Because it can assess hundreds of goods or processes per minute while detecting invisible faults or issues, a system trained to inspect things or monitor a manufacturing asset can quickly exceed human abilities.

There are numerous different ways to write in digital art, including utilising a touch-screen surface, keyboard, digital pen, stylus, electronic hand gloves, and so forth. However, in our system, we combine hand gesture detection together with Python programming to produce genuine human–machine interaction. Air writing differs from its predecessors in the way that our team is utilising newer technologies and Python (a user-friendly language) and is being made keeping in mind the latest trends and demands in market, while ensuring the software is a smooth working, compatible, and easy to implement product of programming.

Fine art created before the advent of digital technology is referred to as traditional art. Pen and paper, as well as chalk and a board, are all used in the traditional method of writing. Traditional and digital art are interdependent and related to one another. The needs of human existence are, by far, the most important driving force for social change, even though this is not necessarily the will of the people [2]. In art, the similar situation arises. The proposed framework wishes to effectively understand the fundamental information on the differences between digital art and traditional art given the current circumstances, where both forms of art are equally advantageous. The core idea behind advanced craftsmanship is creating a hand signal recognition system that allows for digital writing.

Furthermore, this provides a feasible alternative to paper, as well as whiteboards/blackboards, and is an economical as well as eco-friendly alternative to these. Wasted paper is not fresh information. Writing, sketching, doodling, and other paper-wasting activities are common. A4 size paper typically requires 5 L of water to produce, 93% of writing is done on paper made from trees, paper makes up 50% of company trash, 25% of landfill garbage, and the list is unending. Paper waste damages the environment by using huge amounts of water, cutting down trees, and producing tonnes of trash. This problem can be rapidly fixed using air writing.

2 Related Work

Contactless gesture recognition pairs the primary aspect of an air canvas with secondary features like optical character recognition, screenshot, duplication (copy-paste), and so on. The efficacy of this system lies in its diverse utility in fields like education, self-recognition, art, formal demonstrations, and others.

2.1 Literature Survey

Air writing, as defined by Al Abir, M.D. et al., is the practice of writing letters or words in open space while moving your body. In their article titled Deep Learning Based Air-Writing Recognition with the Choice of Proper Interpolation Technique published on 16 December 2021, the authors define air-writing recognition as a special example of gesture recognition in which gestures correlate to characters and digits written in the air [3]. In this work, they used time-series data to identify letters and digits in the English alphabet. It is challenging to prepare sensor data for deep learning systems while minimising loss of data.

Can Uysal et al. clarify how the need for human–machine interactions is growing as robots become more and more vital to human life. Traditional devices such as a touch screen or keyboard may not be desired in future leisure, residential, and industrial applications. In their February 2017-released paper RF-Wri: An Efficient Framework for RF-Based Device-Free Air-Writing Recognition [4], a brand-new framework for device-free air-writing recognition that can accurately categorise all capital letters (26) using inexpensive SDR modules is provided.

According to Muhammad Arsalan and Avik Santra, radar technology is essential for the contact-free detection of gestures or movements of hands, which creates a different and natural type of computer–human interaction. Writing linguistic symbols or words in empty space with hand gestures is known as “air-writing”. Character recognition in air-writing based on network of radars for human–machine interface is a study that was released on 1st October 2019 [5]. The authors of this research suggest that the radars in a network can operate as a virtual whiteboard on which users can create random symbols or numbers in the air with hand-held markers.

Dr. R. Srinivasan and Pavithra Ramasamy explain that the “Finger Motion Tracking System” aims to recognise the English character written in the air using a finger. In this work titled An Economical Air Writing System Converting Finger Movements to Text Using Web Camera published on 16 January 2016, the recognised character is translated into text and shown on the screen [6]. The conversion of finger motions to text is one of the freshly expanding fields of research in the field of gesture recognition, which is what the authors of this paper propose as they develop this system.

In the recent years, Ashutosh Kr. et al.’s paper titled Air Writing using Python identified writing in the air as one among the most captivating and arduous research areas in the fields of pattern recognition and image processing. Date: 5 May 2022 [2]. They argue in this paper that the framework might undermine conventional writing techniques. It provides a simple technique for taking notes, negating the need to carry a cell phone about with you.

Writing in the air has emerged as one of the most engrossing and exigent research areas in the domains of pattern recognition and image processing, according to Ashutosh Kr. et al.’s paper titled Air Writing using Python published in recent years on the 5th of May 2022 [2]. In this paper, they contend that the framework might jeopardise accepted writing practices. It offers a straightforward method for taking notes, eliminating the need to carry a smartphone about.

3 Proposed Methodology

The most exciting part of our framework is this. Functionalities involve several aspects of writing. In this manner, the total number of actions is equal to the number of gestures used to run the framework.

The following stages control how the project is worked on as a whole:

1. The first step is to get started on the task. It shows how the entire programme functions.
2. The second stage, known as image collection, describes taking photos using a computer device’s web camera.

3. Getting to the datapoints of the images from the camera is stated as the third step, or image processing.
4. The final step in the process is the projection of output onto the screen, which is accomplished by connecting the subsequent foci and extending it there.

3.1 Analysis and Framework

The core capabilities we kept in mind for our framework are:

(a) Writing mode
(b) Colour mode
(c) Clear
(d) OCR
(e) Copy paste
(f) Screenshot
(g) Word document automation.

3.2 Details of Hardware and Software

Software Requirements

Visual Studio (Code Editor)
Python 3.10
OpenCV
NumPy.

Hardware Requirements

Specifications of PC's minimum requirement
Minimum 4 GB RAM
Good Internet connectivity
Web Cam.

3.3 Design Details

See Fig. 1.

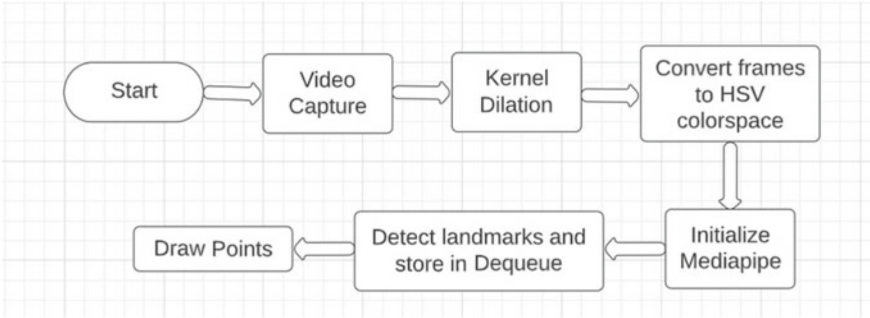


Fig. 1 Flowchart of project design

3.4 Methodology

A PC with a web camera built in it is the only requirement for this work. We will set-up the computer or screen to read anything the client types in front of it. In order to identify the scribbling the client will draw on the air, we used Open CV and we used Python as the coding language. First a fingertip detection model [7] is implemented which then is used to execute the air-writing module. The code will then be tested for results followed by which the code is updated to incorporate a digital canvas module, OCR module, and duplication module. The completion of this takes place by implementing screen capture module and then testing the overall application for success.

The various modules implemented in the system are:

Writing Mode

Here, the framework will follow the fingertip co-ordinates and will store them.

Colour Mode

The client can change the shade of the text out of the different accessible tones.

Clear Mode

If there is some mistake in writing or diagram served as an input to the system, the user can clear the canvas and redo their work.

Copy Paste Mode

The characteristic of this feature is to avail the user to copy and paste the inputs given by them.

Screenshot Mode

It allows the user to take a screenshot of the input.

Word Document Automation

Creation of word document including all the input data.

OCR Mode

Optical character recognition of the input captured in the screenshot.

4 System Accuracy

Module 1: Writing mode: 99%.

99 out of 100 instances gave a successful finger detection, hence module 1 accuracy is 99%.

Module 2: Colour mode: 100%.

100 out of 100 instances gave successful colour change in texts, hence module 2 accuracy is 100%.

Module 3: Clear mode: 97%.

97 out of 100 instances the user could clear the digital canvas, hence module 3 accuracy is 97%.

Module 4: Copy paste mode: 93%.

93 out of 100 instances the user could copy paste the inputs given by them, hence module 4 accuracy is 93%.

Module 5: Screenshot mode: 98%.

98 out of 100 instances successfully took the screenshot of the digital air canvas, hence module 5 accuracy is 98%.

Module 6: Word document automation mode: 98%.

98 out of 100 instances gave a successful creation of word document containing the user input, hence module 6 accuracy is 98%.

Module 7: OCR mode: 68%.

68 out of 100 instances gave a successful optical recognition of the user input, hence module 7 accuracy is 68%.

5 Result and Analysis

The proposed framework allows users to drought down the points from the fingertips to the blank canvas, then it also offers a wide range of colour options such as red, blue, green, black, and yellow. Anything that is being written in the air in front of the camera will be demonstrated on the screen and displayed on the air canvas. The proposed work also offers features to take screenshot of the content on the air canvas and automatically saves it in a folder that is dated in real time without any human intervention. To clear the air canvas the "clear" button needs to be selected. These buttons can be seen at the top of the screen.

Features of screenshots can also be accessed by just selecting one button, word documents containing screenshots of the particular folder or real-time screenshot taken from the air canvas can be made just by selecting one button displayed on the screen. The screenshots inside the folder are numbered serially with the help of the code and no manual process. The feature to copy the image from the air canvas and paste it directly inside the computer's clipboard is also available. Ability to perform OCR on the given canvas and detect any words given to it is also present (Figs. 2, 3, 4 and 5).

Writing Mode

The framework first detects the object that is the palm of the hand and the fingertips are detected and the tip of the index finger is detected as the pointer for air writing. Pinching of index finger and the thumb, ceases the functionality to enable accurate and clean writing.

Colour Mode

The user can change colour of texts using these tones like red, blue, yellow, and green and can be utilised to make effective visual display.

Clear Mode

The proposed system needs a signal to add a backspace. After a specific session has concurred, the user can clear the canvas to begin anew for the next part of their work.

Fig. 2 Screenshot of air-writing implementation on the digital canvas

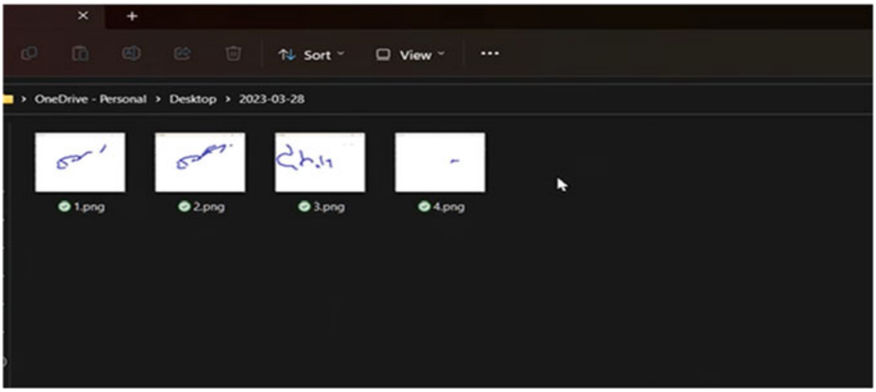


Fig. 3 Screenshot of the automatically created folder

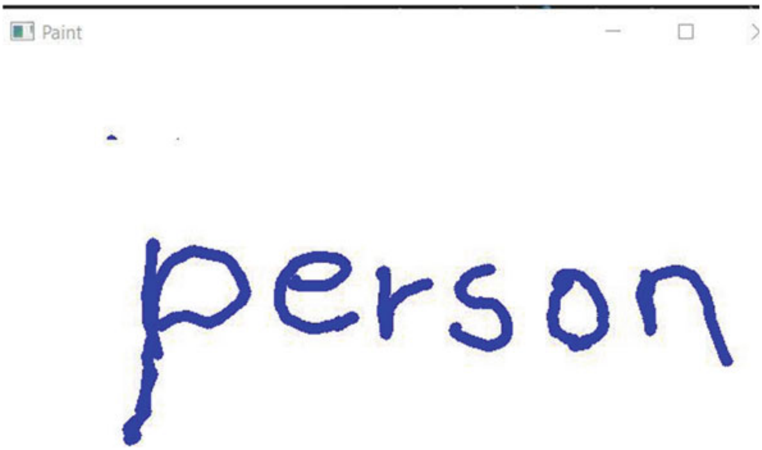


Fig. 4 Screenshot of input for OCR

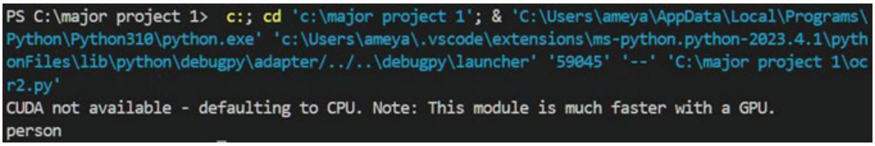


Fig. 5 Output for OCR mode

Copy Paste Mode

In this module, the feature to copy paste the digital inputs is given by the user on air canvas.

Screenshot Mode

In this mode, the user has been enabled to take a screenshot of the digital canvas where the output has been drawn by the user.

Word Document Automation

In this module, the user has been provided with a feature to create a word document of all the data input by them through the air canvas by making a compilation of all output screens in a word file format.

OCR Mode

The module here enables the user to apply optical character recognition over a screenshot of the digital canvas and convert the abstractly written text and numerals to a more digital-friendly format.

6 Conclusion

Contactless gesture recognition using air canvas is a major project designed and implemented by our team which focuses on the primary feature of air writing. The add on features include digital white canvas, virtual ink with customisable colour palette, optical character recognition (OCR), duplication (copy and paste), and screen capture. We have used Python as the primary programming language and visual code editor as the platform for the implementation. This project aims to provide an eco-friendly alternative to paper which may be applied in the fields of education, art, development of cognitive abilities in children, formal or informal presentations, and so on. Furthermore, air writing can be utilised to assist with self-regulation in autistic children. Air writing is especially beneficial for user interfaces where the user cannot type on the keyboard or write on the touchpad/touch screen, or for text input for intelligent system control. This project is entirely dependent on computer vision and Python libraries, particularly NumPy and OpenCV.

Future work regards our proposed system involves bettering of the mapping accuracy and increasing the precision of the OCR module. Furthermore, integration of the text generated by OCR with within word, paint, excel programmes can boost its utility widely in the computer and teaching industry and will be beneficial for commercialisation of this technology globally. The OCR feature can also be bettered and perfected for languages other than English and using natural language processing technology it can be redefined phenomenally.

Air Wwiting can also serve as a communication aid for those who have hearing loss. They may either use AR to show their air-written text or turn it into voice. With little interruption, one may swiftly write in the air and continue working. In addition, there is no requirement of paper when writing in the air. Everything is preserved digitally. Several other features may be added upon our current work such as user GUI, augmentation protocols, voice recognition, and so on to expedite its efficacy in aforesaid use cases.

References

1. Amma C, Georgi M, Schultz T (2014) Airwriting: a wearable handwriting recognition system. Pers Ubiquit Comput 18:191–203
2. Pandey AK, Dheeraj, Tripathi M, Vidyotma (2022) Air writing using python
3. Abir FA, Siam MA, Sayeed A, Hasan MAM, Shin J (2021) Deep learning based air-writing recognition with the choice of proper interpolation technique
4. Uysal C, Filik T (2021) RF-Wri: an efficient framework for RF-based device-free air-writing recognition. IEEE Sens J 21:17906–17916
5. Arsalan M, Santra A (2019) Character recognition in air-writing based on network of radars for human-machine interface
6. Ramasamy P, Srinivasan R (2016) An economical air writing system converting finger movements to text using web camera
7. Chen YH, Huang CH, Syu SW, Kuo TY, Su PC (2021) Egocentric-view fingertip detection for airwriting based on convolutional neural networks. Sensors 21:4382

An Efficient COVID-19-Based Disease Detection on X-Ray Images Using CNN Model

Kumar Keshamoni, L. Koteswara Rao, and D. Subba Rao

Abstract The use of artificial intelligence (AI) and convolutional neural networks (CNNs) in medicine has enabled the analysis and categorisation of X-ray images for various diagnoses. This study proposes a deep learning architecture for the diagnosis of COVID-19 by categorising chest X-ray images. However, the limited availability of a large and high-quality chest X-ray image dataset posed a challenge in developing a reliable and accurate CNN classifier. To address this, the dataset underwent multiple pre-processing stages, including dataset balancing, picture analysis by medical professionals, and data augmentation, to achieve the best performance. The proposed CNN model achieved an overall accuracy of 90.27%, with 91.06% precision, 82.40% recall, and 85.61% FSCORE, with only 9.72% incorrect predictions when validated with a different set of COVID-19 X-ray images. Furthermore, a comparison with other machine learning techniques demonstrated that the suggested model outperformed existing models, including KNN, Naive Bayes, LSTM, SVM, ResNet, InceptionV3, Decision Tree, Logistic Regression, and MobileNet, when evaluated against an independent testing set. Therefore, this study presents an efficient approach for COVID-19 disease detection on X-ray images using a CNN model.

Keywords COVID-19 · X-ray · CNN model · Accuracy · Dataset · CNN · Recognition · Classification · Pooling · BCE

K. Keshamoni · L. Koteswara Rao (✉)
Department of ECE, Koneru Lakshmaiah Education Foundation, Hyderabad, India
e-mail: koteswararao@klh.edu.in

K. Keshamoni
e-mail: kumar.keshamoni@gmail.com

D. Subba Rao
Department of ECE, Siddhartha Institute of Engineering and Technology, Hyderabad, India
e-mail: subbu.dasari@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_33

1 Introduction

The "SARS-CoV-2" coronavirus, often known as the "Severe Acute Respiratory Syndrome" virus, was demonstrated to spread rapidly in late 2019. A coronavirus outbreak a virus with Chinese origin has been identified as the year's primary culprit (COVID-19). In March 2020, the WHO declared the situation to be a pandemic [1, 2]. According to research produced and maintained by state governments and international healthcare organisations, millions of people worldwide were impacted by the epidemic. The most dangerous ailment caused by COVID-19 is pneumonia, an infection that harms the lungs. Dyspnea, a high fever, a runny nose, and a cough are just a few of the symptoms of illness. The most common method for making this diagnosis is to examine chest X-ray images for abnormalities [3]. Another type of electromagnetic radiation that can penetrate is X-rays. To reach the targeted human body parts' internal organs, radiation is applied to these regions. The X-ray image shows the interior of the body in black and white. The use of X-rays is one of the oldest and most popular diagnostic methods in medicine. This is because they display the thoracic cavity, which comprises soft tissues like the lungs, blood vessels, and airways in addition to the bone structures of the chest and spine. Chest X-rays are used to diagnose lung illnesses, pneumonia, and other lung conditions [4]. For diagnosing COVID-19, X-ray imaging is more useful than conventional diagnostic methods. This approach has a number of advantages, including low cost, availability at various X-ray facilities, non-invasiveness, speedy processing, and cost efficiency. X-ray imaging is a superior choice for extensive testing and a quick and simple way to discover a pandemic like COVID-19 in the light of the current global healthcare crisis.

In the last ten years, research concentration has increased by a factor of ten as a result of deep learning and ANNs. Deep ANNs have outperformed conventional models in a number of crucial benchmarks. ANNs have advanced to the point where they are currently regarded as "state of the art" in a number of academic and industrial application disciplines, including natural language processing, speech recognition, image processing, and biological sciences [5]. The development of ANNs has a lot of potential applications in the medical industry, particularly in the processing and interpretation of data and images related to health. Both medical personnel and testing apparatus are urgently needed, as has been observed recently in many regions of the world. There is a connection between identifying COVID-19 carriers and learning how to assess and categorise chest X-rays, particularly in the light of the current epidemic. By analysing the results of a chest X-ray, the CNN can be used to determine if a person is healthy or COVID-19-infected. Early studies in this field have produced encouraging outcomes in terms of accuracy and other performance criteria for quickly and affordably diagnosing illnesses. In this study, COVID-19 X-ray images were more accurately classified using CNN with additional layers. While neural network topologies can process input in both one and three dimensions, they can only analyse images in two dimensions. CNN, a kind of DNN that is frequently used for processing visual pictures, was developed using the visual system of the human brain as its basis. Before the CNN model was trained, the dataset was obtained [6]. The model's extremely short and unbalanced training dataset was made better using the data augmentation approach. The model may end up with extra features as a result. Several image-flipping and image-rotation methods have been employed to provide additional data. To keep the number of photographs with distinct class names in balance, more instances of the minority class have been added to the collection. 551 COVID-19 photographs and 165 ordinary photos made up the batch of photos used to create the CNN model. The CNN model was evaluated on a test set after processes for enhancing the data and balancing the dataset. The CNN model's performance was then evaluated using a variety of performance criteria. For the suggested CNN model, metrics including accuracy, precision, recall, FSCORE, and incorrect predictions must be independently confirmed. In order to demonstrate that CNN is better than other machine learning techniques, its performance has been compared to that of numerous other machine learning models.

The following are a few of the study's key conclusions:

(i) A CNN with more convolutional layers, as the six-layer CNN suggested in this study, is the greatest tool for diagnosing COVID-19.
(ii) Several images are necessary for precise and efficient image classification using CNN models.
(iii) By adding new data to an existing, limited dataset, the data augmentation technique is highly effective at significantly improving the performance of the CNN model.
(iv) Since it makes CNNs invariant, the data augmentation approach is preferable for picture classification.
(v) The proposed CNN model outperformed the most recent ML techniques statistically considerably.
(vi) By combining CNN-based diagnosis with X-ray imaging, the medical industry may find it very beneficial to handle mass testing situations during pandemics like COVID-19 [5].

1.1 Purpose

Using a convolutional neural network (CNN) model, this study aims to develop an efficient and accurate COVID-19 detection system based on chest X-ray images. Multiple pre-processing stages are used to achieve the best results, addressing the issue of the limited availability of high-quality chest X-ray image datasets. The proposed CNN model outperforms other machine learning techniques and obtains an overall accuracy of 90.27%, with 9.72% of predictions being incorrect. The purpose of the study is to assist medical professionals in the early diagnosis and treatment of the disease.

1.2 Approach

Using a convolutional neural network (CNN) model, the authors propose an effective method for COVID-19 disease detection on X-ray images. Dataset balancing, medical professionals' image analysis, and data augmentation all helped to solve the problem of the scarcity of high-quality chest X-ray image datasets. Six convolutional layers, two fully connected layers, and a softmax activation function comprise the proposed CNN model. The model obtained high precision, recall, accuracy, and FSCORE. Validated, the model outperformed other machine learning techniques. The pre-processing phases were essential for attaining optimal performance. The recommended CNN model consists of a flattening layer, three fully connected layers, eight activation function levels, eight batch normalisation layers, six convolutional layers, six max pooling layers, and six dropout layers. Utilised were the binary cross-entropy loss function and the Adam optimizer. The CNN model proposed can be modified to classify images of furniture, room designs, and other elements of interior design.

2 Literature Survey

Deep learning is increasingly being employed in medical applications lately, particularly image-based diagnostics. Particularly when they had to figure out how to interpret medical images for computer vision issues, deep learning-based models performed really well. ANNs fared better than other frequently employed models and image analysis techniques [7, 8]. Because CNNs analyse and categorise medical images so successfully, they have become the industry standard [9, 10]. Lung disease [10], malarial parasites in thin blood smear images [11], breast cancer [12], wireless endoscopy images [13], interstitial lung disease [14], CAD-based diagnosis in chest radiography [15], classification-based diagnosis of skin cancer [16], and automatic diagnosis of various chest diseases using chest X-ray image classification [17] are just a few of the conditions that CNN has been used to diagnose. At [10], malaria has been found. Researchers have carried out researches and worked on projects that are essential to the diagnosis, treatment, and management of the COVID-19 virus since it was first identified in December 2019. Finding COVID-19 patients and treating them with AI image analysis methods have been discussed by researchers [18].

Data from lung CT scans can be analysed using deep learning techniques to precisely detect COVID-19 [18]. Researchers have developed a method for the open-source diagnosis of COVID-19 based on a detailed examination of CNN. In this study, COVID-19 patients were identified from X-ray scans using a specialised CNN configuration. This detection technique was said to have worked well. Concerns were raised about the X-ray dataset, which featured images of healthy persons, COVID-19 patients, and people who had common pneumonia [19]. Patients who tested positive for COVID-19 in this study were automatically classified using cutting-edge CNN architectures. Transfer learning has a hopeful accuracy of 97.82% for recognising COVID-19, according to the study's findings. The validity and adaptability of deep

CNNs of the decompose-, transfer-, and compose kinds with regard to the classification of chest X-ray images for the identification of COVID-19 were studied in another recent study that is pertinent to the subject [20]. Within the last ten years, this investigation was carried out. According to the authors' report, the study's findings are 95.12% accurate, 97.91% sensitive, and 91.87% specific. The dataset's paucity of X-rays and poor-quality photographs of COVID-19 patients reduced the models' accuracy. Preparing datasets, adding new data, and creating a CNN with more layers are the main objectives of this project. This will enhance the performance of the COVID-19 diagnosis. In the part that follows, this will be discussed in more detail.

3 Workflow

Figure 1 shows that the first step was to collect the primary dataset, which included pictures of people who were confirmed to have COVID-19 and pictures of people who were healthy controls. The dataset was then examined by the study's medical experts, who excluded any X-rays with poor diagnostic accuracy or quality. The resulting dataset was exceptionally well-maintained since all of the X-rays were of high quality and easily interpreted using their knowledge. In the third stage, the dataset was expanded using conventional methods. The model training process was then used with the obtained data set. After training, the model's capability to detect illnesses was put to the test. The suggested CNN model was tested using both the core dataset and a separate validation dataset. The fraction of X-ray images that belong to each prediction class is given in Table 1 along with the distribution of datasets into training, test, and validation sets.

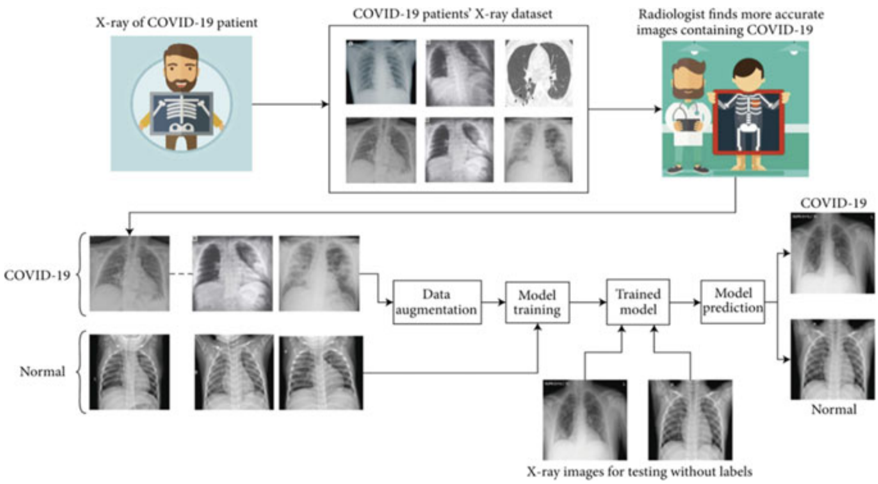


Fig. 1 System development workflow

Table 1 Training and testing of different images count from dataset

Datasets	Normal images	COVID-19 images	Total images
Total data	165	551	716
Testing data	17	56	73
Training data	148	495	643
Independent validation data	50	50	100

4 Implementation

The experiments in this report utilised 716 X-ray images. A total of 716 X-rays were collected, of which 551 were from COVID-19 patients and 42 were from healthy people or people with other conditions, including pneumonia. [6]. The fundamental dataset consists of 551 samples from two COVID-19 classes. Pre-processing was required for promising results because the sample was imbalanced. Just 54% of the original dataset was accurate when CNN was created, which was insufficient for the intended use. In this instance, performance metrics of up to 90.27% accuracy, 91.06% precision, 82.40% recall, 85.61% FSCORE, and 9.72% wrong predictions were attained.

In this study, we've used ECG and X-ray datasets to forecast COVID-19 disease at your request. Both datasets were constructed using several machine learning and deep learning techniques, including KNN, SVM, Logistic Regression, Decision Tree, Naive Bayes, CNN, LSTM, InceptionV3, ResNet, and MobileNet in deep learning.

To implement this paper effectively, the following modules have been put in place.

(a) Home: this module shows precautionary measures to users about COVID-19.
(b) Preliminary Analysis: using this module, we will ask some questions to user on their symptoms and based on symptoms COVID and non-COVID will be detected.
(c) COVID analysis using X-RAY: using this module, user can upload X-ray image and then algorithms will be applied on images to detect presence of COVID or NOT.

In above screen in red colour text, we can see based on symptoms "COVID disease detected" and similarly you can select various options and get output. Now click on "COVID analysis based on X-ray images" link to get below screen (Figs. 2 and 3).

In above screen selecting and uploading "Person 1" folder and then click on "Upload" button to get below output.

In above screen for Person 1 COVID detected in X-rays images. In the same image, we can see detection result, and similarly, you can upload and test other folders (Figs. 4 and 5).

In above screen selecting and uploading "Person 2" folder and below are the prediction outputs.

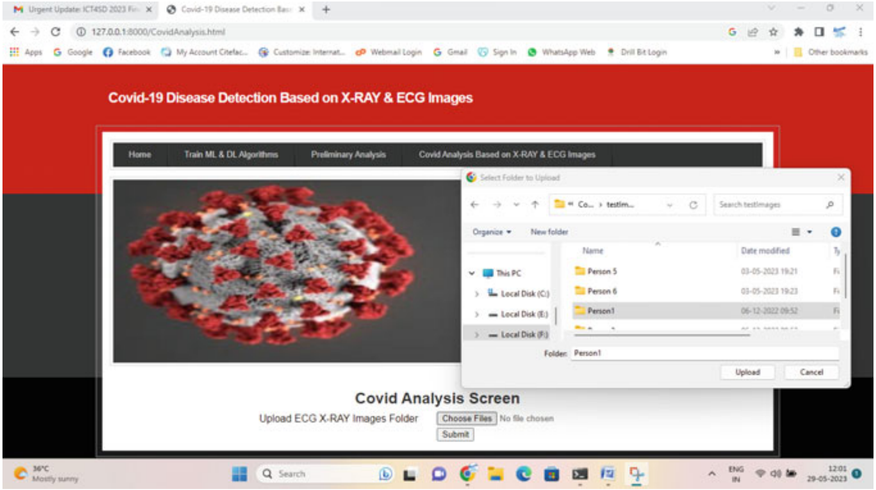


Fig. 2 Uploading "Person 1" folder for COVID test

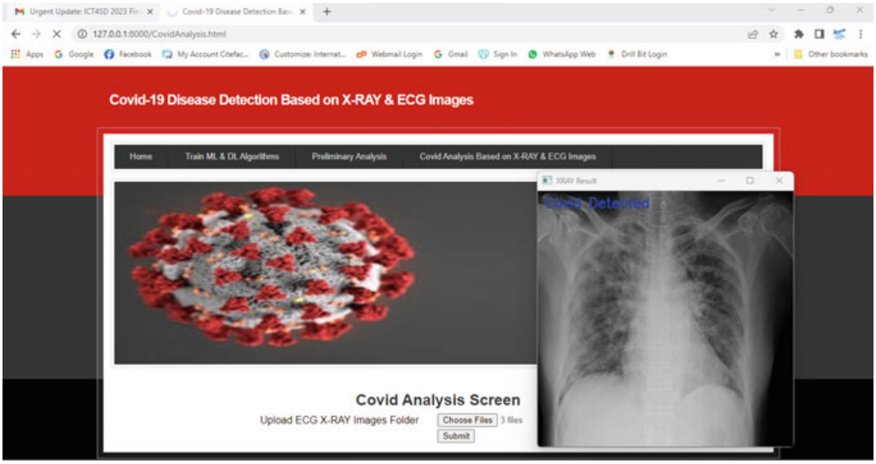


Fig. 3 Person 1 COVID detected in X-rays images

In above screen for Person 2 No COVID detected in X-rays image (Figs. 6 and 7).

In above screen uploading "Person 3" folder and below is the prediction output.

In above screen X-ray "COVID detected" and now close above image to get below output (Fig. 8).

In above screen in blue colour text, we can see detection output. Similarly, we can upload images and perform prediction.

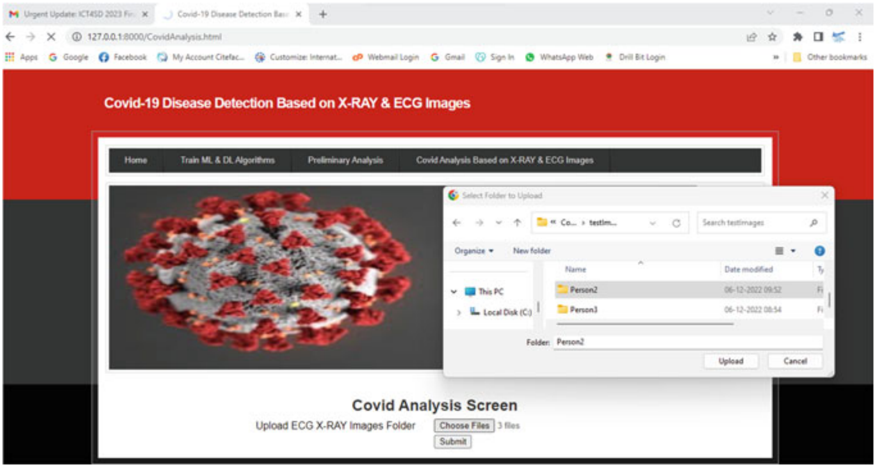


Fig. 4 Uploading "Person 2" folder and below are the prediction output

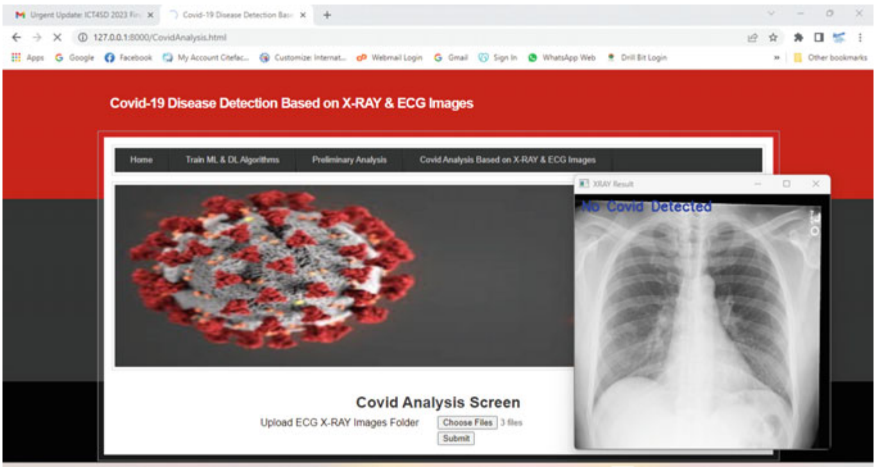


Fig. 5 Person 2 No COVID detected in X-rays image

4.1 Convolutional Neural Networks (CNNs)

CNNs were created using human vision as its model. Thanks to CNN, machines can perceive like humans. CNNs can be used for recognition, classification, and natural language processing [21]. CNNs are deep neural networks that activate nonlinearly, with maximum pooling, and convolutionally. The CNN's convolutional layer performs the "convolution" process. The layer inputs are subjected to these convolutional layer kernels. A feature map is produced by convolving all of these outputs from convolutional layers. Since images are by their very nature nonlinear,

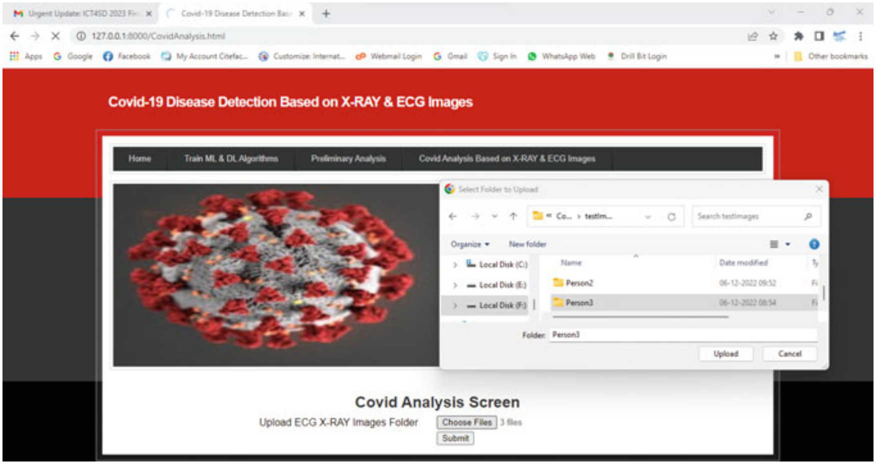


Fig. 6 Uploading "Person 3" folder and below are the prediction output

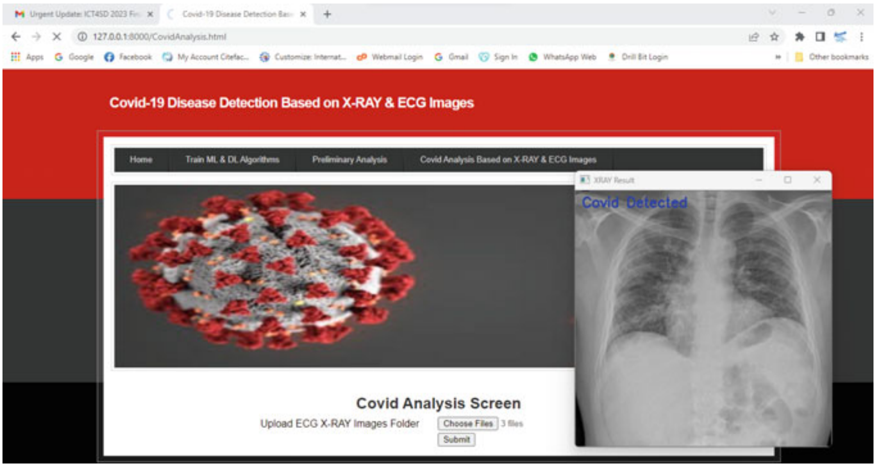


Fig. 7 X-ray "COVID detected" result

this work uses a convolutional layer and the rectified linear unit (ReLU) in the activation functions to enhance the nonlinearity of the input image. As a result, CNN with ReLU is now easier and faster. The output z of the ReLU is "zero" for all negative inputs and "constant" for positive inputs, as shown in Fig. 9.

The subsampling layer process, also known as pooling, is also required by CNN. Each feature map from a convolution layer is individually processed by the pooling layer, by reducing overfitting and extracted features, restoring the pertinent features, and decreasing the feature map. The CNN pooling can be total, average, or maximal. Because other researchers might not be able to recognise acute characteristics, max

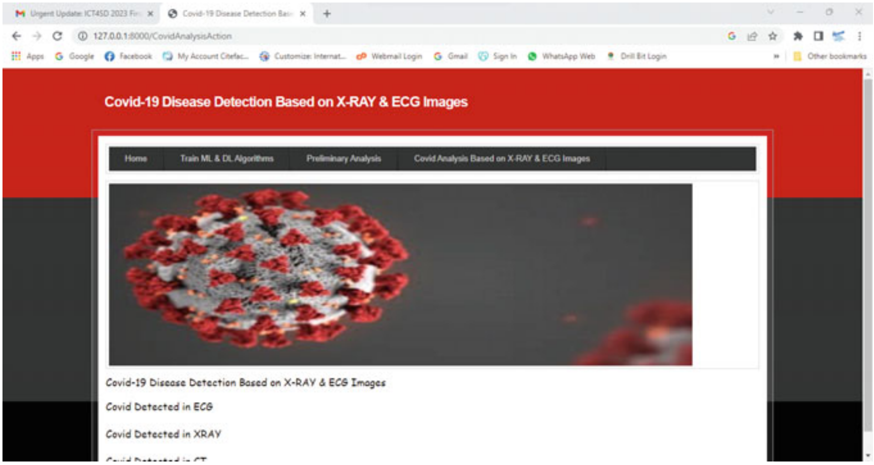
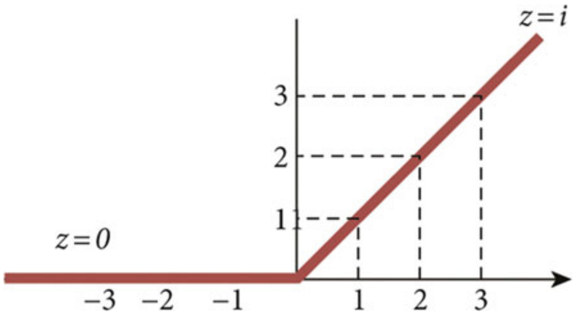


Fig. 8 Final output display

Fig. 9 Function demonstration for ReLU



pooling was used in this investigation. This study used the batch normalisation layer to train a very deep neural network. In order to standardise the input layer and speed up hidden unit learning, scaling and activation are adjusted. Dropping randomly chosen neurons during training decreases over fitting with the 20% dropout layer. To convert the output of the convolutional layer into a one-dimensional feature vector, the CNN utilised in the study comprises a flattening layer at the very end. The convolutional layer's pixel data is organised into a vector by the flattening layer. After being flattened, vector data is supplied to the dense or completely linked layers of CNN. In an entirely linked layer, every neuron is intimately attached to the one above it. Dense layers classify the image by class label using the flattened output from convolution and pooling layers. The class likelihood of a feature is shown by flattened feature set numbers. Based on these probabilities, the fully connected, densely coupled network categorises itself.

4.2 CNN's Picks for Interior Design

The 38 layers of the suggested CNN model include a flattening layer, three fully connected layers, eight activation function levels, eight batch normalisation layers, six convolutional (Conv2D) layers, six max pooling layers, and six dropout layers. A 150 by 150 RGB image with the coordinates 150, 150, 3 is the input image. The filter size increases after every two Con2D layers, while the kernel size remains constant at 3. Con2D uses 64 filters for its first and second layers, 128 filters for its third and fourth layers, 256 filters for its fifth and sixth layers, and so on. The 20% dropout layer, batch normalisation layer, activation layer using the ReLU function, and max pooling layer with a 2:2 pooling size come after each Con2D layer. The layers for maximum pooling, batch normalisation, activation, and dropout come after the final Con2D layer. The 256 output neurons are managed by these layers. The last pooling and convolutional layer produces a three-dimensional matrix, which is subsequently converted into a vector for three dense layers by a flattening layer.

CNN uses the binary cross-entropy (BCE) loss function for binary classification. As binary classification only needs one output node to categorise data into one of two classes, the BCE loss function feeds the output value to a sigmoid activation function. The sigmoid activation function yields results between 0 and 1. Calculated is the projected-to-actual class discrepancy. The "Adam" optimiser reduces learning model loss by changing attribute weight and learning rate. Figure 10 shows the model's design, and Table 2 lists the model's parameters. The CNN was utilised in early experiments in a number of different convolution layer configurations. The model's convolution layers are generated in succession. The CNN was tested and examined using just one convolutional layer. A two-layer CNN was built, results were analysed, etc. Iterations of the process were carried out until the model's outputs were reliable and beneficial. The model's output showed six convolutional layers. The outcome of each model increment is contained in a result.

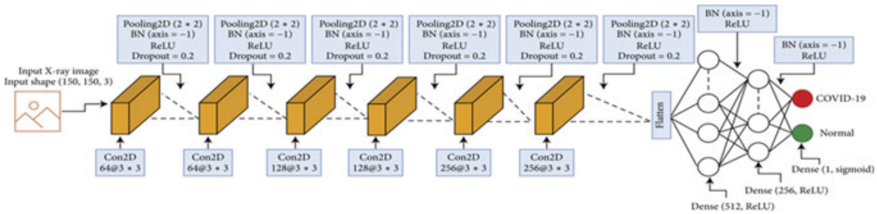


Fig. 10 Architecture of the CNN model

Table 2 Parameter and its values

Parameters	Value
Activation functions	ReLU, sigmoid
Input dimension	(150, 150, 3)
Dropout rate	20%
Max pooling	2×2
Filter to learn	64, 128, 256
Batch normalisation	Axis = -1
Loss function	Binary cross-entropy
Kernel size	3×3
Optimizer	Adam
Epochs	50

5 Conclusions

There were 900 X-ray pictures after pre-processing. The suggested CNN was trained and tested on two subsets. There were 800 X-ray pictures in the training dataset: 400 COVID-19 and 400 normal. The testing dataset also contains 100 X-rays, including 50 from the COVID-19 positive and normal classes. 800 X-ray pictures from a 25% validation subset were input to the model. In order to train the model, 600 X-ray images are used, and 200 are used to validate it. The proposed CNN model contained 38 layers, including 3 fully connected, 6 convolutional, 6 maximum pooling, 6 dropout, 8 activation functions, 8 batch normalisations, and 6 dropout layers. The CNN model produced a precision of 1.0 and using the test data subset from the dataset analysed for this work and the model limitation values indicated in Table 3, an accuracy of 100% was achieved. Accuracy, recall, precision, FSCORE, and false prediction rate were used in this study to evaluate performance. Table 3 displays the parameter results.

Table 3 Test data model performance

Evaluation parameters	%
Accuracy	90.28
Recall	80.36
Precision	90.54
FSCORE	84.05
False prediction rate	9.72

Fig. 11 Primary/main confusion matrix of CNN's model

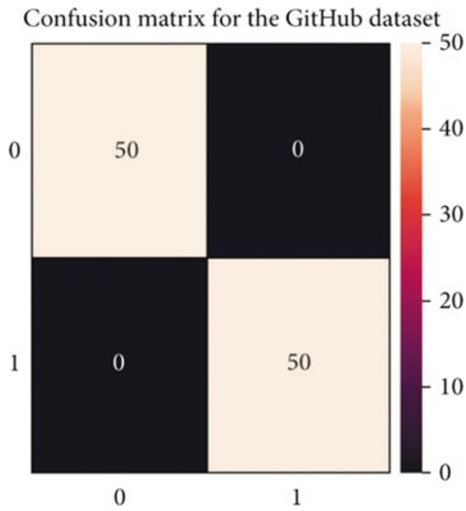


Fig. 12 CNN model training and testing accuracy plot achieved by the

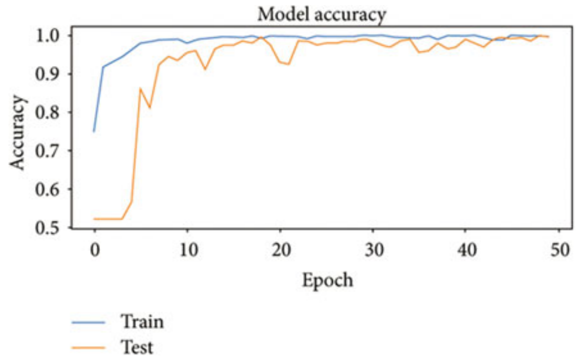
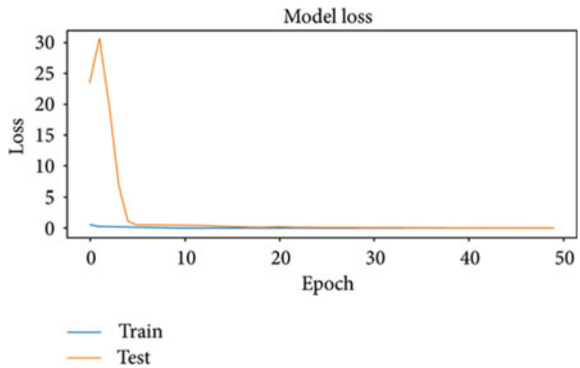


Fig. 13 CNN model training and testing loss plot by the



The matrix of model confusion is shown in Fig. 11. Training and testing accuracy and loss curves are displayed in Figs. 12 and 13. According to the confusion matrix, the CNN model test includes 50 X-rays from the primary dataset, 50 of which are COVID-19 images and 50 of which are normal.

The confusion matrix in Fig. 11 demonstrates that the CNN model correctly predicts each of the 100 images. Accuracy of CNN training and validation is shown as a graph in Fig. 12. The orange curve represents validation accuracy, while the blue curve represents CNN training accuracy. Figure 12 shows that after 5 and 25 epochs, the CNN's training and validation accuracy are still constant. Losses from CNN training and validation are shown in Fig. 13. CNN's training loss is low right away, but its validation loss doesn't become any better until five epochs have passed. The CNN model used in this investigation operates as shown above.

The accuracy charts for CNN model training and testing are displayed in Fig. 12. The suggested CNN model exercised and testing losses is shown in Fig. 13. Because the training loss is 31 in the first epoch, 0.9 after five epochs, 0.0011 after 23 epochs, and 0.000058 in the final epoch, Fig. 13 demonstrates that the proposed CNN converges quickly.

In order to evaluate the model's proficiency, K fold cross-validation was done. Tenfold cross-validation was employed in this investigation. After 10 iterations, the results receive a score of 99.67% (0.15%).

Using COVID-19 Independent Validation Data, CNN tested.

To evaluate the trained model's performance in COVID-19 classification from X-ray pictures, an autonomous dataset was used [22]. 100 COVID-19 X-rays were used in the independent testing. 100 typical images were used for testing. The same model performs well, as evidenced by its 0.995 accuracy, 1.000 precision, and other demonstrated performance metrics.

Figure 14 display the $N \times N$ confusion matrix for the CNN model on an autonomous test dataset from IEEE Data Port. According to the confusion matrix,

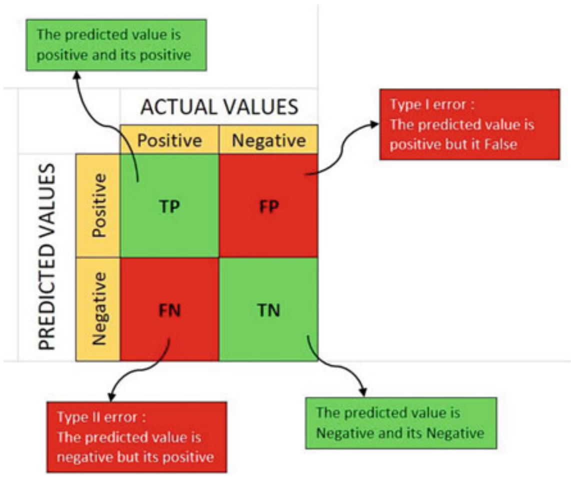


Fig. 14 $N \times N$ matrix of confusion

Fig. 15 Binary classification problem (2 × 2 matrixes)

		Predicted	
		Negative **(N)** -	Positive **(P)** +
Actual	Negative -	True Negative **(TN)**	**False Positive (FP) Type I Error**
	Positive +	**False Negative (FN) Type II Error**	True Positive **(TP)**

both target classes generate an equal amount of images. One COVID-19 image is misclassified by the model (Fig. 15).

The CNN model test takes 50 X-rays from the primary dataset, 50 of which are COVID-19 images and 50 of which are normal, in accordance with the confusion matrix. The number of things a classifier properly and erroneously classifies is shown in a table called a "confusion matrix". It is used to evaluate a classification model's efficacy. It can be used to determine performance metrics such as accuracy, precision, recall, and FSCORE to evaluate the efficiency of a classification model.

Confusion matrices are often used because they give a more accurate measure of how well a model works than classification accuracy. There is no data on how many instances were wrongly labelled in the classification accuracy study. Imagine your data is split into two classes, with 15% of the data being in class B and 85% of the data being in class A. Imagine that all instances of class A are accurately classified by your classification model, but all instances of class B are wrongly classified. The model's accuracy in this case is 85%. Therefore, it is sad that Class B is misclassified. The confusion matrix, on the other hand, gives a clearer insight into your classifier's performance by showing examples that were correctly and incorrectly classified for each class.

There are two methods for evaluating a model's accuracy. Accuracy can be defined as the percentage of numbers that were correctly predicted.

Confusion Matrix for Understanding: We can identify the measures we're seeking by using the basic phrases given below. When both the actual and predicted numbers are positive, this is known as a "true positive" (TP). When the prediction and the actual number are both negative, this is referred to as a "true negative" (TN).

- False positives (FP): when the prediction is accurate but the outcome is unfavourable. Referred to as the Type 1 error as well.
- False negatives (FN) are predictions that turn out to be positive when they should have been negative. It is also known as a Type 2 error.

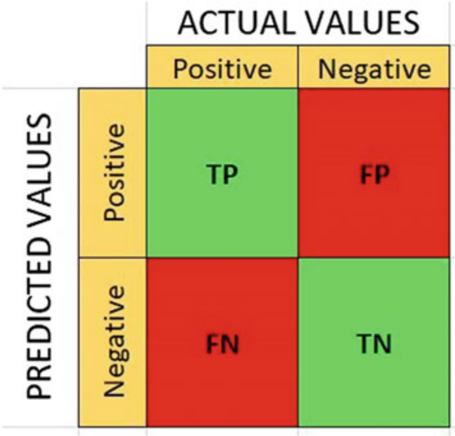


Fig. 16 Confusion matrix for the binary classification

A 2 × 2 matrix with 4 values is what we would have for a binary classification problem:

- The target variable can be given either positive or negative values. Actual values for the target variable are displayed in the columns, while projected values are displayed in the rows (Fig. 16).

Fig. 17 Wrongly classified X-ray image

A mislabelled image was evaluated by experts. The X-ray came from an early COVID-19 patient. Clear patterns that could have been seen in standard X-ray scans are not present in the image. Figure 17 displays the misidentified COVID-19-positive X-ray from the model.

This comparison and the CNN model's overall performance demonstrate that COVID-19 might be diagnosed using X-ray image categorization.

Figure 18 compares CNN with several machine learning techniques using two bar graphs. The test data bar graph represents the results of the first test scenario, which evaluated each model on a portion of the dataset. The findings of the second test scenario, in which each model was evaluated using a separate test set, are displayed in the independent validation data bar graph.

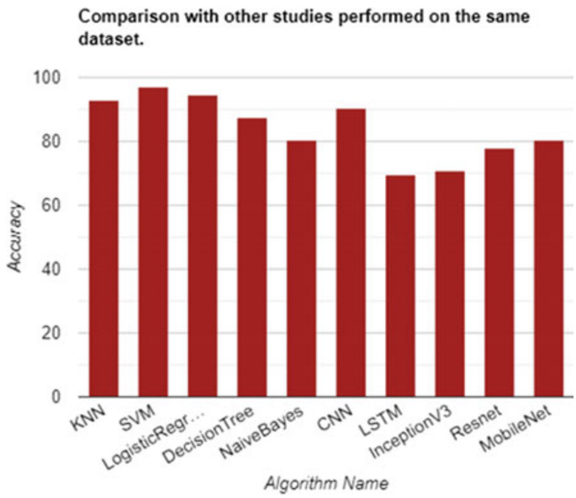


Fig. 18 Comparison with related research using the same dataset

Table 4 Comparison with related research using the same dataset

Algorithm name	Accuracy	Precision	Recall	FSCORE	False prediction rate
KNN	93.06	86.84	95.69	90.17	6.94
SVM	97.22	98.33	92.86	95.31	2.78
Logistic regression	94.44	91.13	91.13	91.13	5.56
Decision tree	87.50	79.69	86.82	82.31	12.50
Naive Bayes	80.56	73.91	85.22	75.77	19.44
CNN	90.28	90.54	80.36	84.05	9.72
LSTM	69.44	44.53	46.88	45.00	30.56
InceptionV3	70.83	50.00	50.00	49.18	29.17
ResNet	77.78	38.89	50.00	43.75	22.22
MobileNet	80.56	90.00	56.25	55.56	19.44

The CNN-based method that was suggested was compared to other recent academic works to see how well it worked for the current application. Research findings and urban tales are contrasted in Table 4.

Datasets get trained on various machine and deep learning algorithms such as KNN, SVM, Logistic Regression, Decision Tree, Naive Bayes, CNN, LSTM, InceptionV3, ResNet, and MobileNet.

References

1. Cucinotta D, Vanelli M (2020) WHO declares COVID-19 a pandemic. Acta Biomed Atenei Parm 91:157–160
2. Rustam F, Reshi AA, Mehmood A et al (2020) COVID-19 future forecasting using supervised machine learning models. IEEE Access
3. Cennimo DJ (2020) Coronavirus disease 2019 (COVID-19) clinical presentation, vol 8, pp 101489–101499 [Online]. https://emedicine.medscape.com/article/2500114-clinical#b2
4. X-Ray (Radiography)-Chest (2020) https://www.radiologyinfo.org/en/info.cfm?pg=chestrad#overview
5. Reshi AA, Rustam F, Mehmood A, Alhossan A, Alrabiah Z, Ahmad A, Alsuwailem H, Choi GS (2021) An efficient CNN model for COVID-19 disease detection based on X-ray image classification. In: Complexity and robustness trade-off for traditional and deep models, vol 2021, pp 1–12. Article ID 6621607
6. Cohen JP (2020) Github Covid19 X-ray dataset [Online]. https://github.com/ieee8023/covid-chestxray-dataset
7. Lundervold AS, Lundervold A (2019) An overview of deep learning in medical imaging focusing on MRI. Z Med Phys 29(2):102–127
8. Ahmad M (2021) Ground truth labeling and samples selection for hyperspectral image classification. Optik 230. Article ID 166267
9. Kayalibay B, Jensen G, van der Smagt P (2017) CNN-based segmentation of medical imaging data. http://arxiv.org/abs/1701.03056
10. Li Q, Cai W, Wang X, Zhou Y, Feng DD, Chen M (2014) Medical image classification with convolutional neural network. In: Proceedings of the 2014 13th international conference on control automation robotics and vision (ICARCV). Singapore, pp 844–848
11. Umer M, Sadiq S, Ahmad M, Ullah S, Choi GS, Mehmood A (2020) A novel stacked CNN for malarial parasite detection in Thin blood smear images. IEEE Access 8:93782–93792
12. Rouhi R, Jafari M, Kasaei S, Keshavarzian P (2015) Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Syst Appl 42(3):990–1002
13. Sharif M, Attique Khan M, Rashid M, Yasmin M, Afza F, Tanik UJ (2019) Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. J Exp Theoret Artif Intell 1–23
14. Asada N, Doi K, MacMahon H et al (1990) Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: pilot study. Radiology 177(3):857–860
15. Katsuragawa S, Doi K (2007) Computer-aided diagnosis in chest radiography. Comput Med Imaging Graph 31(4–5):212–223
16. Esteva A, Kuprel B, Novoa RA et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118
17. Dong Y, Pan Y, Zhang J, Xu W (2017) Learning to read chest X-ray images from 16000+ examples using CNN. In: Proceedings of the 2017 IEEE/ACM international conference on connected health: applications, systems and engineering technologies (CHASE). Philadelphia, PA, USA, pp 51–57

18. Dong D, Tang Z, Wang S et al (2020) The role of imaging in the detection and management of COVID-19: a review. IEEE Rev Biomed Eng 14:16–19
19. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43:635–640
20. Abbas A, Abdelsamea MM, Gaber MM (2020) Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. http://arxiv.org/abs/2003.13815
21. Mane DT, Kulkarni UV (2017) A survey on supervised convolutional neural network and its major applications. Int J Rough Sets Data Anal 4(3):71–82
22. Chen ZH (2020) Mask-RCNN detection of COVID-19 pneumonia symptoms by employing stacked autoencoders in deep unsupervised learning on low-dose high resolution CT. IEEE Dataport

Performance Analysis of 5G NR Cell Using Classical Scheduling Algorithms and Different MIMO

Rajesh Kumar, Deepak Sinwar, and Vijander Singh

Abstract 5G NR has been deployed in most of the countries and covers three major use case scenarios, viz. enhanced mobile broadband (eMBB), ultra-reliable low latency communication (URLLC), and massive machine type of communication (mMTC). To enable the 5G services on the existing 4G infrastructure, the non-standalone architecture (NSA) of 5G NR is equally popular. For NSA deployment, many countries are utilizing the 3.5 GHz band on time division duplex (TDD) transmission. For both uplink and downlink transmissions, TDD is utilizing same frequency bands and scheduling of resources in such environments is quite challenging. Here, the performance of 5G NR is analysed on 20 MHz and 100 MHz channel bandwidth in terms of resource utilization by comparing the performance of different scheduling algorithms, i.e. Professional Fair, Round Robin, and Best CQI on different MIMO (64×64, 32×32, 16×16, 8×8) set-up and block error rates. The performance of Best CQI algorithm is found to be better than other scheduling algorithms under considerations.

Keywords New radio · Time division duplex · Best CQI · Professional Fair · Round Robin · 3.5 GHz band · MIMO

1 Introduction

The fifth generation (5G) of wireless technology indicates a significant leap forward in mobile connectivity, promising to deliver faster speeds, lower latency, and greater capacity than previous generations. To achieve these ambitious goals, 5G introduces

R. Kumar · D. Sinwar (✉)
Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, Rajasthan 303007, India
e-mail: deepak.sinwar@gmail.com

V. Singh
Department of ICT and Science, Faculty of Information Technology and Electrical Engineering, NTNU–Norwegian University of Science and Technology, Ålesund, Norway
e-mail: vijander.singh@ntnu.no

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_34

a new air interface called New Radio (NR), which operates in both sub-6 GHz and millimetre-wave (mmWave) frequency bands [1]. It is difficult to assure the performance of 5G NR to the end users because of several key factors (i.e. antenna design, modulation and coding schemes, scheduling algorithms, cell performance, radio frequency planning, etc.) that are responsible to determine quality of services (QoS) [2].

The communication in 5G NR can be downlink and uplink. Downlink communication refers to the communication from gNB to user equipment (UEs), whereas the uplink communication is directed from UEs to gNB. The downlink in 5G NR is designed to deliver ten times more data rate and reliable connectivity than 4G Networks. 5G NR leverages advanced technologies such as massive multiple-input multiple-output (MIMO), beamforming, and dynamic spectrum sharing to optimize the downlink performance and improve network capacity and coverage [3]. Downlink communication in 5G NR operates in both sub-6 GHz and millimetre-wave (mmWave) frequency bands, enabling it to support a wide range of use cases and deployment scenarios, viz. enhanced mobile broadband (eMBB), Ultra-Reliable low latency communication (URLLC), and massive machine type of communication (mMTC). However, optimizing the downlink performance in 5G NR requires careful planning and design [4]. To maximize the network capacity and coverage, MIMO technology plays an important role to allow transmission of multiple data streams simultaneously using multiple antennas. MIMO improves the performance of uplink and downlink transmission in 5G NR by enhancing signal quality, reducing interference and increasing the throughput. However, to utilize the full benefits of MIMO, it is essential to use an effective scheduling algorithm for intelligent allocation of resources to the different users based on the channel quality and QoS requirements [5]. A scheduling algorithm determines how network resources, such as time, frequency, and power, are allocated to different users at different times in the resource grid. These resources are based on various factors such as user priority, traffic demand, and channel quality.

The design of scheduling algorithm in 5G NR should be flexible and adaptive that support a wide range of use cases and deployment scenarios. Because of these different use case scenarios, the scheduling algorithm must update the resource grid as per the user's requirements. The scheduling algorithm uses a combination of channel state information (CSI), beamforming, and MIMO precoding to optimize the allocation of resources and maximize the network performance. It is quite challenging that the scheduling algorithm in 5G NR must achieve high network efficiency, low latency, and high throughput. It must be adaptable in varying network conditions, user requirements, and traffic demands to ensure that network resources are allocated in the most efficient and effective manner [6]. Several scheduling algorithms are available in 5G NR such as Best CQI, Professional Fair, and Round Robin.

Best CQI (Channel Quality Indicator) scheduling algorithm selects the user with the highest CQI value and allocates resources to that user. CQI is a measure of the channel quality between the user equipment (UE) and the base station (BS), and a higher CQI value indicates a better channel quality. Best CQI scheduling algorithm is suitable for enhanced mobile broadband (eMBB) use cases where high data rates are required. Professional Fair (PF) scheduling algorithm considers both the channel quality and the amount of data that needs to be transmitted to each user. It aims to achieve fairness among users while maximizing network throughput. PF scheduling algorithm is suitable for use cases where fairness is important, such as video streaming and voice over IP (VoIP). Round Robin (RR) scheduling algorithm allocates resources to users in a cyclical manner, ensuring that each user gets an equal share of network resources [7]. RR scheduling algorithm is suitable for use cases where fairness and equal distribution of resources are essential, such as Internet of Things (IoT) devices and massive Machine Type Communications (mMTC).

In terms of performance, Best CQI scheduling algorithm can achieve high data rates and high throughput, but it may not achieve fairness among users. PF scheduling algorithm can achieve fairness among users, but it may not fully utilize the network resources. RR scheduling algorithm can achieve fairness and equal distribution of resources, but it may not fully utilize the network resources or provide high data rates [8]. In summary, the choice of scheduling algorithm in 5G NR depends on the specific use case and network requirements. Best CQI scheduling algorithm is suitable for eMBB use cases where high data rates are required, while PF scheduling algorithm is suitable for use cases where fairness is important. RR scheduling algorithm is suitable for use cases where fairness and equal distribution of resources are essential. Here, the problem of identifying best scheduling algorithm under varying MIMO schemes is analysed on different channel bandwidths. In short, the major contribution of the work as mentioned as follows:

- Experimental analysis of different scheduling algorithms such as Best CQI, Professional Fair, and Round Robin.
- Performance analysis of three well-known scheduling algorithms, i.e. Best CQI, PF, RR using different MIMO schemes (64×64, 32×32, 16×16, 8×8) on 20 MHz and 100 MHz channel bandwidths.

In short, the rest of the paper organized as follows: Sect. 2 will discuss related works in terms of cell performance of 5G NR. Section 3 will describe the system model and experimental set-up for evaluating the cell performance of 5G NR. Section 4 will present the results and discussion, followed by Sect. 5, which will conclude the paper.

2 Literature Review

This section highlights various extensive literature utilizing below sub-6 GHz bands and numerous simulation parameters during deployment of 5G services. The deployment of 5G is the essential and most difficult task in the practical scenario. There are lot of factors that need to be considered for the deployment of 5G such as population density, existing telecom infrastructure and cost. Oughton et al. [9] presented an insight to the capacity, coverage, and cost of different 5G infrastructure strategies in the Netherlands. They discussed the modelling approach related to various factors and deployments strategies, viz. (1) upgrading existing 4G infrastructure, (2) deploying 5G infrastructure in high-frequency bands, (3) deploying 5G infrastructure in low-frequency bands, and (4) a hybrid strategy that combines the deployment of 5G infrastructure in both high- and low-frequency bands. The results show that the hybrid strategy offers the best combination of coverage, capacity, and cost. On the other hand, Moses and Moses [10] discussed the design and performance of compact MIMO antennas for 5G handheld devices. They addressed the challenges of designing MIMO antennas for compact devices and proposes a new approach to achieve self-decoupling and good isolation between the antenna elements. The new approach is based on the usage of a T-shaped parasitic element. The simulation results demonstrated high antenna efficiency and a low correlation coefficient between the elements. The MIMO technique is an important pillar to gain higher throughput among users. Therefore, Lun et al. [11] highlight the challenges of providing high-speed broadband connectivity to rural areas using 5G technology. The Deployment of the C-Band is equally popular as using 3.5 GHz band carrier frequency for this purpose. The C-band offers a balance between coverage and capacity. The authors also highlight unique characteristics of rural areas such as low population density and challenging terrain. On the other hand, 3.5 GHz frequency band is equally popular in private area networks. Esa et al. [12] discussed on the planning of 5G networks in industrial areas using the 3.5 GHz frequency band. They examine the advantages of the 3.5 GHz frequency band, including its balance between coverage and capacity in machine types of communication. They also highlight the challenges of deploying 5G networks in industrial areas related to gaining higher reliability and low latency communication in Indonesia.

The 3.5 GHz frequency band utilizes different antenna diversity techniques (i.e. spatial diversity, polarization diversity, and pattern diversity) as compared to LTE infrastructure. In this regard, Alimi et al. [13] discussed on the usage of antenna diversity techniques to improve the performance of 5G fixed wireless access networks. The simulation scenario demonstrates that antenna diversity techniques can significantly improve the performance of 5G fixed wireless access networks, particularly in terms of coverage, capacity, and reliability. On the other hand, Ahmad et al. [14] spotlight on the design and experimental analysis of a reconfigurable 5G antenna for sub-6 GHz wireless applications. They discuss the challenges of designing antennas that can operate on multiple frequency bands for 5G networks which require high data rate and low latency communication. The simulated results demonstrate the performance of the antenna across multiple frequency bands with low returns loss and high radiation efficiency. Overall, the study provides valuable insights into the design and experimental analysis of reconfigurable 5G antennas for sub-6 GHz wireless applications. Hikmaturokhman et al. [15] presented similar kind of comparative analysis with varied frequency and three frequency bands, viz. 700 MHz, 3.5 GHz, and 26 GHz in the context of Jakarta. The analysis includes the evaluation of path loss, delay spread, and coherence bandwidth for each frequency band using different channel models. The simulation results demonstrate that the choice of channel model and frequency band significantly affects the performance of 5G networks in terms of coverage, capacity, and reliability. They also suggest that the varied frequencies and channel models are important contexts while optimization of the 5G network.

It is evident that the identification of 5G band is an important aspect for the initial deployment of 5G. In this regard, Ekawibowo et al. [16] focused on identifying the suitable frequency bands for the initial deployment of 5G technology in Indonesia. They analyse various 5G bands in context of spectrum availability, regulatory framework, and ecosystem readiness in terms of technical and commercial feasibility. The simulation results demonstrate that the mid-band spectrum, specifically the 2.3 and 2.6 GHz bands, is the most suitable for the initial deployment of 5G technology. They stated that the mid-band spectrum offers a balance between coverage and capacity, and its availability for 5G technology deployment is relatively high.

Technical and commercial feasibility is an important context that needs to be considered during deployment of the 5G services. Sheikh et al. [17] discussed on the analysis of indoor solutions for providing coverage at the 3.5 and 28 GHz frequency bands for 5G systems. The simulation results demonstrate that indoor solutions, such as small cells and millimetre-wave indoor systems, are the most suitable for providing coverage at the 3.5 and 28 GHz frequency bands for 5G systems. They shortlist the frequency band while considering the coverage requirements, deployment scenarios, and technical feasibility.

3 System Model and Experimental Set-Up

To analyse the cell performance of 5G NR using classical scheduling algorithms, a series of experiments was conducted. This section discusses not only the system model but also the simulation environment specifically with respect to types of system model and simulation parameters. To evaluate the network performance, we have utilized a system model of 5G NR with different MIMO antenna configurations. The model is configured to estimate the channel performance of 5G NR in the downlink direction using channel state information reference signals (CSI-RS). The MIMO scheduling decision is made by gNB based on these characteristics. Here, spatial multiplexing technique applies to perform multi-layer transmission that limits the number of transmit and receive antennas [18] as shown in Fig. 1.

Furthermore, the precoding technique matches the transmission layer to the physical antenna ports. The gNB calculates the CSI-RS based on the channel condition and sends it to the UEs [18] as shown in Fig. 2. The CSI is the important parameter in which gNB calculates the PMI and CQI values. Based on these parameters, gNB scheduler chooses the number of DL transmission layer, the precoding matrix, and the selection of the modulation and coding scheme.

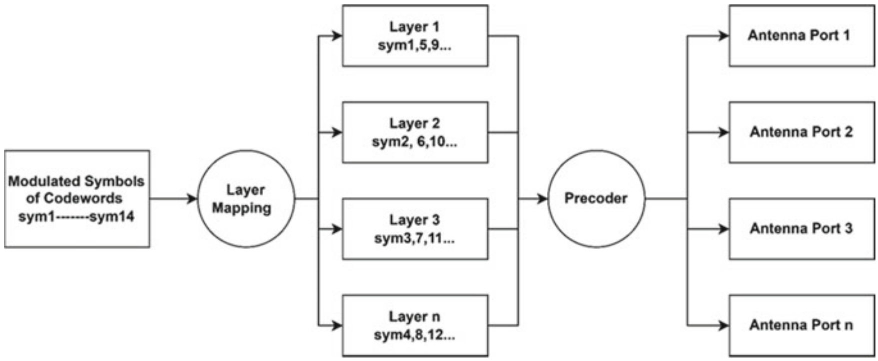


Fig. 1 Spatial multiplexing and precoding technique of 5G NR

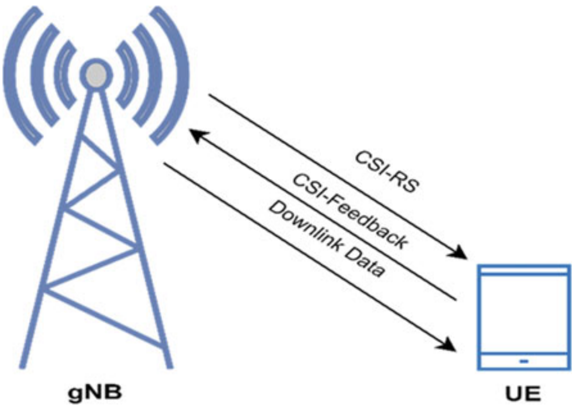


Fig. 2 Channel state information reporting by gNB

3.1 Experimental Set-Up

Based on this system model, an experimental set-up was created on MATLAB 2021a to evaluate the cell throughput with four UEs. Experimental studies were conducted on 20 MHz and 100 MHz channel bandwidths with different MIMO set-up (8 × 8, 16 × 16, 32 × 32, 64 × 64) at the transmitter and receiver end. Here, 64 × 64 MIMO indicates, 64 transmitter elements, and 64 receiver elements. The simulation parameters are given in Table 1. The experimental set-up consists of several scenarios with different values of scheduling algorithms, MIMO set-up, CSI row number, sub-band size, and channel bandwidth.

Table 1 Simulation parameters considered during performance analysis of 5G NR cell

Parameter name	Value
Number of frames	5
Scheduling type	Symbol-based
Maximum coverage of gNB	1500 m
Number of RBs	51 and 273
SCS	30
Carrier frequency	3.5 GHz
CSI row number	8, 11
Panel dimension	[8 1]
Scheduling algorithm	Best CQI, RR, PF
Channel model	CDL-C
Slot periodicity and offset	[10 2]
PMI and CQI mode	Sub-band
Sub-band size	8 and 16
MIMO (Tx and Rx)	[8 × 8, 16 × 16, 32 × 32, 64 × 64]

4 Results and Discussion

To analyse the performance of 5G NR cell, a series of experiments was conducted based on simulation parameters as discussed in the previous section. MIMO permits the usage of multiple antenna technology to simultaneously transfer and receive the data stream. Here, the performance of 5G NR cell is analysed on four performance metrics, viz. peak throughput, achieved average downlink spectral efficiency, achieved average downlink throughput, and block error rate (BLER) as follows.

4.1 Peak Throughput

As shown in Fig. 3, MIMO technology in 5G NR significantly improves the performance. For downlink transmission, a significant improvement in the peak throughput can be observed in case of a 64 × 64 MIMO as compared to other MIMO (i.e. 32 × 32, 16 × 16, and 8 × 8) in both 20 MHz and 100 MHz bands.

It can be viewed from Fig. 3 that 64 × 64 MIMO provides almost double peak throughput as compared to the 32 × 32 MIMO. In the same way, peak throughput of 16 × 16 MIMO is found near double as compared to the 8 × 8 MIMO. By increasing the transmitter and receiver elements in MIMO, a significant increase in the signal quality can be observed with reduced interferences. In addition, no impact of scheduling algorithm was observed in case of peak throughput for the

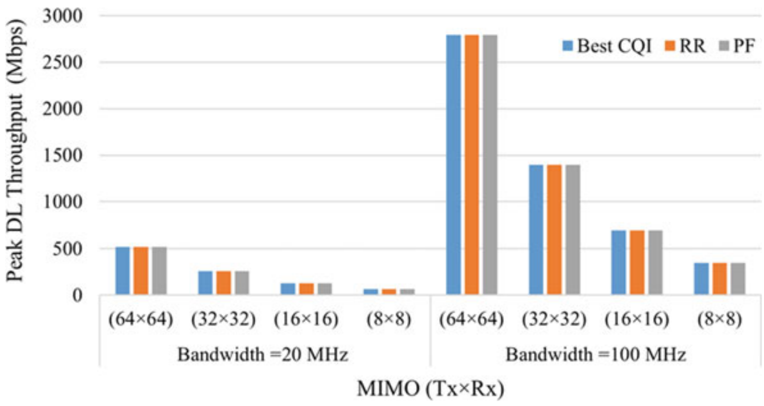


Fig. 3 Peak DL throughput (20 and 100 MHz bandwidth)

same scenario as shown in Fig. 3. In 5G NR, the peak throughput is found to be directly proportional to the channel bandwidth.

4.2 Achieved Average Downlink Throughput and Spectral Efficiency

To effectively utilize the resources as per QoS requirements, MIMO supports effective beamforming and precoding techniques. It also improves the diversity of the system, reduces the effects of fading, and improves the SNR value that leads to higher throughput gain. However, the MIMO capacity can also be measured according to effective spectrum utilization. Spectral efficiency can be measured in terms of amount of information that can be transmitted over a given amount of radio frequency. To validate the same, we evaluated the achieved average downlink spectral efficiency based on simulation parameters as depicted in Table 1. Figure 4a, b represents the achieved average downlink throughput and achieved average spectral efficiency.

If the spectral efficiency is higher, a higher data rate can be achieved, but it can be affected by the number of factors such as modulation and coding scheme, bandwidth, number of antennas, and interference level. Figure 4 shows that the achieved average spectral efficiency is greatly impacted in 100 MHz channel. Several reasons, i.e. such as interference, higher noise level, overhead, nonlinear distortion, and channel coherence time are contributing towards lower spectral efficiency. However, the wide channel should provide higher data rate, but it brings a new challenge of lower spectral efficiency. On the other hand, as shown in Fig. 4c, d, the achieved average throughput increases as the channel bandwidth increases. But in case 32 × 32 MIMO, the performance of 16 × 16 MIMO is found almost equal. Theoretical analysis indicated that several factors such as increasing complexity and inter-antenna spacing are key contributors for the same. The 32 × 32 MIMO set-up has large antenna which can

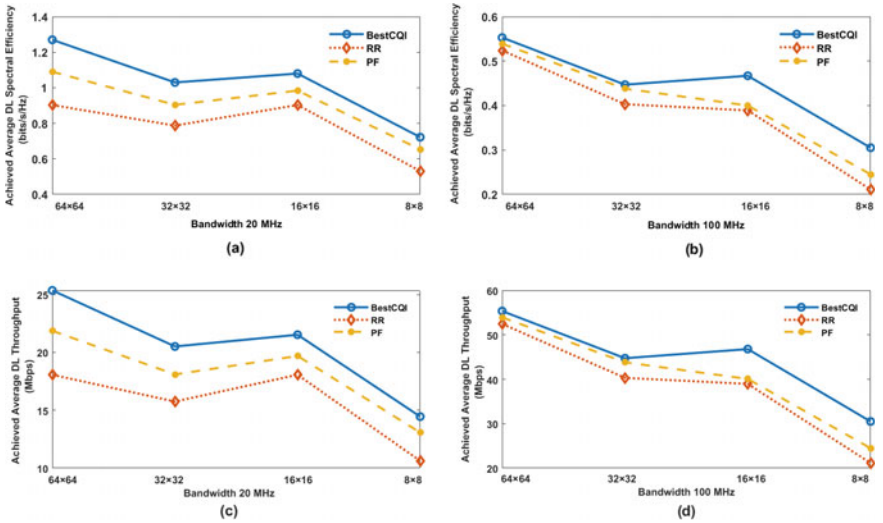


Fig. 4 Achieved average downlink spectral efficiency and achieved average downlink throughput with different MIMO setup on 20 MHz and 100 MHz channel bandwidths

result in greater spacing between antennas, as the higher antennas correlation, leading to reduced spatial diversity and lower performance. It can be viewed from Fig. 4 that the performance of Best CQI is better than RR and PF scheduling algorithms.

4.3 Block Error Rate (BLER)

BLER is one of the important metrics that calculate the number of erroneous blocks with respect to number of transmitted blocks. A higher BLER indicates lower quality of the radio link. Based on simulation parameters as depicted in Table 1, we calculated the BLER for four users on different MIMO and scheduling algorithms as shown in Fig. 5. As per the performance of scheduling algorithms is concerned, the Best CQI outperforms other two algorithms under consideration. As shown in Fig. 5, Best CQI algorithm has a lower BLER value than RR and PF algorithms.

Based on the experimental evaluations, it can be stated that the performance of Best CQI algorithm is found to be better than RR and PF. A sample downlink throughput, goodput, resource share, and buffer status of Best CQI under 20 MHz channel bandwidth are shown in Fig. 6. The buffer status represents the amount of data waiting that needs to be transmitted by the UEs. If a UE with high CQI has a low buffer status, it may not be able to fully utilize the allocated resources, thereby leading to suboptimal system throughput. Similarly, if a UE with low CQI has a high buffer status, may not receive sufficient resources to transmit its data, thereby leading to a lower system throughput.

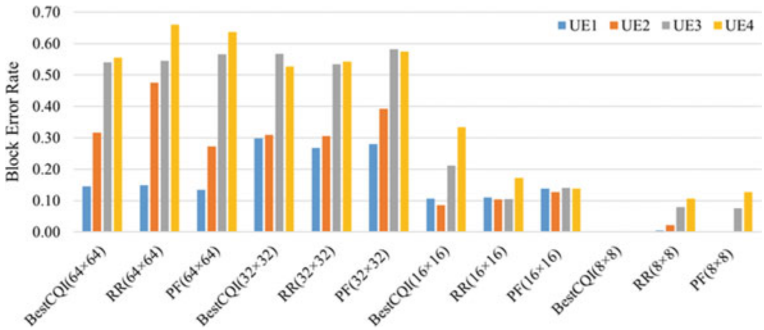


Fig. 5 Average BLER of all users under different MIMO setup

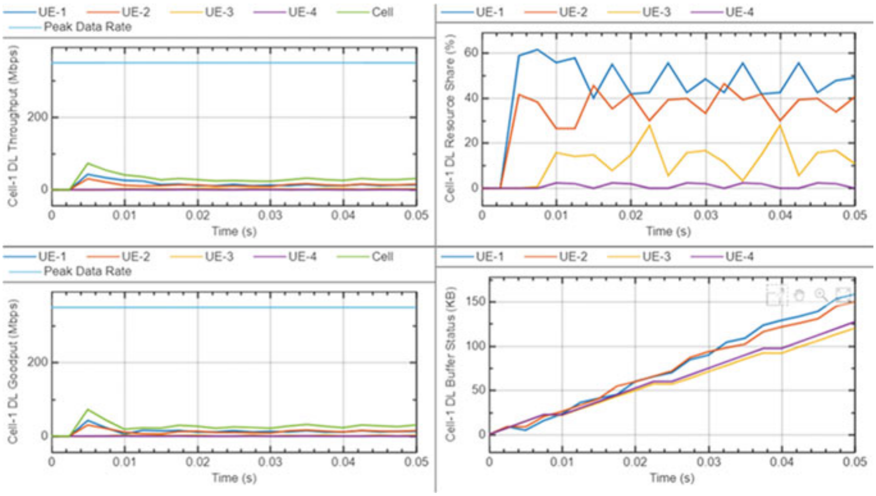


Fig. 6 Best CQI cell performance under 20 MHz channel bandwidth

Overall, the buffer status of a UE can have a significant impact on the performance of scheduling algorithms in 5G NR, as it determines the priority of a UE in receiving resources from the network.

4.4 Discussion

5G NR below 6 GHz band is utilizing the services for different traffic scenarios such eMBB and URLLC, but the major challenge here is the scheduling and allocation of the resources through gNB. As shown experimental results, Best CQI algorithm provides higher throughput as compared to the RR and PF but does not support fairness. The RR algorithm works in a cyclical manner to provide equal opportunities to all, but it does not consider the channel quality. The PF algorithm is the average performer that work on the weight management concept and achieves fairness among users. For effectively scheduling the resources among users, a new scheduling strategy is required that will improve the system throughput. However, there are number of techniques available (i.e. increasing the available bandwidth, choosing higher order modulation scheme, utilizing the advanced MIMO technique, increasing the number of antennas, and optimizing the network parameters) that can be applied to achieve the higher throughput in 5G NR. On the other hand, BLER is one of the important metrics that contribute to the higher throughput in 5G NR. In RR algorithm, resources are allocated in cyclical manner to all the UEs without worried about the channel condition and data rate. BLER of RR algorithm impacts the system throughput by causing delay in the transmission block or increase amount of re-transmission. This phenomenon degraded the user experience. However, in the PF resource are allocated based on the fairness criteria which consider the data rate requirement of each user and channel quality. Higher BLER in PF greatly impact the performance by affecting the channel quality and fairness criteria that leads to less resource suboptimal resource allocation. On the other hand, Best CQI algorithm allocates the resources based on highest CQI value which directly represent the estimated channel quality. In case of higher BLER, the Best CQI algorithm is unable to compute the accurate value of CQI which results in significant reduction in the system throughput. Overall BLER is one of the important metrics that effects the decision of the scheduling algorithm while allocating the resources effectively. It is highly recommended to consider the BLER metrics for scheduling algorithm while implementation of the resource allocation strategy to ensure higher user throughput and user experience.

5 Conclusion

The MIMO technique is a key enabling technology that can achieve higher throughput gains and enable simultaneous data transfer among users in 5G NR. Effective utilization of MIMO can provide opportunities for parallel transmission of data among users, improving cell performance and reliability. Therefore, in this experimental study different MIMO scenarios of 64×64, 32×32, 16×16, and 8×8 were compared using three scheduling algorithms, viz. Best CQI, RR, PF under 20 MHz and 100 MHz channel bandwidths. Experimental results indicated that the highest throughput was achieved with 64×64 MIMO using the Best CQI scheduling algorithm. However, a major challenge was observed in achieving lower downlink achieved spectral efficiency in 100 MHz channel bandwidth with higher MIMO. Based on these experiments, it was observed that the standard scheduling algorithm may not perform well in resource scheduling and cell performance. In future, the focus should be on designing an efficient scheduling algorithm for MIMO that can achieve higher throughput for effective scheduling of resources.

References

1. Attaran M (2021) The impact of 5G on the evolution of intelligent automation and industry digitization. J Ambient Intell Humaniz Comput 0123456789. https://doi.org/10.1007/s12652-020-02521-x
2. Abdel Hakeem SA, Hady AA, Kim HW (2020) Current and future developments to improve 5G-newradio performance in vehicle-to-everything communications. Telecommun Syst 75(3):331–353. https://doi.org/10.1007/s11235-020-00704-7
3. Takeda K, Xu H, Kim T, Schober K, Lin X (2020) Understanding the heart of the 5G air interface: an overview of physical downlink control channel for 5G new radio. IEEE Commun Stand Mag 4(3):22–29. https://doi.org/10.1109/MCOMSTD.001.1900048
4. Fitriasari HI, Lestari AI, Luhurkinanti DL, Sari RF (2021) Performance evaluation of downlink multi-user OFDMA scheduling in 5G new radio (NR). In: ECTI-CON 2021–2021 18th international conference on electrical engineering/electronics, computer, telecommunications and information technology smart electrical system technology proceeding, pp 219–223. https://doi.org/10.1109/ECTI-CON51831.2021.9454694
5. Chen Z, Yuan XT, Ren J, Yuan T (2022) An ultra-wideband MIMO antenna for 5G smartphone. AEU Int J Electron Commun 154:154301. https://doi.org/10.1016/j.aeue.2022.154301
6. Yin H, Zhang L, Roy S (2021) Multiplexing URLLC traffic within eMBB services in 5G NR: fair scheduling. IEEE Trans Commun 69(2):1080–1093. https://doi.org/10.1109/TCOMM.2020.3035582
7. Mamane A, Fattah F, El Ghazi M, Balboul Y, El Bekkali M, Mazer S (2021) Proportional fair buffer scheduling algorithm for 5G enhanced mobile broadband. Int J Electr Comput Eng 11(5):4165–4173. https://doi.org/10.11591/ijece.v11i5.pp4165-4173
8. Li L, Shao W, Zhou X (2021) A flexible scheduling algorithm for the 5th-generation networks. Intell Converg Netw 2(2):101–107. https://doi.org/10.23919/icn.2020.0017
9. Oughton EJ, Frias Z, van der Gaast S, van der Berg R (2019) Assessing the capacity, coverage and cost of 5G infrastructure strategies: analysis of the Netherlands. Telemat Inform 37:50–69. https://doi.org/10.1016/j.tele.2019.01.003
10. Moses ATZ, Moses N (2021) Compact self decoupled MIMO antenna pairs covering 3.4–3.6 GHz band for 5G handheld device applications. In: AEU Int J Electron Commun 141:153971. https://doi.org/10.1016/j.aeue.2021.153971
11. Lun J, Frenger P, Furuskar A, Trojer E (2019) 5G new radio for rural broadband: how to achieve long-range coverage on the 3.5 GHz band. In: IEEE vehicular technology conference, vol 2019, pp 5–10. https://doi.org/10.1109/VTCFall.2019.8891556
12. Rai Nur Esa ARD, Hikmaturokhman A (2020) 5G NR planning at mmWave frequency: study case in Indonesia industrial area. In: Proceeding—2020 2nd international conference on industrial electrical and electronics. ICIEE 2020, pp 205–210. https://doi.org/10.1109/ICIEE49813.2020.9277451
13. Alimi IA, Patel RK, Muga NJ, Monteiro PP (2020) Performance analysis of 5G fixed wireless access networks with antenna diversity techniques. Wirel Pers Commun 113(3):1541–1565. https://doi.org/10.1007/s11277-020-07399-8
14. Ahmad I, Dildar H, Khan WUR, Shah SAA, Ullah S, Ullah S, Umar SM, Albreem MA, Alsharif MH, Vasudevan K (2021) Design and experimental analysis of multiband compound reconfigurable 5G antenna for sub-6 GHz wireless applications. Wirel Commun Mob Comput 2021. https://doi.org/10.1155/2021/5588105
15. Hikmaturokhman A, Suryanegara M, Ramli K (2019) A comparative analysis of 5G channel model with varied frequency: a case study in Jakarta. In: 2019 7th international conference on smart computing and communication. ICSCC 2019, pp 3–7. https://doi.org/10.1109/ICSCC.2019.8843632
16. Ekawibowo SA, Pamungkas MP, Hakimi R (2018) Analysis of 5G band candidates for initial deployment in Indonesia. In: Proceeding 2018 4th international conference on wireless and telematics. ICWT 2018, pp 3–8. https://doi.org/10.1109/ICWT.2018.8527780

17. Sheikh MU, Ghavimi F, Ruttik K, Jantti R (2019) Analysis of indoor solutions for provision of indoor coverage at 3.5 GHz and 28 GHz for 5G system. In: 2019 26th international conference on telecommunications. ICT 2019, pp 340–345. https://doi.org/10.1109/ICT.2019.8798826
18. MathWorks, NR cell performance evaluation with MIMO [Online]. Available: https://in.mathworks.com/help/5g/ug/nr-cell-performance-evaluation-with-mimo.html

A Study and Overview on Current Trends and Technology in Mobile Applications and Its Development

Harish Rathod and Sanjay Agal

Abstract Nowadays, everyone uses a mobile phones. Everything is embedded in the mobile phones. Day by day numbers of features are introduce in the mobile phone as well as in mobile applications too. Peoples are using mobile phone and mobile application to simplify their daily life task. It has been discovered that smartphone apps are becoming an extremely crucial aspect of people's lives. As we can see, mobile applications are used in almost every industry. Different technologies are used in different mobile applications. In this paper we are focusing on currently available of different types of mobile applications in the market. In this paper we discussing currently which kind mobile applications and technologies are widely are used in the market? We also talk about how much effort and money will be spent developing mobile apps. The objectives of this review paper are to understand the usage of smartphones and mobile data, understand current trends in mobile applications, understand mobile application development life cycle, understand mobile application development time and cost, and understand revenue generated from mobile application. According to the survey, India has the most smartphone usage, and mobile data rates are quite low, the Android and iOS platforms are used to create the majority of mobile applications. In mobile applications, technologies such as geo-based location, augmented reality, beacons, and the Internet of Things (IoT) are commonly used.

Keywords Mobile application · Mobile operating system · Application development

H. Rathod (✉)
Gujarat Technological University, Ahmedabad, Gujarat 382424, India
e-mail: hirathodphd2021@gmail.com

S. Agal
Dr. V. R. Godhania College of Engineering and Technology, Porbandar, Gujarat 360575, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_35

1 Introduction

Since the last ten years, the information technology industry has grown at a breakneck pace. Lots of software are available in the market. Nowadays almost every daily life tasks of human beings are carry out and monitor by software systems. In the early days software system was used only at specific places such as space research centers, government offices and private companies and organizations to make their task easy. After that many peoples were used software system to expand their business and make their office work easy but at that time software systems can be used by either standalone software or websites. One can use a desktop or laptop to access standalone or web application. Every normal people cannot effort to buy a personal computer or laptop. So it could be the limitations for expanding of business. From the very short of span a smartphones has introduced and it is chipper and easy to use as compare to personal computer and laptop. Nowadays every next person using a smartphone. People are using smartphone not only for phone calls but they are using smartphones for the purpose from learning to earning. People can download a vast variety of mobile application from the app store. Nowadays people are using mobile applications for their different needs such as education purpose, shopping, health and business. Different mobile application uses different technologies. This paper divided into following sections. Section 2.1—Smartphones and mobile data usage. Section 2.2 Current trends in Mobile Applications and Technology used in Mobile Applications. Section 2.12—Mobile Application Development Life Cycle. Section 2.13—Mobile Application Development Cost and Time. Section 2.14—Revenue generation from mobile application. Section 3—Summary.

2 Current Trends and Technologies in Mobile

It has been observed that the use of smartphones and their applications has increased dramatically over the previous two decades. Mobile apps are more user-friendly than desktop or laptop apps. We've compiled a list of current mobile app trends and technology.

2.1 Smartphones and Mobile Data Usage

The demand for smartphones is increasing from last ten years. Because any application from a smartphone is much easier to use than a desktop or laptop application. According to Ericsson Mobility Report Indians are on first rank to spend their times on smartphone. This is not limited to only talk and messages but Indians are also increasing to spend their times on streaming application and social networking sites [1]. In 2019, Indians consumed an average of more than 15 GB of data each month, a figure that is expected to rise in 2020 and 2021. In India, data usage per person climbed to more than 12 GB in 2020, and in 2021, it would increase to 18.8 GB per person. In 2019, the average amount of data used per person in Western Europe was more than 7 GB. It was increased to 11 GB per user per month in 2020. In North America, data usage per person was more than 8 GB in 2019, and is expected to rise to more than 11 GB in 2020. By 2021, this figure would have surpassed 14 GB per person. At the end of 2021 it was reached to 15.2 GB. In middle and East Europe usage of data in 2019 was 5.1 GB per person. In 2021 it was increased to 9.9 GB. In 2020, per person smartphone traffic was per month average 16.1 GB. In 2021, it was increased to 18.4 GB. One of the reason for increasing of smartphone traffic is work from home. India had 8 crore smartphones in 2021, and this number will rise to 1.2 billion by 2027. India will have 4G dominance by 2027. By the end of 2027, India will have 50 million 5G subscribers. India boasts the world's cheapest mobile phone rates. According to the Mobile Data Processing Report, India has the lowest mobile package rate per 1 GB in the world followed by Israel, Kirgizstan, Italy, and Ukraine. According to the survey, the majority of India's middle class is youthful and technologically sophisticated. The Indian market is strong enough to withstand new technologies. India's smartphone business is booming. In India, the smartphone Market competition is severe, and data is also incredibly inexpensive. In India, 1 GB of mobile data costs $0.09, in Israel, $0.11, in Kyrgyzstan, $0.21, in Italy, $0.43, and in Ukraine, $0.46. In St. Helena, mobile data rates are the most expensive. It costs 583 times as much as India. Fiber infrastructure is better in these countries. Customers will spend a total of 930 h of the year, i.e., 39 days on mobile in the coming days. According to the survey, 45 lakh crore hours were spent on mobile Internet devices in these countries. In 2015, people all throughout the world spent only 80 min on their mobile devices watching the Internet. This time period has now been extended to 130 min. Smartphone availability, speed, a good screen, and app innovation have all contributed to a large surge in mobile Internet usage.

2.2 Current Trends and Technologies

During this pandemic situation almost everyone from child to old person has change their life style most of the time during pandemic they spent on mobile phones. Almost every things for daily needs for life are available online. People are like to buy their house hold consumable items online in this pandemic situation. Mobile applications are available for download from a variety of app stores. To access different mobile applications, you can use a different mobile operating system. Smartphones based on Android and iOS are widely used. Table 1 shows the market share of operating systems. From Table 1 we can say that, Android and iOS capture 99.30% market worldwide [2]. As a result, the Android and iOS operating systems are the most popular among app developers. Traditional software development is still viable in the market but mobile application provide some extra advanced features that a typical desktop or web applications may fail to provide it.

Table 1 Market share of operating system—March 2022–March 2023

Operating system	Market share (%)
iOS	28.37
Android	70.93
KiaOS	0.12
Samsung	0.36
Windows	0.02

For example availability of Bluetooth, Wi-Fi, Sensors, Compass, GPS, etc. makes the mobile phone a much powerful deployment device than a PC [3]. From the app stores many numbers of mobile applications can be downloaded. People are expanding their business and earning money through mobile applications, because a huge number of mobile users are available than a desktop or PC. I have compiled a list of the most popular mobile applications as well as the technologies that are used in them.

2.3 Social Media Mobile Applications

Social media platforms such as Facebook and Twitter have overtaken traditional websites. People prefer to share their thought and experience through the Social Media instead of websites [4]. People cannot picture their lives without the use of the Social Media mobile app. People start their morning with check the news feed on WhatsApp, Snapchat, and Linked etc. Mobile application can be classified by different parameters such as Audience, Purpose, and Platform. Based on the number of monthly active users [5]. Table 2 depicts very famous mobile chat application based on active user. WhatsApp application has highest active users. There are many technologies used in Social Media mobile applications based on the functionality provided by it. A trending technologies for Social Media mobile applications are: Location-based services, Voice interfaces, Live video streaming, Interaction with other services, Networks for building relationships, Networks for media sharing [6]. Top ten countries that spend the most time on social media as shown in Table 3. With 156 min, India is ranked 15th on the list. In comparison, the global average is 142 min.

Table 2 User of chat application—October 2021

Social media mobile application	Active users (millions)
WeChat	1251
Facebook messenger	1300
Snapchat	538
Telegram	550
WhatsApp	2000

Table 3 Top ten countries that spent the most time on social media

S. No.	Name of Country	Time in minute
1	Philippines	230
2	Nigeria	222
3	Brazil	218
4	Kenya	214
5	South Africa	212
6	Colombia	210
7	Mexico	203
8	Argentina	201
9	Indonesia	194
10	UAE	191

2.4 On-Demand Mobile Application

During the COVID-19 pandemic outbreak, people now prefer getting most of the things delivered at their doorstep to stay at home as much as possible. People's access to services has been transformed thanks to on-demand platforms. According to a survey conducted by the Aspen Institute, more than 85 million US citizen uses minimum one on-demand application [7]. In the on-demand app development market, 45 million people in the United States have provided their skills. Yearly, 22.4 million users spend $57.6 billion on on-demand services [8]. Only in the final quarter of 2017, on-demand mobile app development spending totaled $10,293 billion. Revenue from on-demand app development will reach $335 billion by 2025.

Types of on-demand apps are person to person, enterprise to person, enterprise to enterprise [9]. In the market, there are a variety of on-demand mobile applications. Some of the well-known on-demand mobile applications are Uber, Zomato, Swiggy, etc. A trending technologies for on-demand mobile applications are: Location-based services, payment methods, order tracking, deliver status, social networking integration, push notifications, and Chabot. Challenges are legal challenges and security challenges. Opportunities in on-demand apps are real time technology, background history and capability, interactive maps, pricing, and distance-wise priority [10].

2.5 Mobile Health Application

In this pandemic situation it is not necessary to visit the doctor personally. Mobile applications provide the feature of video conferencing so that customers can directly consult the doctors. In 2020, the global market for mobile Health application was valued at $24.93 billion. The global market was expected to rise by 65.7% in 2020. In 2021, the market is predicted to be worth $38.89 billion, rising to $314.60 billion by 2028 [11]. Worldwide China had 65% of mobile health app users in 2020, followed by India with 63%, Australia with 45%, and Netherland with 33%, and France with 33% [12]. Worldwide some the well-known mobile health applications are introduce here. (1) The media safe pill minder app, this mobile application keep track on your medicine so you don't forget to take them. It has features like doctor visit management and appointment reminders, and it's available on Android and iOS. This also offers health-monitoring tools. (2) The Medical ID is free Android mobile application that allows the patients to create their multiple medical profiles. (3) The MyFitnessPal is one of the most popular health-tracking applications available for Android and iOS. This mobile application is all-in-one workout, mindfulness, food tracking, sleep, and fertility tracker. You can use the app to establish weight and exercise goals, make meal plans, and connect it with most fitness trackers and watches. (4) The Diabetes: M this application is available in Android and iOS platform, that is aimed to assist diabetics take some of the guessing out of controlling their glucose levels. (5) The Apple Health is the iOS health "brain" and tracker app that is included with all Apple devices. Sleep, eating, activity, symptoms, heart rate, and respiration are all monitored [13].

2.6 IoT-Based Mobile Applications

The Internet of Things (IoT) is a collection of components or people who are known as "things" that are integrated with software, electronics, a network, and sensors to collect and exchange data. The purpose of the Internet of Things is to extend Internet connectivity from traditional devices such as computers, smartphones, and tablets to relatively simple items such as household appliances [14]. There are numerous domains like health, transportation, education, etc., in which IoT can be applied to tackle our day-to-day challenges. The Internet of Things has a positive impact on our civilization. Mobile phones are used by almost everyone on the planet. As a result, the optimum platform for IoT is a mobile application. As a result, anyone can utilize it without difficulty. Mobile applications are affecting IoT in a variety of ways, making the technology more diversified and versatile. The following are the most notable examples: Smart Homes, Wearables, Healthcare and Medicine, Retail, Agricultures, Smart cities, etc. In recent years, the Internet of Things has experienced remarkable expansion, which does not appear to be slowing down anytime soon. IoT technology, which includes software, services, connections, and devices, has a global market. By 2025, more than 64 billion IoT devices are estimated to be online and actively working [15]. The post-COVID-19 worldwide IoT market is predicted to increase at a Compound Annual Growth Rate (CAGR) of 16.7% from 2021 to 2026, from USD 300.3 billion in 2021 to USD 650.5 billion in 2026 [16]. IoT can be used in a Learning Management System, according to Khaleel Mershad and Pilar Wakim. Experimentation, Remote Lectures, Data Sharing, Students' Assessment, Classroom Applications, and Classroom Monitoring were among the LMS applications showcased [17]. The advantages of IoT-based mobile applications are numerous. A couple of these advantages are listed here like cost effective, technical optimization, improved data collection, reduce waste, improved customer engagement. The disadvantages of IoT-based mobile applications are security, privacy, flexibility, complexity.

2.7 Augmented Reality in Mobile Application

During the year 2020, augmented reality technology will have experienced great growth. Commercial use of the technology has increased because of widely used by market leaders like Microsoft, Apple, Google, Facebook, and Amazon. Augmented reality has a market value of $15.3 billion. By the end of 2020, AR active devices were estimated to rise to 598 million units and are projected to grow to 1.73 billion by 2024 [18]. AR features in mobile apps can be used for a variety of purposes. Many applications which is based on augmented reality can allow the customer to apply different hair style and makeup on their faces virtually so customer can check how they look. One of the well-known beauty product company L'Oreal Paris provide this kind of application to their customers. Instagram and Snapchat uses AR for face filters [19]. AR can be used for Navigation. There are different kind of Navigation are available such as Car Navigation, Pedestrian Navigation, Indoor Navigation [20, 21]. A mobile application based on augmented reality is now being used in medical science for learning purposes. NitLabEduca, a smartphone application based on augmented reality, was created by Jacks Fernandes and his team. NitLabEduca was discovered to be a resource capable of spatial abstraction as well as functional comprehension of spinal cord pathways. Furthermore, NitLabEduca has shown strong usability in the teaching–learning process and can be used to supplement printed materials, enhancing student and health professional training [22].

2.8 Virtual Event Streaming Mobile Application

With the world confronting the COVID-19 epidemic, staying at home has become the norm. Many offices have announced work from home options, students are studying in the comfort of their own homes, artists are not permitted to hold concerts, conferences, or training sessions, and even fitness classes are now available online. According to Research and Markets "From 2019 to 2025, the global video live broadcasting solutions market is expected to develop at a CAGR of 17.80%" [23]. As more individuals look for outlets to watch live video feeds, this sector is poised to soar to new heights. In 2020, according to Grand View Research, The worldwide market of video streaming was $50 billion. It will be expected in increase up to $61 billion. Artificial intelligence and blockchain technology which enhance the video quality are expected to hike the market of video streaming. Video production includes cinematography, editing, scriptwriting, voice-overs, and a variety of other elements and uploading all rely on AI [24]. The live streaming platforms available in the markets are Tango, Streamlabs, Youtube live, Facebook live, Twitch, Instagram live, Livestream, Vimeo, Hakuna, Broadcast Me, etc. [25]. Issue in live streaming applications are camera issues, meeting management issues, microphone issues, and work from home issues [26].

2.9 Beacon Technologies and Mobile Applications

Beacons have been among the best effective ways to connect people in a variety of businesses, especially retail and marketing. The crux of beacon technology is the beacon device. Beacons are a type of small computer. Beacons used IoT to search human in its network and provide information. Beacons act as a guide, such as compass. Beacons guide the smartphone users. Apple's iBeacon technology was the first to be introduced. It was soon followed by Google's Eddystone. If a person passing through beacon network then the nearest beacon will send a message to the persons' mobile and guide them. To make a beacon-based control system, you'll need three things: At least one or more bacon device, mobile application, user's permission. A latest version of iOS and Android is recommended to support beacons. The Apple beacon standard is known as iBeacon, while Google's is known as Eddystone. Types of beacons are Standard beacon, Small/Portable beacon, USB beacon, Video beacon, AI beacon, Sticker beacon, Parent beacon, Dedicated beacon [27]. Crowd Compass Beacons, Event2Mobile app, Locate Beacon app, Medical Appointment Scheduling Management, Launch here app, BeHere app, Proximitask app, Mingleton app, At the Ballpark app, Event base app, NearBee app are some of the available smartphone applications that utilize beacon technology [28, 29].

2.10 Mobile Game Application

According to Oxford Dictionaries, game is defined as: "*An activity that one engages in for amusement or for fun*". The basic definition of the game is to play or to compete between two or more players with the rules set to achieve the objectives of the game. The game consists of various types or genres for example sports, simulation, adventure, escape, and so on. One of the gaming tool is mobile application. **Lots of free and paid mobile games can be downloaded from Google play stores and apple store**. In this pandemic situation many people are playing mobile games to spend their times. However, the game not only focuses on people to play but provides an activity in the learning process as well [30]. During the forthcoming years, 2021–2026, the handheld gaming market is estimated to grow at a CAGR of 12.3%. The increasing acceptance of current technologies for generating games, as well as the growing smartphone usage, are important contributors in the growth of the mobile gaming business. The gaming industry is growing around the world, and smartphones are playing a big part in it. The gaming business has gained scalability as a result of the introduction of mobile games. Platforms such as Facebook and Instagram have begun to produce new mobile games in order to assure high product distinctiveness and to benefit from engaging games to improve their advertising tactics [31]. Most popular mobile games are: "*Earn to dia 2, Beach Buggy Racing, Hay Day, Cover fire, Clash of Clans*" [32].

2.11 Mobile Application for Trading Stock Market

Mobile applications has been started dominating our everyday lives especially in the lockdown. One area where use of mobile applications are growing is the stock market. Many people have been using the mobile trading apps of stock broking brands since last couple of years. But the lockdown has witnessed a surge in stock trading and investing activity on mobile phones to the extent of 45% in May 2020 as compared to 24% in last year for the same. With 1.2 million new accounts opening at CDSL in March and April, stock broking brands apps like Zerodha and Upstox have seen around 20% rise in first time investors mainly under the age of 30 years. Many people jumped on this rise feeling the justified valuation of stock prices post March lows. The pandemic situation provided more time to investors and users of mobile trading apps to focus on stock prices in a more focused manner. Mobile trading apps are extremely popular among modern day technology investors [33]. Benefits of online trading are: Convenient, Information, Extra Power, Control at Individuals, Global Trade etc. [34]. Trading apps have grown increasingly popular among teenagers. People are increasingly opting for mobile trading over traditional trading. Furthermore, cryptocurrency have become the latest craze among teenagers. India is the world's second-largest cryptocurrency adopter. "*Zerodha Kite, Upstox PRO, Angel Broking, Motilal Oswal, 5Paisa, and India Infoline*" are the safest online stock trading smartphone applications in India [35].

2.12 Mobile Application Development Life Cycle

The mobile application development life cycle is largely no different than the development life cycle of web or desktop application [36]. There are five major portion of the process. (1) Requirement gathering and analysis of the requirement. (2) Design the user interface of the mobile application. (3) Development of the mobile application. (4) Testing the application and identify bugs. (5) Launch the mobile application. While developing a mobile application is fundamentally similar to traditional web/desktop development in terms of process and architecture, there are a few things to keep in mind. Multitasking, form factor, device and operating system fragmentation, and limited resources are all things to keep in mind.

2.13 Mobile Application Development and Cost

The cost to develop an app will vary according to the platform and its features. Client need to ensure whether they are planning to build the app on android, iOS, or both and about the features they will include in the app. Other major factors that influence the cost of an app are the design, premium APIs, as well as technologies. A simple on-demand mobile application will take around 150–200 h, and its development cost will be $6000–$8000. The cost of developing an app can range anything from $2000 to $300,000. A simple mobile application cost less while complex mobile application cost more depending on the functionality. The average time required to develop an app can go from 2 to 3 weeks to 9+ months. It consider multiple variables like the number of features and their complexity, uniqueness of the design, development complications, use of third-party services, and so on.

2.14 Revenue Generation from Mobile Application

Mobile applications can be created for a variety of reasons, including benefit to society or income generation. One of the best ways to get consistent revenue is to use a mobile application. Many organizations, companies, small business groups, shopkeepers, and others converted their businesses to internet mode during the pandemic scenario. They did a lot of their business through smartphone apps. They earned money by providing their services through smartphone applications. There are a range of revenue models which can be used with mobile applications. Approaches for generating profits are user subscription, affiliate deals, in-app advertising, service fees, in-app purchases, crowdfunding, brand sponsorship, E-commerce, donations, merchandise [3, 37]. Worldwide income from mobile apps is estimated to cross over 318 billion dollars by 2020. In comparison to the previous year, this was an increase of more than 60 billion dollars. The most profitable area that year was mobile gaming, which generated over 200 billion dollars in sales, followed by social networking mobile apps, which generated over 31 billion dollars. Over the following five years, sales across most segments will rise, bringing total revenue to roughly 613 billion dollars by 2025.

3 Summary

The goal of this review article is to find out what's new in the field of mobile applications. Recently, various kind of mobile applications and technology are most extensively employed around the world. We gave data on smartphone and mobile data usage in various nations around the world. We showcased nine different sorts of mobile applications and technologies that are frequently used worldwide. The life cycle of mobile application development was described, followed by revenue generating from mobile applications. In this pandemic situation everyone in the world using a mobile phones for their daily life needs in every aspects. India has the greatest rate of smartphone and mobile internet usage. In the market, there are a variety of mobile platforms on which mobile applications can run. From the study we can say that Android and iOS platforms are widely used in the world. From the app stores, you may get a wide range of mobile applications for various purposes. Among these mobile applications current trending mobile applications categories are social media applications, on-demand applications, event streaming applications, gaming applications. Mobile health applications, IoT-based mobile applications, augmented reality, beacon technology, location-based technologies, and artificial intelligence are all popular technologies. The development process of mobile application is almost same as desktop or web applications but some considerations are multitasking, form factor, device and operating system fragmentation and limited resources. Mobile game applications have the largest income potential. The duration of time it takes to design a mobile application is determined by its complexity. The development time is usually between 150 and 300 h. The cost of development can run anywhere from \$2000 to \$30,000. The cost of development is determined by the amount of features and their complexity, as well as the design's distinctiveness.

Users increasingly need a large number of mobile applications with several features. As the functionality of a mobile application grows, so does the complexity of the app, and so do the failure rates of mobile apps. Many mobile applications provide good functionality and services, but they are not widely used, and mobile applications are gradually being phased out.

References

1. Ericsson Team (2021) Mobile internet usage. https://www.ericsson.com/en/reports-and-papers/mobility-report. Accessed 05 Jan 2022
2. Simpson R (2021) Mobile operating system market share worldwide. https://gs.statcounter.com/os-market-share/mobile/worldwide. Accesses 25 Sept 2021
3. Thomas CG, Devi J, A study and overview of the mobile app development industry. Int J Appl Eng Manag Lett (IJAEML) 5(1):115–130. https://doi.org/10.5281/zenodo.4966320
4. Ramos RF, Rita P, Moro S, From institutional websites to social media and mobile applications: a usability perspective. Eur Res Manag Bus Econ 25:138–143
5. Statista Research Department (2021) Most popular global mobile messenger apps based on number active users. https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/. Accessed 15 Oct 2021
6. Beklemysheva A (2019) Technologies in social media. https://steelkiwi.com/blog/social-network-application-development-types-challenges-technologies-costs/. Accessed 17 Oct 2021
7. Patel U (2021) On demand app. https://www.tristatetechnology.com/blog/best-on-demand-app-development-ideas-in-2021/. Accessed 09 Oct 2021
8. Patel B (2021) Users of on demand app. https://www.spaceotechnologies.com/on-demand-app-ideas/. Accessed 09 Oct 2021
9. Woxapp team (2021) Types of on demand app. https://woxapp.com/industries/on-demand-app-development/. Accessed 10 Oct 2021

10. Arafat S, Challenges for on-demand technology and the unseen opportunities. Int J Bus Hum Technol 8(2). https://doi.org/10.30845/ijbht.v8n2p1
11. Fortune Business Insight Team (2020) The global market for mobile health application. https://www.fortunebusinessinsights.com/mhealth-apps-market-102020. Accessed 10 Feb 2022
12. Statista Team (2022) The number of mobile health application users in selected country. https://www.statista.com/forecasts/1181436/share-of-health-app-users-by-country. Accessed 15 Feb 2022
13. Knerl L (2021) The best mobile health care application for patient. https://www.hp.com/us-en/shop/tech-takes/best-healthcare-apps-for-patients. Accessed 13 Feb 2022
14. Smith J (2022). Internet of things. https://www.guru99.com/iot-tutorial.html. Accessed 22 Feb 2022
15. Sinha D (2019) Domain of IoT applications. https://www.iotforall.com/mobile-iot. Accessed 18 Feb 2022
16. IoT Market Team (2022) The Market of IoT. https://www.marketsandmarkets.com/Market-Reports/internet-of-things-market-573.html. Accessed 21 Feb 2022
17. Mershad K, Wakim P, A learning management system enhanced with internet of things applications. J Educ Learn 7(3). ISSN 1927-5250. E-ISSN 1927-5269. https://doi.org/10.5539/jel.v7n3p23
18. Makarov A (2021) Augmented reality technology. https://mobidev.biz/blog/augmented-reality-future-trends-2018-2020. Accessed 25 Oct 2021
19. Buildfire (2021) Augmented reality. https://buildfire.com/mobile-app-development-trends/. Accessed 01 Nov 2021
20. Sharma MK, Chachaundiya S, Vishal, Augmented reality navigation. Int J Eng Res Technol (IJERT) 9(6)
21. Tamhane M, Sarjekar P, Gupta C, Karche VM, Campus navigation using augmented reality. Int J Sci Eng Res 9(4)
22. Fernandes J, Teles A, Teixeira S, An augmented reality-based mobile application facilitates the learning about the spinal cord. Educ Sci
23. Research and Market Team (2021) Virtual event streaming. https://www.researchandmarkets.com/reports/4763596/video-live-stream-solutions-market-global. Accessed 05 Dec 2021
24. Grand View Research Team (2021) Video production. https://www.grandviewresearch.com/industry-analysis/video-streaming-market. Accessed 08 Dec 2021
25. Pate B (2021) List of virtual events application. https://www.spaceotechnologies.com/mobile-app-based-live-streaming-platforms/. Accessed 12 Dec 2021
26. Karl KA, Peluchette JV, Aghakhani N, Virtual work meetings during the COVID-19 pandemic: the good, bad, and ugly. Small Group Res 1–13. https://doi.org/10.1177/10464964211015286
27. Intellectsoft Team (2020) How beacon technology works. https://www.intellectsoft.net/blog/what-are-beacons-and-how-do-they-work/. Accessed 20 Dec 2021
28. Malik N (2021) List of beacon mobile application. https://blog.beaconstac.com/2015/02/10-best-ibeacon-apps-that-are-leading-the-pack/. Accessed 24 Dec 2021
29. Zea HEC, Solano SR, Rondon MM, Puente KG et al, Prototype of beacon-mobile application for medical appointment scheduling management. In: Conference paper. https://doi.org/10.18687/LACCEI2019.1.1.145
30. Yahya FA, Ahmad I, Roszali FF, Sarudin N, A review on mobile game learning applications trends. Int J Eng Trends Technol (IJETT)
31. Mordor Intelligence (2021) Mobile gaming market. https://www.mordorintelligence.com/industry-reports/mobile-games-market. Accessed 02 Jan 2022
32. Diwas (2022) List of mobile games. https://www.seasiainfotech.com/blog/most-popular-mobile-games/. Accessed 10 Jan 2022
33. Malhotra S, Study of features of mobile trading apps: a silver lining of pandemic, J Glob Inf Bus Strat 12(1)
34. Sharma S, Anute N, Ingale D, A study of online trading system in India. J Manag Res Anal 8(3):139–142. Content available at: https://www.ipinnovative.com/open-access-journals

35. Startup Talky Team (2021) Safest online trading mobile application in India. https://startuptalky.com/safest-online-trading-apps/. Accessed 15 Jan 2022
36. Microsoft Team (2021) Mobile application development. https://docs.microsoft.com/en-us/xamarin/cross-platform/get-started/introduction-to-mobile-sdlc. Accessed 18 Jan 2022
37. Kontsevoi B (2020) Approaches for generating revenue from mobile applications. https://www.forbes.com/sites/forbestechcouncil/2020/04/28/mobile-app-monetization-part-1-revenue-generation-models/?sh=6aac9b1602df. Accessed 03 Feb 2022

Heart Rate Monitoring Using External Camera

Chandan Singh Rawat, Tanya Dubey, Sagar Pujari, Yash Bhise, and Hridesh Kamal

Abstract Heart rate is a crucial indicator of people's physiological state. Patients, who are in need of regular check-ups, find the existing modes of electrode-based and probe-based monitoring inconvenient and uncomfortable. To eliminate this inconvenience and provide a comfortable monitoring environment for the patients, the non-contact heart rate monitoring is introduced. Latest research has proposed numerous methodologies for the non-contact detection of heart rate (HR) based on subtle color changes of the face due to cardiovascular activities, which are invisible to human eyes but can be captured by digital cameras. We propose a framework which utilizes ICA algorithms and power spectral analysis. It works well on stationary subjects under well controlled conditions, but the performance slightly degrades if the videos are recorded under more challenging conditions, specifically when subjects' motions and illumination variations are involved.

Keywords Heart rate · Non-contact · Contactless · PPG · BPM · Camera

1 Introduction

Heart rate (HR) is a crucial sign of someone's physiological condition. The skin contact requirement of conventional HR monitoring techniques like electrocardiograms, sphygmomanometry, and pulse oximetry makes them uncomfortable and cumbersome. On the other hand, there are commercial cameras everywhere these

C. S. Rawat (✉) · T. Dubey · S. Pujari · Y. Bhise · H. Kamal
VES Institute of Technology, Chembur, Mumbai, India
e-mail: chandansingh.rawat@ves.ac.in

S. Pujari
e-mail: 2019sagar.pujari@ves.ac.in

Y. Bhise
e-mail: 2019yash.bhise@ves.ac.in

H. Kamal
e-mail: 2019hridesh.kamal@ves.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_36

days, including webcams, security cameras, and cell phone cameras. There are a lot of possible uses for the method of remote HR monitoring utilizing common cameras. However, it is still a difficult task because the change brought on by the cardiac pulse is very small compared to the many other factors that can also cause fluctuation of the local skin's gray value. It has been reported that skin color change brought on by cardiac pulse can be captured by common cameras for HR measurement. Lighting changes and subject movements are two of these elements that are crucial. In this article, we provide a novel HR measurement framework that can lessen noise caused by illumination variations and subjects' motions.

Although it is a common issue in biomedical research, video (of visible light) processing is rarely addressed with HR measurement research. This study shows how computer vision techniques can assist in addressing the issue of remote HR measurement.

Recent studies have demonstrated that, in well controlled lab settings, HR can be determined from facial videos. However, under real-world circumstances, the task is more challenging since a number of variables could taint the pulse signal obtained from the face area [1].

Human heart pumps blood in the entire body through blood vessels and capillaries. This pumping of blood is carried out at regular intervals depending on the rate at which the heart beats. When this blood is pumped it also reaches the facial area leading to a very slight change in the person's overall color complexion. The color changes during the heart pump and becomes normal after each pump. This information could be used in determining the person's heart beat as it would correspond to the color change.

Section 1 explains the significance of heart rate monitoring and gives a general introduction to the methods of monitoring heart rate. Section 2 describes the review of literature where a brief of existing methods for heart monitoring has been discussed. Section 3 consists of a contactless method for measuring heart rate along with the system's block diagram and detailed working. It also covers the technology stack used for its implementation. Section 4 consists of the results and analysis section where the real-world application statistics have been provided.

2 Background and Literature Review

This section contains a summary of various methods and techniques, used for non-contact heart rate monitoring, that we have studied to get a better understanding on the topic. This study has been performed using 8 different research papers which have been mentioned in the references section.

Both commercial and academic interests are drawn to the topic of remote, non-intrusive HR measurement. Photoplethysmography (PPG) has been used in numerous earlier attempts at remote heart rate monitoring. The blood volume pulse (BVP), measured at peripheral body tissues (such as the palm or fingertip), is typically employed as an indicator of cardiac cycle measurement because the blood volume of microvascular throughout the body fluctuates along with cardiac pulse. The PPG method's basic idea is to illuminate the skin with an LED before measuring how much light is reflected or transmitted to a photodiode. PPG can determine the local blood volume pulse because the quantity of light absorption depends on blood volume. Although it is possible to monitor HR using PPG-based settings without any contact, this technique still needs specialized lighting and sensors [1].

Recently, numerous articles put forth color-based approaches for measuring HR remotely with commonly available cameras. Poh et al. [2] investigated the viability of measuring HR using web-cam recorded face videos. They used the Viola-Jones face detector to identify the region of interest (ROI) and calculated the mean pixel values of the ROI for each frame from three different color channels. The PPG signal was subsequently shifted into the frequency domain to determine the frequency with the maximum power as the HR frequency after Independent Component Analysis (ICA) was used to isolate the PPG signal from the three color traces. Poh's findings demonstrated that ICA separated sources can obtain higher accuracy for measuring HR in comparison to the raw green trace.

The built-in camera of a smartphone was used by Kwon et al. [3] to capture facial movies, and they retrieved HR using both the raw green trace and the ICA separated sources. They discovered that, in contrast to Poh's findings, ICA marginally decreased performance. Later, Poh et al.'s [4] approach was enhanced by including a number of temporal filters both before and after ICA. When comparing their self-collected data to the new method, HR was measured with extremely high accuracy.

Balakrishnan et al. [5] proposed a motion-based approach. They monitored minute head oscillations brought on by the heartbeat and used principal component analysis (PCA) to separate the pulse signal from the paths of several feature points they were tracking. On the videos they had amassed on their own, the technique displayed encouraging results. Since the approach depends on motion tracking, the individuals are restricted in that they cannot move voluntarily. They suggested that one useful future step would be to measure HR on moving subjects.

In this section, a brief summary of the papers studied was presented. It talks about two main approaches for remote heart rate monitoring—motion-based methods and color-based methods. From the literature review, it was concluded that color-based methods produce more accurate results. Keeping in mind the limitations and drawbacks of the previous methods, the problem statement for the project was formulated.

3 Methodology

The proposed system consists of a methodology for determining the heart rate of a person without the use of any physical contact with the person.

As depicted in Fig. 1, first the subject's video is recorded (minimum 30 s footage). The video captures the subject's ROI which would be further useful in determining the heart rate by performing various functions on the same. The ROI consists of the

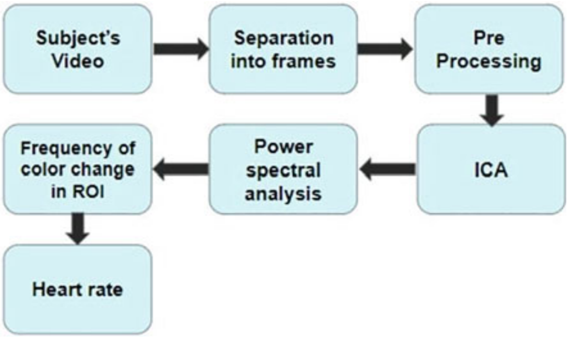


Fig. 1 Block diagram of the proposed system

subject's cheek and temporal area. Then captured video is then broken down into individual frames (each frame consists of the ROI precisely captured and marked for applying procedures).

First, intensity rectification is performed. This is done to eliminate the errors in the RGB color spectrum in ROI due to ambient lighting. Then Independent Component Analysis (ICA) is performed to extract only the information which is useful and the rest of the information which is redundant is discarded. In our case the red and green spectrums had maximum color variations whereas blue was discarded as it did not show a notable change in its value with respect to time. All this is done on a spatial level.

The above information is then converted into frequency domain using fast Fourier transform. A fast Fourier transform (FFT) is used in the power spectrum analysis method to determine the frequency spectrum of a particular signal from its variation. It is helpful for choosing the right sampling rates as well as figuring out how much noise is present with the signal. Once the power spectral analysis is performed, a relation between color change in RIO and time is established. This gives information regarding the frequency of color change with respect to time of the ROI. The frequency of color change within a 1 min window frame would thus give an estimated heart rate of the subject.

4 Algorithm and Implementation

Firstly, the data is acquired using the OpenCV library in Python. The input rate of the video is set at 15 frames per second to ensure adequate number of frames for smoother calculation. The raw data is then put through a series of basic processing steps to make it useful for HR analysis' descriptive phases. In processing, the color is magnified to amplify the hues present in the video feed (Fig. 2).

This data of the video is used to make kernels and initialize the filters to match the input video data. Cut-off frequencies for the BPF are set in this step.

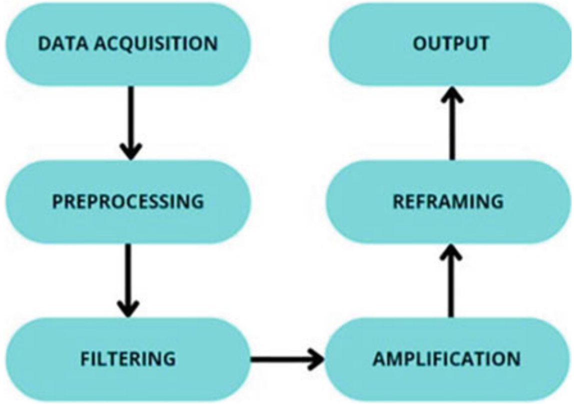


Fig. 2 General process of HR detection

The Gaussian pyramid [6] is also created to suit the input video. Once the filters and kernels are set up, the variables for the calculation of Fast Fourier Transform (FFT), Inverse Fast Fourier Transform (IFFT), and main BPM are initialized. In this step, the video data is first converted to frequency domain for faster calculation. This code uses FFT and IFFT to transform an image between the spatial and frequency domain. The FFT is used to calculate a Discrete Fourier Transform (DFT) using a smaller number of multiplications. A time-domain sequence is transformed into an equivalent frequency-domain sequence via the DFT. A frequency-domain sequence is transformed into an analogous time-domain sequence by the inverse DFT, which executes the opposite operation. The FFT approach keeps all of the original data. A bandpass filter is then applied to the altered image.

Once a pulse is detected, this periodic pulse is used to calculate an estimated Heart Rate of the subject in the frame. Then an IFFT of this pulse is taken and amplified to generate the main output. An output on the same window is displayed to the user along with the calculated beats per minute (BPM).

5 Results and Discussion

HR tests were done in parallel with another heart rate monitor device called the Oximeter. This ensured the accuracy of the measurement. Errors of 0–12% were detected, but it is hard to calculate the real error rate, as the monitor itself exhibits intrinsic errors.

The subjects were asked to sit in front of the laptop, then adjust the positioning of the laptop in such a way that the face of the subject was completely visible in the fixed ROI detection region of the GUI window.

Once the calculation process is initiated, the subject has to move as little as possible. This ensures that the code does not have to address any movements that could potentially cause an error in the algorithm. The current code cannot handle abrupt movements and can cause errors.

The lighting conditions around the subject are maintained bright. But the lighting is not exactly focused on the subject to ensure uniform illumination levels throughout the process there is no direct throw of light on the subject. A direct throw of light would cause an extremely bright spot of light on the subject's ROI hence causing significant errors in the subject's HR.

A mean value of the HR is calculated from the multiple frames of the subject recorded. This mean value was displayed in the GUI and the terminal window of the interpreter used.

The formula used for the calculation on Error (%):

$$\text{Error}\ (\%) = \left(\frac{\text{Calculated BPM} - \text{Actual BPM}}{\text{Actual BPM}}\right) * 100 \tag{1}$$

It can be observed that the HR readings taken using the oximeter and the HR readings taken using the algorithm are similar giving an error percentage of upto 12%. The accuracy of the algorithm is highly dependent on the stability of the subject and their background surroundings being captured in the frame.

The HR reading varies on the condition of the subject. A higher reading is obtained if the reading is taken right after the subject has been indulged in some physical activity or if the subject is stressed out. A more optimum HR reading is observed if the subject is calm and composed.

A study was conducted with the participation of a hundred subjects and the observations from both the proposed algorithm and an oximeter were documented to measure and evaluate the performance of this solution. The results were tabulated and are presented in Fig. 3. The red line represents the HR measurements using an oximeter whereas the blue line indicates the corresponding measurement using the algorithm. Both methods of measurement were performed simultaneously to ensure accurate comparison.

As visible in the graph, the algorithm closely follows the measurements obtained from the oximeter. Further, Fig. 4 provides a closer look at the difference between the algorithm's estimates and the actual results for different subjects using a bar graph. Another method to look at our findings is using a Bland–Altman plot. In heart rate monitoring, the Bland–Altman plot is commonly used to evaluate the agreement between two different heart rate measurement methods. For the purpose of our study, we evaluate the agreement between the proposed algorithm and readings measured using an oximeter.

In order to construct a Bland–Altman plot, the difference between the two measurement methods is calculated for each pair of measurements. The average of the two methods is also calculated for each pair of measurements. The difference and average values are then plotted on the *y* and *x*-axis, respectively. A horizontal line is drawn at the mean difference between the two methods, and two additional lines are drawn at the limits of agreement, which are typically set at ± 1.96 times the standard deviation of the differences.

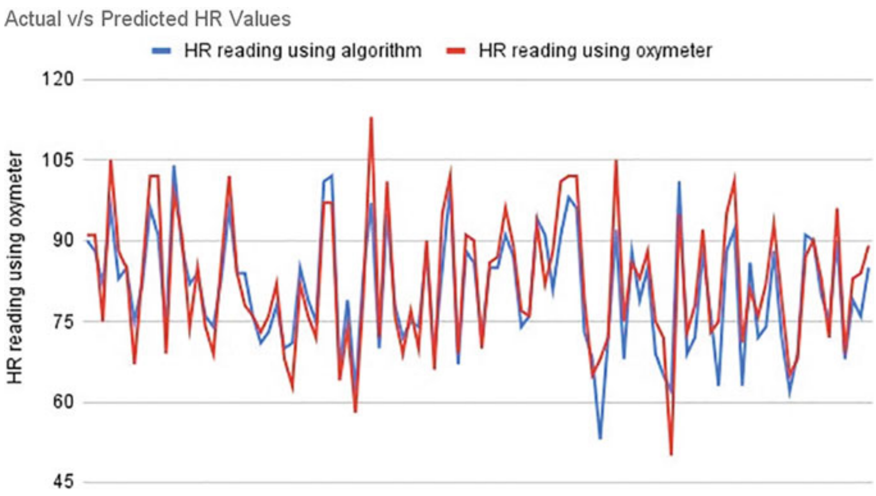


Fig. 3 Graphical representation of the results

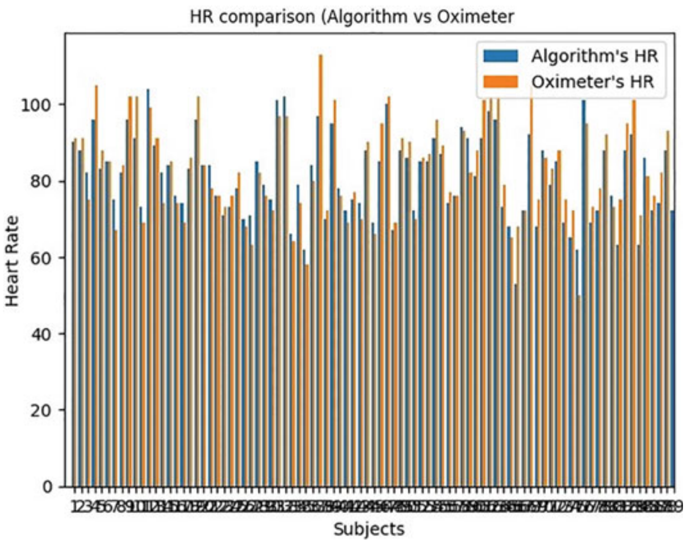


Fig. 4 Graphical comparison between algorithm and oximeter's readings

Interpretation of the Bland–Altman plot involves examining the distribution of the differences between the two methods and comparing it to the limits of agreement. If the differences are normally distributed and fall within the limits of agreement, then the two methods can be considered to be in good agreement. For this project, referring to Fig. 5, we find that only a miniscule fraction of measurements fall outside the limits of agreement, calculated at − 9.07 for lower limit and 12.27 for upper limit, respectively. Further, performing a Shapiro–Wilk test provides us with a *p*-value of 0.68 which does support a normal distribution.

It was also observed during the study that the subjects who perform physical activities and exercise on a daily basis had a lower HR reading as compared to the ones who don't. The subjects with the medical condition of hypertension also showed higher readings of BPM. However, it was established that only people with hypertension don't show higher BPM but some people with normal or low blood pressure also showed higher HR readings up to 105 bpm pertaining to other factors such as stress or fever.

In our research study, we found a strong positive correlation between the two methods of HR measurement, with a correlational coefficient of 0.89. This indicates a high degree of linear association between the two the proposed method and the Oximeter readings.

In addition, we computed two measures of error in our model. The root mean square error was found to be 5.56, indicating that our model's predictions deviated from the actual values by an average of 5.56 bpm. The mean absolute deviation was found to be 4.38, indicating that the average absolute difference between the calculated values and actual values was 4.38 bpm. These measures help show the reliability and precision of our model, making it a viable method for portable, non-contact HR monitoring.

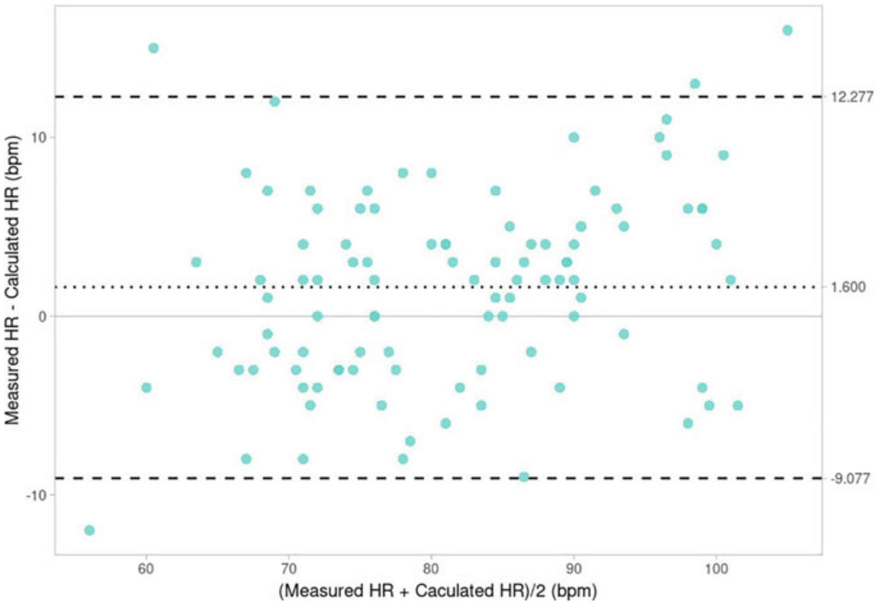


Fig. 5 Bland–Altman plot of the observations from the study. The dotted line indicates a mean bias of 1.6 bpm and the dashed lines indicate the limits of agreement at − 9.077 and 12.277 for our dataset

HR reading is an important factor when monitoring a person's health. A normal HR for healthy adults ranges from 60 to 100 bpm. A lower HR reading signifies a healthy heart, that is, good cardiovascular fitness.

The project of heart rate monitoring using an external camera has some limitations that should be taken into consideration. First and foremost, the accuracy of the heart rate measurement using this method may be affected by various factors, such as lighting conditions, camera position, and motion artifacts. Moreover, the method may not be suitable for individuals who wear makeup or accessories that can reflect light. Additionally, the camera-based heart rate monitoring may not be practical for continuous monitoring or for use during physical activities that involve a lot of movement. Finally, the method may not be as reliable as traditional heart rate monitoring methods, such as using a chest strap or a finger clip, particularly in clinical settings where accuracy is crucial. Therefore, while the camera-based heart rate monitoring has the potential to be a convenient and non-invasive method, it should be used with caution and its limitations should be carefully considered.

6 Conclusion

Heart rate monitoring is essential, as unusual changes related to the cardiovascular system help to obtain a diagnosis. Among all the pulse monitoring methods, the non-contact ones are considered to become the most useful in daily life and IoT applications. The methods proposed above may help increase the accuracy of detection of HR and improve the quality of detection in the case of dynamic subjects.

In this project, we saw the reliability of rPPG when used in conjunction with color magnification and Gaussian Pyramids for upscaling the input video stream over a 30 s observation window. The proposed solution was able to show a correlation coefficients of 0.90 with a pulse oximeter while displaying bias and precision close to commercially available rPPG equipment.

This paper represents a good proof of concept for this method, but there are areas which remain unexplored.

References

1. Wang C, Pun T, Chanel G (2018) A comparative survey of methods for remote heart rate detection from frontal face videos. Front Bioeng Biotechnol 6:33
2. Poh M-Z, McDuff DJ, Picard RW (2010) Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Opt Express 18(10):10762–10774
3. Kwon S, Kim H, Park KS (2012) Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. In: 2012 annual international conference of the IEEE engineering in medicine and biology society. IEEE
4. Poh M-Z, McDuff DJ, Picard RW (2010) Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans Biomed Eng 58(1):7–11

5. Balakrishnan G, Durand F, Guttag J (2013) Detecting pulse from head motions in video. In: Proceedings of the IEEE conference on computer vision and pattern recognition
6. Olkkonen H, Pesola P (1996) Gaussian pyramid wavelet transform for multiresolution analysis of images. Graph Models Image Process 58(4):394–398
7. Lomaliza J-P, Park H (2019) Improved heart-rate measurement from mobile face videos. Electronics 8(6):663
8. Ayesha AH, Qiao D, Zulkernine F (2021) Heart rate monitoring using PPG with smartphone camera. In: 2021 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE

Classification of Gait Abnormalities Using Transfer Learning with EMG Scalogram Features

Pranshu C. B. S. Negi, S. S. Pandey, Shiru Sharma, and Neeraj Sharma

Abstract Gait abnormalities can significantly impact the mobility and quality of life of individuals, thus making its early diagnosis crucial for proper treatment planning and rehabilitation. In this study scalograms generated from EMG signals of two important gait abnormalities, rheumatoid arthritis, and prolapsed intervertebral disc are classified using transfer learning. Scalogram has an advantage when dealing with high-noise data with abrupt transitions making it an excellent choice for classifying movement patterns. When classified using only CNN an accuracy of 91.1%, precision of 92.8%, recall of 92.9%, AUC of 0.98, and a PRC of 0.97 were obtained. For transfer learning, VGG16, VGG19, ResNet50, Inceptionv3, InceptionResNet, MobileNet and MobileNetv2, and DenseNet Large were incorporated along with previous CNN. DenseNet Large achieved highest accuracy of 97.5% along with 96.2% precision, 96.2% recall, an AUC of 0.99 and a PRC of 0.99. The use of transfer learning provided a significant boost to performance of the model. The proposed method of using scalograms with transfer learning can be used to accurately diagnose gait abnormalities and allow medical professionals to design treatment and rehabilitation plan.

Keywords Electromyography · Scalogram · Gait analysis · Transfer learning

P. C. B. S. Negi (✉) · S. Sharma · N. Sharma
Indian Institute of Technology (Banaras Hindu University), Varanasi, UP 221005, India
e-mail: pranshucbnegi.rs.bme19@itbhu.ac.in

S. S. Pandey
Institute of Medical Sciences, Banaras Hindu University, Varanasi, UP 221005, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_37

1 Introduction

Gait is the pattern of movement during walking or running. Gait analysis is the study of gait patterns in order to diagnose abnormalities in the muscle [1]. These can be caused by either neurological or non-neurological conditions. Lesions in the central nervous system or peripheral neurons cause neurological gait disorders. Some examples of neurological disorders are Parkinson's disease, Hemiplegia, Amyotrophic Lateral Sclerosis, and Huntington's disease [2].

On the other hand, musculoskeletal problems or traumatic spinal or brain injury mainly contribute to the onset of non-neurological gait abnormalities [3]. A prominent example of such a severe condition is Rheumatoid Arthritis (RA), a chronic inflammatory illness that causes pain, swelling, and stiffness in the joints along with an altered posture and slower pace when walking [4]. Another prevalent problem is prolapsed intervertebral disc (PIVD) or herniated disc. PIVD causes back pain and alteration of gait [5]. It is more common in adults over the age of 40, affecting 1–3% of the total population, while RA is common in middle aged women [6].

Gait analysis can be performed using visual observation, force plate measurements, motion capture technology, and sensor-based methods. Sensor-based approaches employ modalities such as Inertial Measurement Units (IMUs) and Electromyography (EMG) [7]. EMG is a method of measuring electrical activity in muscles. Researchers can better understand the underlying processes of gait and identify possible areas for improvement in patients with gait disorders by assessing muscle activation during walking [8]. Researchers and clinicians can develop targeted therapies to improve gait function by recognizing these aberrant muscle activation patterns [9].

The analysis of EMG signals can be carried out using signal-processing techniques [10]. The standard signal processing with respect to EMG signal involves the following steps: (i) Filtering: To remove unwanted noise from the EMG signal, (ii) Rectification: To convert the EMG signal to extract its absolute, and (iii) Feature extraction: To identify specific features of the EMG signal in time, frequency, and spatial domain, that helps in classification of data/disease. One such feature extraction technique is wavelet transform. Once the wavelets are obtained, a scalogram can be generated that shows relationship between frequency and time scalograms help identify signal patterns that may not be apparent in the raw data [11]. The nature of generated scalogram depends of several different types of wavelets like Morlet, Mexican hat, and Haar wavelets. The choice of wavelet depends on the characteristics of the signal being analyzed and the specific features of interest [12]. Once the entire processing step is done, the data is split into training and testing dataset and fed into the classifier.

One of the most used types of neural networks in deep learning is the convolutional neural network (CNN). A bulk of image classification task is performed using CNNs due to their ability to learn spatial hierarchies of features in an image [13]. A CNN consists of (i) Convolution Layer, which extract local features such as edges and corners, (ii) Pooling Layer, which pools of combines the pixels in each grid, (iii) Fully Connected Layers, perform computation on the extracted image features to generate predictions [14]. A popular machine learning technique that is currently being used by researchers is transfer learning. Transfer learning involves the integration of a pre-trained network on one task with a network to perform a different classification task. One main advantage of transfer learning is that it significantly cuts the amount of data required for new task [15]. Due to this, researchers can quickly train image classifiers with a decent accuracy. Some of the popular transfer learning architectures are: (i) VGG16: A convolutional neural network with 16 layers, trained on the ImageNet dataset. (ii) VGG19: A convolutional neural network with 19 layers, trained on the ImageNet dataset. (iii) ResNet50: A deep residual network with 50 layers, trained on the ImageNet dataset. (iv) InceptionV3: A convolutional neural network with inception modules, trained on the ImageNet dataset. (v) InceptionResNetV2: A deep residual network with inception modules, trained on the ImageNet dataset. (vi) MobileNet: A lightweight convolutional neural network, trained on the ImageNet dataset. (vii) MobileNetV2: A lightweight convolutional neural network with improved performance, trained on the ImageNet dataset. (viii) DenseNet: A densely connected convolutional neural network, trained on the ImageNet dataset. These networks are pre-trained on millions of parameters are easily accessible through TensorFlow [16] library or GitHub.

In the present study we have implemented classification of EMG scalogram signals using both a simple CNN and comparing it with the transfer learning architectures mentioned above. The accurate diagnosis of gait abnormality can help medical professionals in devising interventions and rehabilitation plans and improve the quality of life of patients.

2 Materials and Methods

2.1 Data Collection and Preprocessing

In this study, the participants were selected from Sir Sunderlal Hospital (IMS), Varanasi, and other local clinics. The Institute Ethics Committee at the Institute of Medical Science (Banaras Hindu University) approved the study, and the subjects provided proper consent. A total of five subjects participated, including one with Rheumatoid arthritis and two with PIVD. Two healthy subjects were included as controls. The subjects ranged in age from 25 to 69 years old, with an average weight of (75 $\pm$ 12) kg and an average height of (166 $\pm$ 6) cm. The study recorded EMG data from two major leg muscles, the tibialis anterior (TA), and medial gastrocnemius

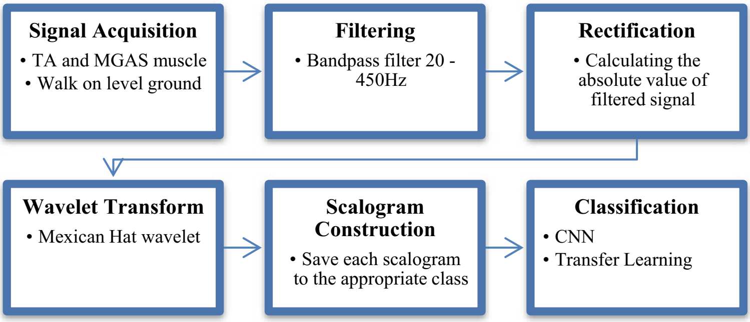


Fig. 1 Flowchart of present study

(MGAS), using two DataLITE Wireless Surface EMG LE230 devices (Biometrics Ltd., Newport, United Kingdom) with a sampling frequency of 1000 Hz. During each trial, the subjects were asked to walk on level ground without support. The recorded data was processed using Python programming libraries, including Numpy, Pandas, PyWavelet, and SciPy [17], and the waveforms were visualized using the matplotlib library. The complete process is shown in Fig. 1.

2.2 *Scalogram Generation*

The continuous wavelet transform (CWT) analyses signals in the time–frequency domain by convolving the signal with a series of scaled and shifted wavelet functions [18]. It generates a 2D plot called a scalogram, with time on the x-axis and frequency on the y-axis, showing the magnitude of the wavelet coefficients. For this study, Mexican Hat Wavelet was used [19]. The continuous wavelet transform (CWT) of a function $f(t)$ is given by the convolution of $f(t)$ with a scaled, and a translated version of the wavelet function $\psi(t)$ is given as:

$$\mathrm{CWT}(f, a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{b}\right)\mathrm{d}t \tag{1}$$

where a is the scale factor and b is the translation factor. The scalogram of the function representing energy at scale s is defined as:

$$S(a) = \|Wf(a, t)\| = \sqrt{\int_{-\infty}^{\infty} f(t)|Wf(a, t)|^2\mathrm{d}t} \tag{2}$$

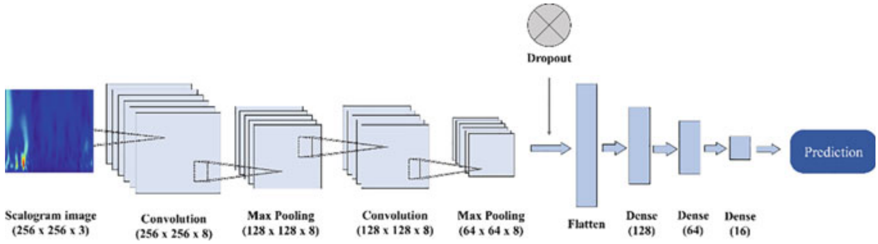


Fig. 2 Schematic diagram of a CNN

Once the scalogram images were obtained, they were split into training and testing datasets in the ratio of 75% training images to 25% testing images. A total of 1000 scalogram images of all classes were obtained.

2.3 Transfer Learning

For a comparative perspective the scalogram images were first classified using a generic CNN with convolutional and pooling layer. The hyperparameters of the networks were extensively tuned beforehand. For the CNN, categorical crossentropy is used as loss function and adam is used as optimizer. Further, following architectures were used in present study in conjunction with CNN: VGG16, VGG19, ResNet50, Inceptionv3, InceptionResNet, MobileNet and MobileNetv2, and DenseNet Large. Figure 2 shows the architecture of the CNN used in present study. The models are downloaded using TensorFlow library and the code is written in Python. The classification was done on an Alienware 15R3, with 32 GB RAM and 6 GB Nvidia GTX 1060 GPU. Figure 2 shows the basic schematic of CNN.

3 Results and Discussion

3.1 EMG Waveforms and Scalogram

Figure 3 shows a comparison of sEMG waveforms on each step of signal processing. The difference can be easily observed between healthy and affected EMG waveforms. The pattern irregularity depicts abnormal muscle activity and function in the form of jagged peaks. These abnormal signals indicate various conditions, including muscle weakness, spasms, tremors, and fatigue. Figure 4 shows the continuous wavelet transform of EMG signal along with its scalogram for both muscles.

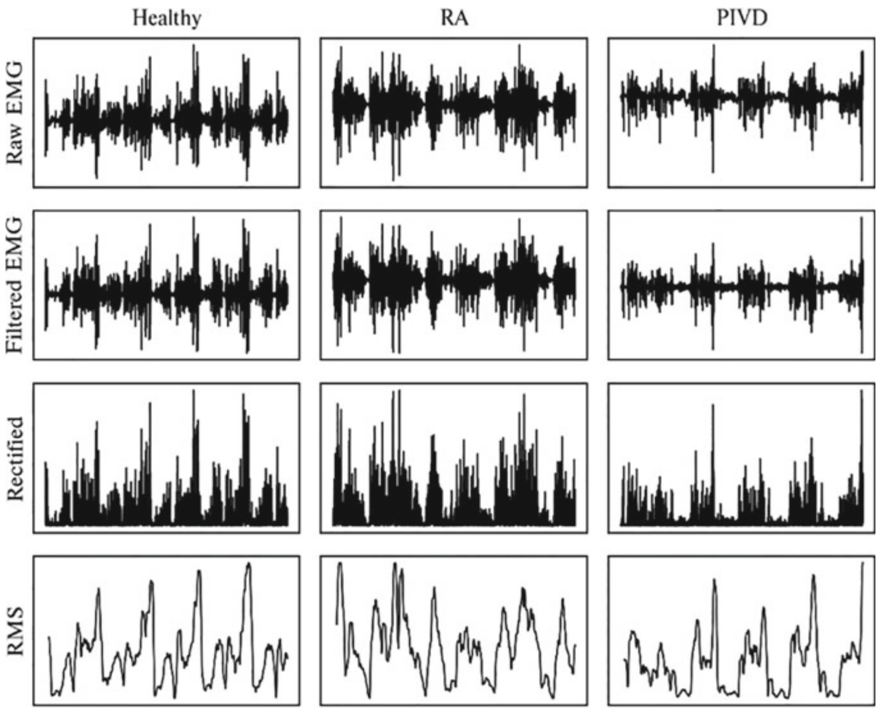


Fig. 3 sEMG waveform on each step of signal processing

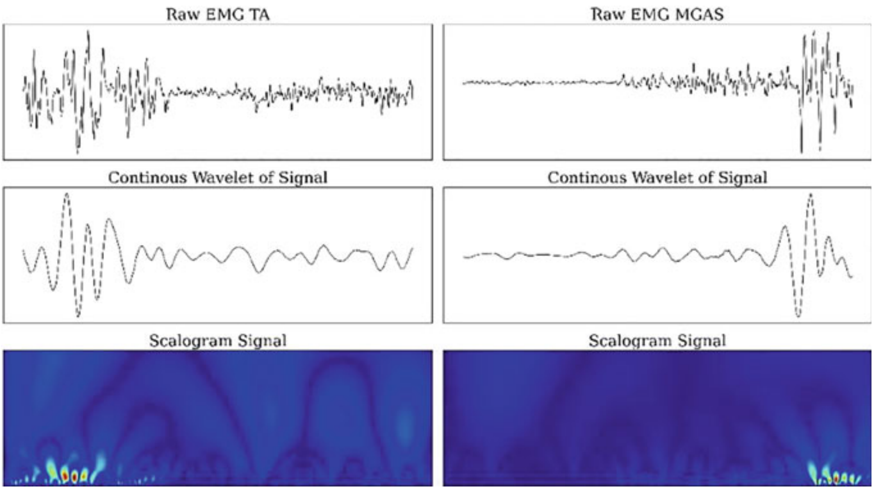


Fig. 4 sEMG waveform with CWT and scalogram

3.2 Classification

The results of CNN training and transfer learning are shown in Table 1. To obtain all aspects of model performance, several metrics apart from accuracy are considered. These are precision, recall, AUC score, and PRC.

Here, for CNN an accuracy of 91.1%, precision of 92.8%, recall of 92.9%, AUC of 0.98, and a PRC of 0.97 were obtained. Further, when transfer learning architectures were incorporated, DenseNet Large achieved the highest accuracy of 97.5% along with 96.2% precision, 96.2% recall, an AUC of 0.99, and a PRC of 0.99. It is evident from the results that incorporation of pre-trained architectures into the CNN greatly enhances the accuracy of classification. A comparative analysis of our proposed model with existing studies is highlighted in Table 2.

Table 1 Classification results of CNN and transfer learning models

Model	Accuracy (%)	Precision (%)	Recall (%)	AUC	PRC
CNN only	91.1	92.8	92.9	0.98	0.97
VGG16	92.4	88.6	88.6	0.9816	0.9657
VGG19	93.2	89.8	89.8	0.9795	0.9607
ResNet50	95.7	93.6	93.6	0.9861	0.9771
Inceptionv3	89.8	84.8	84.5	0.9568	0.9302
InceptionResNet	80.4	75.6	61.0	0.8617	0.7714
MobileNet	93.4	90.2	90.2	0.9806	0.9615
MobileNetv2	94.4	91.7	91.7	0.9867	0.977
DenseNet large	97.5	96.2	96.2	0.9982	0.9965

Table 2 Comparison of present work with existing literature

Authors	Task	Model	Result (%)
Fricke et al. [20]	Gait abnormality	CNN, SVM, KNN	97.3
Katiyar et al. [21]	Gait abnormality	Self-organizing maps	91
Badura et al. [22]	Pain state classification	AdaBoost	85
Xiong et al. [23]	Gait tracking	PCA and regression	87.1
Popescu et al. [24]	Gait in Parkinson's disease	CNN	96.5
Le et al. [25]	Gait analysis	Autoencoder	96.2
Negi et al. [26]	Terrain walk	SVM and DNN	95
Zhang et al. [27]	Multi-modal gait analysis	CNN	96.05
Mokdad et al. [28]	Gait abnormality	SVM, LDA, and KNN	95.8
Kumar et al. [29]	Gait abnormality	RNN	91.3
This study	Gait abnormality	Transfer learning	97.5

The results of the present study suggest that transfer learning architecture is an effective tool for classifying EMG recordings using scalograms. In almost every case, the performance of model has been improved. By enhancing the feature extraction and classification capabilities of the CNN, the proposed method can enable researchers and medical professionals to identify and treat a wide range of diseases and conditions more effectively.

4 Conclusion

In the present study authors classified the scalogram images obtained from EMG signals of healthy subjects and patients with RA and PIVD. For a comparative analysis, eight popular networks namely, VGG16, VGG19, ResNet50, Inceptionv3, InceptionResNet, MobileNet and MobileNetv2, and DenseNet Large were utilized. When standalone CNN is used for classification, an accuracy of 91.1%, precision of 92.8%, recall of 92.9%, AUC of 0.98, and a PRC of 0.97 were obtained. By combining CNN with other models, the general accuracy saw considerable improvement. The highest accuracy of 97.5% was given by DenseNet Large with 96.2% precision, 96.2% recall, an AUC of 0.99, and a PRC of 0.99. The proposed transfer learning model, has performed exceptionally well for classifying EMG data from patients with RA and PIVD. Analysis and classification of sEMG signals, is of great importance to clinicians and researchers, and can be further utilized for treatment planning, rehabilitation and for evaluation of treatment outcomes.

References

1. Winter DA (1984) Kinematic and kinetic patterns in human gait: variability and compensating effects. Hum Mov Sci 3(1–2):51–76
2. Hausdorff JM (2005) Gait variability: methods, modeling and meaning. J Neuroeng Rehabil 2(1):1–9
3. Perry J, Davids JR (1992) Gait analysis: normal and pathological function. J Pediatr Orthop 12(6):815
4. Turner DE, Helliwell PS, Siegel KL, Woodburn J (2008) Biomechanics of the foot in rheumatoid arthritis: identifying abnormal function and the factors associated with localised disease 'impact'. Clin Biomech 23(1):93–100
5. Humphreys SC, Eck JC (1999) Clinical evaluation and treatment options for herniated lumbar disc. Am Fam Physician 59(3):575
6. Cross M, Smith E, Hoy D, Carmona L, Wolfe F, Vos T, Williams B, Gabriel S, Lassere M, Johns N, Buchbinder R (2014) The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis 73(7):1316–1322
7. Muro-De-La-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A (2014) Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. Sensors 14:3362–3394
8. Carlo J (2002) De Luca. Surface electromyography: Detection and recording
9. Moon Y, Sung J, An R, Hernandez ME, Sosnoff JJ (2016) Gait variability in people with neurological disorders: a systematic review and meta-analysis. Hum Mov Sci 47:197–208

10. Singh RE, Iqbal K, White G, Holtz JK (2019) A review of EMG techniques for detection of gait disorders. Artif Intell-Appl Med Biol
11. Bialasiewicz JT (2015) Application of wavelet scalogram and coscalogram for analysis of biomedical signals. In: Proceedings of the world congress on electrical engineering and computer systems and science, vol 333
12. Rioul O, Vetterli M (1991) Wavelets and signal processing. IEEE Signal Process Mag 8(4):14–38
13. Negi PC, Negi S, Sharma N (2022) Gait analysis-based identification of neurodegenerative diseases using machine learning techniques. In: 2022 international conference on advances in computing, communication and materials (ICACCM). IEEE, pp 1–6
14. O'Shea K, Nash R (2015) An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458
15. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H, He Q (2020) A comprehensive survey on transfer learning. Proc IEEE 109(1):43–76
16. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S (2016) Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467
17. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, Van Der Walt SJ (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 17(3):261–272
18. Orfanidis SJ (1995). Introduction to signal processing. Prentice-Hall, Inc
19. Kilby J, Hosseini HG (2006) Extracting effective features of SEMG using continuous wavelet transform. In: 2006 international conference of the IEEE engineering in medicine and biology society. IEEE, pp 1704–1707
20. Fricke C, Alizadeh J, Zakhary N, Woost TB, Bogdan M, Classen J (2021) Evaluation of three machine learning algorithms for the automatic classification of EMG patterns in gait disorders. Front Neurol 12:666458
21. Katiyar R, Pathak D, Kumar V (2010) Clinical gait data analysis based on spatio-temporal features. arXiv preprint arXiv:1003.1511
22. Badura A, Masłowska A, Myśliwiec A, Piętka E (2021) Multimodal signal analysis for pain recognition in physiotherapy using wavelet scattering transform. Sensors 21(4):1311
23. Xiong D, Zhang D, Zhao X, Chu Y, Zhao Y (2021) Synergy-based neural interface for human gait tracking with deep learning. IEEE Trans Neural Syst Rehabil Eng 29:2271–2280
24. Popescu N, Channa A, Ifrim R, Neuro-cognitive evaluations using deep learning and wearable sensorial devices. New Approaches 26
25. Le HT, Phung SL, Bouzerdoum A (2018) Human gait recognition with micro-Doppler radar and deep autoencoder. In: 2018 24th international conference on pattern recognition (ICPR). IEEE, pp 3347–3352
26. Negi S, Negi PC, Singh DBV, Sharma N (2021) Comparative analysis of SVM and DNN for multiple terrain classification using hybrid sensor. In: Proceedings of integrated intelligence enable networks and computing: IIENC 2020. Springer Singapore, pp 317–325
27. Zhang YP, Cao GZ, Ling ZQ, He BB, Cheng HR, Li WZ, Cao SB (2021) A real-time gait phase recognition method based on multi-information fusion. In: 2021 18th international conference on ubiquitous robots (UR). IEEE, pp 249–255
28. Mokdad A, Debbal SMEA, Meziani F (2020) Application of the continuous wavelet transform for the analysis of pathological severity degree of electromyograms (EMGs) signals. Polish J Med Phys Eng 26(3):149–154
29. Kumar M, Gautam DP, Bhaskar DV (2022) Effect of machine learning techniques for efficient classification of EMG patterns in gait disorders. IJEER 10(2):117–121

OSM-Based Indoor Positioning System

Ruchi Chakkarwar, Pratika Kamath, Isha Kulkarni, Nithi Poojari, and Pinki Vishwakarma

Abstract Most people rely on the Global Positioning System (GPS) to provide location information, weather forecasts, and navigation. GPS devices use four or more satellites to determine location, which works with great accuracy in outdoor environments and is vital for users worldwide. However, GPS is not efficient for indoor conditions due to the need for an unobstructed line-of-sight connection with satellites. With the construction of buildings over 3 million sq ft floor areas worldwide, there is a real need for positioning and navigation in indoor spaces. Researchers are exploring alternative systems for GPS, using popular wireless technologies like Bluetooth, Wi-Fi, RFID, and Infrared. This proposed work aims to develop a mobile application that can estimate a user's position within a building using Wi-Fi technology.

Keywords Wi-Fi positioning · Indoor position · Positioning using Wi-Fi

R. Chakkarwar (✉) · P. Kamath · I. Kulkarni · N. Poojari · P. Vishwakarma
Shah and Anchor Kutchhi Engineering College, Mumbai, India
e-mail: ruchi.chakkarwar_19@sakec.ac.in

P. Kamath
e-mail: pratika.kamath_19@sakec.ac.in

I. Kulkarni
e-mail: isha.kulkarni_19@sakec.ac.in

N. Poojari
e-mail: nithi.poojari_19@sakec.ac.in

P. Vishwakarma
e-mail: pinki.vishwakarma@sakec.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_38

1 Introduction

Global Positioning System (GPS) is of much importance in this day and age, from content recommendation according to location to news of the geographical area you currently are in; GPS has become a significant discovery in the world of technology. GPS is not suitable for indoor positioning because it relies on a direct line-of-sight connection to at least four satellites to determine location accurately [1, 2]. However, indoor environments with walls, roofs, and other obstructions block the GPS signals, leading to inaccurate or nonexistent location information. The GPS signal weakens considerably when it passes through walls or other physical barriers, and it can be affected by reflections and signal multipath in indoor environments. As a result, GPS receivers may not be able to receive a strong enough signal to determine location accurately indoors, making GPS impractical for indoor positioning.

One Nine Elms (UK) published that around 100 million buildings are present in the world as of 2022 [3], which indicates a huge need for indoor positioning in indoor spaces.

This work proposes an application that uses Wi-Fi signals to determine the indoor position of the user. At present, there is no widely accepted standard for indoor positioning systems that are equivalent to outdoor GPS. Nevertheless, Wi-Fi signals have been proposed as a potential solution for achieving indoor positioning, and numerous proposals have been put forward using this technology. Many of these proposals utilize machine learning techniques to estimate a user's position based on Wi-Fi signal strength and other factors. By analyzing the signal strength of Wi-Fi access points and their locations within a building, machine learning algorithms can create a model that estimates the user's position with reasonable accuracy. While these approaches are not yet as accurate as GPS, they offer a promising way to provide indoor positioning in a variety of settings [1, 4].

Radio Signal Strength (RSS) values from Wi-Fi at a particular location are used to determine the current location of the user [2]. Wi-Fi-based positioning is easier and has attracted a lot of attention in the wireless transmission industry [1, 5].

2 Related Work

Liu et al., here the authors provide a thorough outline of these two categories of positioning strategies after analyzing numerous outstanding articles in the relevant field. Also, they analyze the difficulties and potential directions for advancement in the current technological environment [1].

Ranimol et al., an indoor location method employing Wi-Fi fingerprint is presented in this research. A database is used to measure and record the received signal's intensity. The database contains a list of signal strengths and the places where each one is located. The user's position can be determined using the similarity between the signal that was detected and the ones that were saved in the database. For this, it is possible to leverage the current Wi-Fi infrastructure [2].

Vladimir et al., the initiatives between 2019 and 2021 that use machine learning algorithms for determining indoor position and Wi-Fi as the localization signal are reviewed in this article. Sixty-four papers that were deemed pertinent were found, and they were all carefully examined to gain a better picture of the state of affairs from various angles. The findings demonstrate how neural network-based models, assessed through practical trials, are used for interior positioning based on Wi-Fi patterns. Considering this, many works continue to do a small-area evaluation, which can affect how well the findings are communicated [5].

Sukhjit et al., provides us an assessment on OpenStreetMap. They performed various quality assessments using various parameters and how OpenStreetMap is generating a huge dataset with the help of non-commercialized users, with varying levels of experience. So the assessment becomes vital to give maturity to OpenStreetMap data [4].

Jamal et al. is a collection of experiences and research which has been carried out with OSM as the central and core theme. The book addresses various research topics like applications of OSM in different domains and so on [6].

Md. Rashidujjaman et al., introduces us to a location-based application that works as a audio road guider for both the normal and visually impaired people using OpenStreetMap. In this paper they have discussed the challenges and features of the proposed location-based information system as an example of the usability of OpenStreetMap. They have also discussed the problems faced while developing OSM in a developing region like Bangladesh. Finally, they present the future scope of various location-based services using OpenStreetMap [7].

Repasky et al., the researchers propose a novel technique to improve the existing OpenStreetMap database using aerial imagery. Preliminary tests conducted with the proposed model in their results showed the correlation between the discrimination information measure and proxy for database quality by the authors. In spite of this further investigation is required to check the robustness and nature of this correlation [8].

Shuai et al., in this article, a support vector machine (SVM)-based technique for Wi-Fi indoor 2.5D positioning is put forth. Prior to obtaining the position along the vertical direction, the technique first uses the Wi-Fi signal of each floor to identify the floor by SVM. Then, using a collection of Wi-Fi position fingerprints, the weighted K-nearest neighbor (WKNN) algorithm is used to estimate the plane's location. As a result, three-dimensional movement inside is made possible [9].

Dong et al. this study suggests a brand-new Encoded LSTM network model that can efficiently obtain key features from Wi-Fi RSS data and offer high precision for positioning. According to experimental findings based on a crowd sourced open-source dataset, the encoded LSTM is able to reduce the Wi-Fi fingerprints' dimension from 992 to 64 and obtain a mean positioning error of 7.37 m. The encoded LSTM outperforms 13 commonly utilized positioning algorithms when compared to standard findings based on the same dataset, showing the lowest mean error. In comparison to other traditional LSTM models, the suggested encoded LSTM can also offer a 10% increase in positioning accuracy [10].

Joseph et al., a Wi-Fi multi-mode fusion localization technique using images is suggested in this paper. One of the most popular methods currently in use is Wi-Fi fingerprint 10-calization, but Wi-Fi fingerprint localization suffers from multipath and non-line-of-sight phenomena as well as signal shifts, scattering, and other issues, which results in an unreliable localization impact [2].

Grgić et al., this paper shows and evaluates a practical application of potential Wi-Fi network use for indoor localization. Different worldwide navigation satellite systems typically provide navigation and localization services in outdoor environments (GNSS). There is no line-of-sight between the satellite and the receiver, so indoor localization must depend on alternative methods. In interior spaces, Wi-Fi networks are widely used and can be used for indoor positioning [11].

Haryanto et al., this study contrasts Wi-Fi and geomagnetism as two IPS techniques. The accuracy of the methods used in public spaces is compared in this study. The Earth's Magnetic Fingerprint, which is produced by the HCM8533L instrument, is used in the geomagnetism technique. The Wi-Fi approach makes use of the NodeMCU device's Received Signal Strength (RSS). The position references make use of nearby, operational access locations. All of the noise and disturbances are caused by actual road noise. *K*-nearest neighbors (KNN) and first position are the methods for magnetic fingerprinting. The multi-lateration algorithm is used in the Wi-Fi approach. Every 1.44 m^2, Geomagnetism IPS employs 36 fingerprints [12].

3 Proposed Methodology

A prototype Android application for an indoor positioning system that runs purely on the user's smart phone was created. Using a prepared radio map and the device's built-in Wi-Fi chipset, the application enables locating the location of the device. The application works in the two location fingerprinting phases: calibration phase and positioning phase (Table 1)

A. **Calibration Phase**

Initially, information is acquired about each individual room in the structure. Different locations can be distinguished using the Wi-Fi network inside a building

Table 1 Notation

Notation	Description
q_i	The geographical coordinates of *i*th location
r_i	The *m* RSS values from *m* access points at that location
d	The Euclidean separation between the readings
w_j	The weight of the current coordinates

since they would experience varying Wi-Fi signal levels from different routers. The calibration phase measures the Radio Signal Strengths (RSS) of Wi-Fi access points at various locations within the structure, and this information is later updated in the server database [2, 4]. Each of the n measurements is a tuple and is included in the radio map:

$$(q_i, r_i) \quad i = 1, 2, \ldots n$$

where

$$q_i = (x_i, y_i)$$

are the geographical coordinates of ith location

$$r_i = (r_{i1}, r_{i2}, \ldots r_{im})$$

are the m RSS values from m access points at that location.

The user can choose the dependable access points for that specific building during the calibrations process. This disqualifies readings from other access points from being taken into account.

The application's functionality during calibration includes the following:

(1) Add or remove a building.
(2) Viewing, adding, or removing points for the chosen building.
(3) Use the point from the list to do fingerprinting.
(4) Modify the list of accessible access points by adding or removing dependable access points.

After the information was gathered, the location estimator algorithm used the RSS values from the unknown position to estimate the user's position. This is done in the positioning phase.

B. **Positioning phase**

The application establishes the precise position during the positioning step. To determine the closest neighbor, the RSS values of all the sensed APs are measured and compared to those on the produced radio map. The floor map will display the color grid of the closest neighbor. This approach merely makes use of the radio map as a look-up table.

The weighted K-nearest neighbors algorithm is the location estimation algorithm in use. It involves following steps:

I. **Discover your k closest neighbors in the radio map**

With the current reading of r, rank the n points in ascending order of Euclidean distance. When calculating Euclidean distance, the readings are treated as vectors.

The sorted list's first k readings should be returned.

$$(q_1, r_1), (q_2, r_2) \ldots (q_k, r_k)$$

II. **Determine the current unknown location's coordinates**

Utilize the formula to determine coordinates.

$$q = \sum_{j=1}^{k} \frac{w_j q_j}{\sum_{l=1}^{k} w_l}$$

where all the weights are non-negative

$$w_j = d^{-1}_{(r_i, r)}$$

The Euclidean separation between the readings is known as d.

The coordinates for the jth position are q_j.

The number of nearest neighbors considered (k), a tuning parameter for WKNN, regulates how local the location calculation is. The approach functions as a straightforward look-up table for $k = 1$. The location will be assumed to be anywhere between the calibration points for greater values.

Functionality while positioning:

(1) Load and examine building map.
(2) Calculate the position (in the form of grids as marked in the loaded map).

The measurements made during the calibration phase and the list of friendly Wi-Fi networks are stored on a server, allowing new users to skip this step by downloading the already calibrated measurements. To determine the building and floor map that will be loaded in the programmed dynamically, the Wi-Fi access points around the new device are compared to the saved points at the time of positioning (Fig. 1).

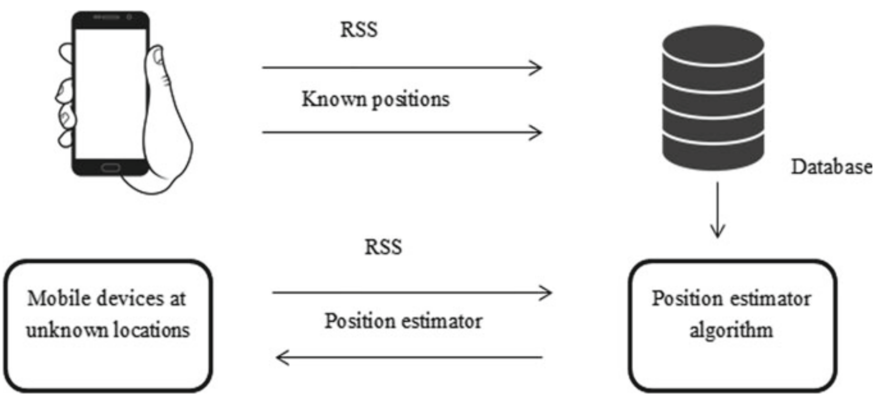


Fig. 1 High level system overview

Practical Difficulties Encountered

The system's level of accuracy is influenced by a number of variables; they were found and their impacts were reduced to the greatest extent possible.

I. **Readings of fluctuating received signal strength (RSS)**

Readings collected from a specific area may vary often and lead to mistakes. This was fixed by averaging numerous measurements taken at the same place.

II. **Finding trustworthy Access points**

The access points that were taken into consideration for placing should be an integral component of the structure permanently and ideally be accessible at all times. Temporary hotspots should not be incorporated into the system since moving them could have a negative effect on the computations.

III. **User orientation**

The strength of the signal received at that place can change depending on how the user holds their smartphone. This can be explained by the user's body blocking some of the impulses coming from a certain direction. In order to account for this effect, orientation-specific data were acquired during the calibration phase. However, even after having unique readings for the four cardinal directions, the expected effect could not be observed. The program had to periodically determine the device orientation from the magnetometer and accelerometer sensors, which proved to be a computationally demanding aspect of this technique.

IV. **Proper positioning of Access points**

A building's access points should be placed so that each site may be uniquely identified using the Wi-Fi fingerprint discussed before. However, there are restrictions on what can be done in this situation.

V. **Normalizing RSS Values**

The RSS value that each device can produce at a specific place varies. This is as a result of Android devices having lower-quality hardware. In order for the application to operate correctly across a variety of devices, the readings need to be properly normalized. For the application to be properly normalized, it must be tested on more than 9000 Android OS-equipped smart phones available on the market.

4 Result Analysis

The main objective of any positioning system is to determine the user's location, so the accuracy of positioning is one of the key parameters to evaluate the performance of indoor positioning techniques.

Using a predefined radio map and the device's built-in Wi-Fi chipset, the application enables pinpointing the location of the device. The application functions in the calibration phase and positioning case for the two location fingerprinting phases. The user can choose the dependable access points for that specific building during the calibrations process. This excludes readings from other access points from being taken into account. The available Wi-Fi access point RSSs are measured from various locations across the structure during the calibration process. When the measurements were obtained for 30 s (during which time readings became significantly more precise), an average value was calculated and recorded in the radio map. The application provides the accurate location during the positioning stage. The RSS values of all sensed APs are measured and compared to the values on the generated radio map to identify the nearest neighbor.

5 Conclusion

Wi-Fi positioning systems can be a feasible solution for indoor location tracking as it is easy to scale with minimal manual intervention. The research explored a highly useful method for developing a GPS replacement that would be appropriate and functional in enclosed spaces. The technology could be implemented on practically all buildings such as shopping malls, hospitals, university campus, airports, without incurring new infrastructure costs since the system uses Wi-Fi signals to estimate a user's location within around 2 m. At this time, the system can only determine the user's position. The position acquired can be utilized to direct the user to his favorite location inside the building by adding a navigation module to the application. Not only can this system also be coupled with augmented reality modules to provide the users with interactive information.

References

1. https://doi.org/10.1049/iet-com.2019.1059. Access Time: 30th March 2023, 22:00
2. Joseph R, Sasi SB (2018) Indoor positioning using Wi-Fi fingerprint. In: 2018 international conference on circuits and systems in digital enterprise technology (ICCSDET), Kottayam, India
3. https://www.onenineelms.co.uk/2022/09/13/how-many-buildings-are-there-in-the-world-2022. Access Time: 25th March 2023, 13:53
4. Assessment of OpenStreetMap Data-A review, IJCA/2013/Sukhjit Sehra, Hardeep Rai, Jaiteg Singh
5. Bellavista-Parent V, Torres-Sospedra J, Perez-Navarro A (2021) New trends in indoor positioning based on Wi-Fi and machine learning: a systematic review. In: 2021 international conference on indoor positioning and indoor navigation (IPIN), Lloret de Mar, Spain
6. Introduction to Open Street map in Geographic Information Science: Experiences, Research and Applications, Open Street Map in GI science /2015/Jamal Jokar Arsanjni, Marco Helbich, Alexander Ziph
7. Location based Information System Md. Rashidujjaman Rifat, Shubrami Moutushy, Syed Ishtiaque Ahmed, Hasan Shahid Ferdous Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology

8. Repasky B, Payne DT, Dick DA (2020) Automatic assessment of open street maps database quality using aerial imagery. In: IEEE, 2020 digital image computing: techniques and applications (DICTA)
9. Zhang S, Guo J, Wang W, Hu J (2018) Indoor 2.5D positioning of Wi-Fi based on SVM. In: 2018 ubiquitous positioning, indoor navigation and location-based services (UPINLBS), Wuhan, China
10. Dong Y, Arslan T, Yang Y (2022) An encoded LSTM network model for Wi-Fi-based indoor positioning. In: 2022 IEEE 12th international conference on indoor positioning and indoor navigation (IPIN), Beijing, China
11. Grgić K, Balić L, Križanović V, Žagar D (2022) An example of indoor positioning possibility using Wi-Fi network and mobile phone. In: 2022 international conference on smart systems and technologies (SST), Osijek, Croatia
12. Haryanto DK, Karyono K, Hutagalung S (2018) The comparison between geo-magnetism and Wi-Fi for indoor positioning system for public places. In: 2018 IEEE international conference on robotics, biomimetics, and intelligent computational systems (robionetics), Bandung, Indonesia
13. https://www.openstreetmap.org/. Access Time: 30th March 2023, 20:00
14. https://play.google.com/store/apps/details?id=net.osmand. Access Time: 30th March 2023, 20:45

A Multilayered Secure Image Steganography Technique for Resisting Regular-Singular Steganalysis Attacks Using Elliptic Curve Cryptography and Genetic Algorithm

Richard Apau, James Ben Hayfron-Acquah, Michael Asante, and Frimpong Twum

Abstract Increasing cyberattacks and data breaches have called for innovative approaches to secure data communication. Although many methods have been proposed, cryptography and steganography have gained the most popularity for protecting the security of secret information. While cryptography works by scrambling the content of a secret message to provide data privacy, steganography focuses on masking secret message existence to ensure secrecy. Fundamentally, cryptography reinforces the data protection capabilities of steganography. However, the ability of steganography to provide data protection for secure data communication has been undermined by steganalysis. Statistical steganalysis mechanisms identify, detect, and accurately estimate the presence of steganography-based hidden information. Regular-singular (RS) analysis scheme is one of the well-known statistical analysis methods that is used to detect steganographic techniques that employs least significant bit (LSB) insertion algorithm for secret message embedding. This paper's objective is geared towards improving the robustness, imperceptibility, and visual quality of embedded image using secure methods by minimizing MSE values and maximizing PNSR values. Genetic algorithm (GA) is utilized to adjust the stego-image pixel values by identifying appropriate bits within the cover image pixel values through the selection of fittest best chromosome. The security is achieved by encrypting and decrypting keys generated using elliptic curve cryptography (ECC). The technique proposed in this paper proved effective and robust against RS statistical steganalysis attacks as it provides multilayered secure data protection capable of resisting image distortion and statistical detection as shown by RS steganalysis and PSNR results.

Keywords Steganography · Cryptography · ECC · LSB · Regular-singular (RS) steganalysis · Genetic algorithm · Imperceptibility · PSNR

R. Apau (✉) · J. B. Hayfron-Acquah · M. Asante · F. Twum
Kwame Nkrumah University of Science and Technology (KNUST), PMB, Kumasi, Ghana
e-mail: rapau@st.knust.edu.gh

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_39

1 Introduction

Cryptographic and steganographic applications are popular technological interventions for protecting data and information [1–3]. These applications secure data based on cryptography and steganography concepts. Cryptography denotes the act of protecting messages from unauthorized agents often by changing its meaning and the principle of steganography is concerned with concealing the contents of a secret message [3], i.e. steganography improves cryptography by not only protecting the message but also hiding its existence in total. Due to this, various combinations of cryptography and steganography techniques exist [4–6]. Rivest, Shamir, and Adlenam (RSA), ECC, digital signature algorithms (DSA), and Deffie-Hellman (DH) key exchange are commonly used cryptographic algorithms for steganographic-based applications. However, relevant literature indicates that RSA and ECC are most effective for data protection [7]. As a result, most steganographic-based applications that employ cryptography (REF) are based on RSA and ECC. Similarly, like cryptography, steganographic schemes include image, video, audio, text, and network steganography. However, the degree of redundancy and convenience associated with images have made them the most popular covert object for steganography [8].

As indicated earlier, the fundamental principle of steganography is to provide the required security and concealment from unauthorized access including eavesdropping [9]. However, the information security and data protection capabilities of steganography have been undermined by the concept of steganalysis. Steganalysis techniques possess the capabilities to reveal the content of hidden messages embedded within cover image [10] and which can lead to disastrous security incidents and breaches. Statistical steganalysis mechanisms such as regular-singular (RS) identify the presence of secret message hidden by steganography using LSB substitution technique [11]. For secure secret data communication to occur, the resistance of steganographic techniques against RS steganalysis is crucial for optimal assurances of data protection as well as information security.

Indeed, while technological advancements allow for different mechanisms to curtail steganalysis attacks, attackers are equally equipped to develop complex and sophisticated techniques to break steganographic applications, i.e. many steganalysis protocols have been developed and tested to have the ability to break steganographic applications and reveal messages concealed in covert objects [12]. Given the exponential growth of attacks on data and the accuracy of steganalytic applications to reveal hidden data and extract its content, a more robust and secure system is needed to resist RS steganalysis to protect data against detection, modification, and distortion.

In recent times, scholars have begun to explore genetic algorithm (GA) as an alternative approach to providing another layer of security for data communication [12]. GA is a metaheuristic search-based algorithm, belonging to the evolutionary algorithm family that utilizes the principles of natural selection and natural genetics, and has the capabilities to enhance data security. However, the potential of GA to provide better security has not seen much utilization in steganographic techniques and cryptographic schemes to provide the much-needed hierarchical and multilayered security resistance to RS statistical steganalysis. Against this background, the paper aims to develop a secure multilayered image steganographic technique resistance to RS steganalytic attacks using ECC and GA. RS attack has become well-known statistical analysis and gained the popularity of detecting hidden message, hence the need for enhancing data protection capabilities of image steganography. The outcome of the proposed technique is expected to outperform existing methods and provide better security for secret data.

The paper is sectioned as follows: Sect. 2 empirical review of related works, exploring previous proposed techniques and methods. In Sect. 3, the methodology which outlines the proposed method as well as steps and procedures followed in the implementation of proposed technique is introduced. In Sect. 4, the outcome of the experimental results is provided and discussed, while Sect. 5 gives detailed security analysis of the proposed technique. The paper finally ends with Sect. 6, which provides conclusion and suggestions for future research.

2 Related Works

Steganalysis focuses on detecting hidden secret data. Fridrich et al. [13] proposed RS steganalytic attack, with the capability to detect both sequentially and randomly embedded secret messages. Since then, plethora of other RS steganalysis techniques have been proposed with the capacity of estimating the length of embedded messages and accurately detecting and revealing secret messages concealed within a cover image.

Various improved steganographic techniques have been proposed aimed at resisting RS statistical steganalysis. Wang et al. [14] used GA to develop image steganography technique**.** The technique was capable of resisting RS analysis attack. LSB substitution method was employed to hide secret messages within the selected cover image to obtain the resulting stego image. The pixel values of the resulting stego image were altered using GA to maintain its statistical characters. The method achieved better visual quality. Lou and Hu [11] developed improved image steganography method using transformation function based on a reversible histogram transformation. The method based on modified LSB function was aimed at resisting two well-known statistical steganalysis attacks, i.e. RS and X^2-detection attacks. The PSNR values obtained from the results demonstrate resistance to both RS steganalytic attack and X^2-detection methods. Kanan and Nazeri [15] proposed image steganography technique that maintains the cover image visual quality after message embedding. The method uses LSB substitution technique to hide that and uses genetic algorithm to preserve the stego image tunable visual quality. Lossless data compression algorithm was used to compress data before data embedding. The method modelled steganography as a search problem and optimization challenge. The results obtained from the experiment proved that the method achieved high payload capacity with improved PSNR values when compared with other existing methods at the time. Baagyere et al. [12] employed genetic algorithm principles of selection, crossover, and mutation to develop a steganography application that combines the security effects of steganography and cryptography concepts. The method used residue number system (RNS) properties to hide encrypted messages into the cover image. Visual quality perception and statistical method analysis of the resulting stego image using standard key metrics showed the proposed system is robust to steganalysis. More recently, Yadav et al. [16] also designed a robust data hiding scheme based on GA, Hamilton path, and embedding cost estimation. The GA enabled the selection of best fittest individual chromosome whose count value is the minimum bit-flip cost among all individual chromosomes. The secret key is generated using the Hamilton path for data embedding and extraction. The results obtained after experiment show that the technique increased the PSNR results by 7% leading to a PSNR value of 41.8 dB.

The above-reviewed famous works, and some other existing works have provided the benchmark through which this study hinges. This addresses the gaps in literature by improving the data hiding schemes using ECC, GA, and modified LSB insertion to realize optimal data security. The technique of combining ECC, GA, and image steganography has not been extensively explored in literature. Even though plethora of studies exist on the effective combination of GA and image steganography, however studies that securely combines the security effect of ECC, GA, and steganography are limited or non-existent. This paper serves as entry point for explorations.

3 The Proposed Method

A secure multilayered approach is adopted to provide maximum security protection to the secret message. Four different iterative processes are followed, each with an improved method of providing a mechanism to ensure the overall performance. The four phases are data encryption using ECC, modified LSB steganography embedding, application of genetic algorithm, and RS analysis. As already indicated, the idea behind combining steganography and cryptography is to complement the security effect of each other. Thus, data is first encrypted using secret key generated through ECC implementation. Thereafter, the encrypted secret message is hidden within the pixel values of selected cover image. Genetic algorithm is applied to adjust the stego-image pixel values to harden statistical analysis capabilities to reveal the concealed secret message. Finally, performance evaluation of the proposed system is determined using standardized criteria to estimate the imperceptibility of the algorithm against attacks through RS analysis. A comparison of the results obtained is done with few existing methods. Figure 1 shows the step-by-step procedure adopted for embedding messages.

Generally, protocols for the implementation of ECC specify the domain parameters to be used. The application of elliptic curve cryptography works by choosing a maximum prime number value, a curve equation and picking a point that lies on the curve. The chosen point is the public key used for encryption whereas a random

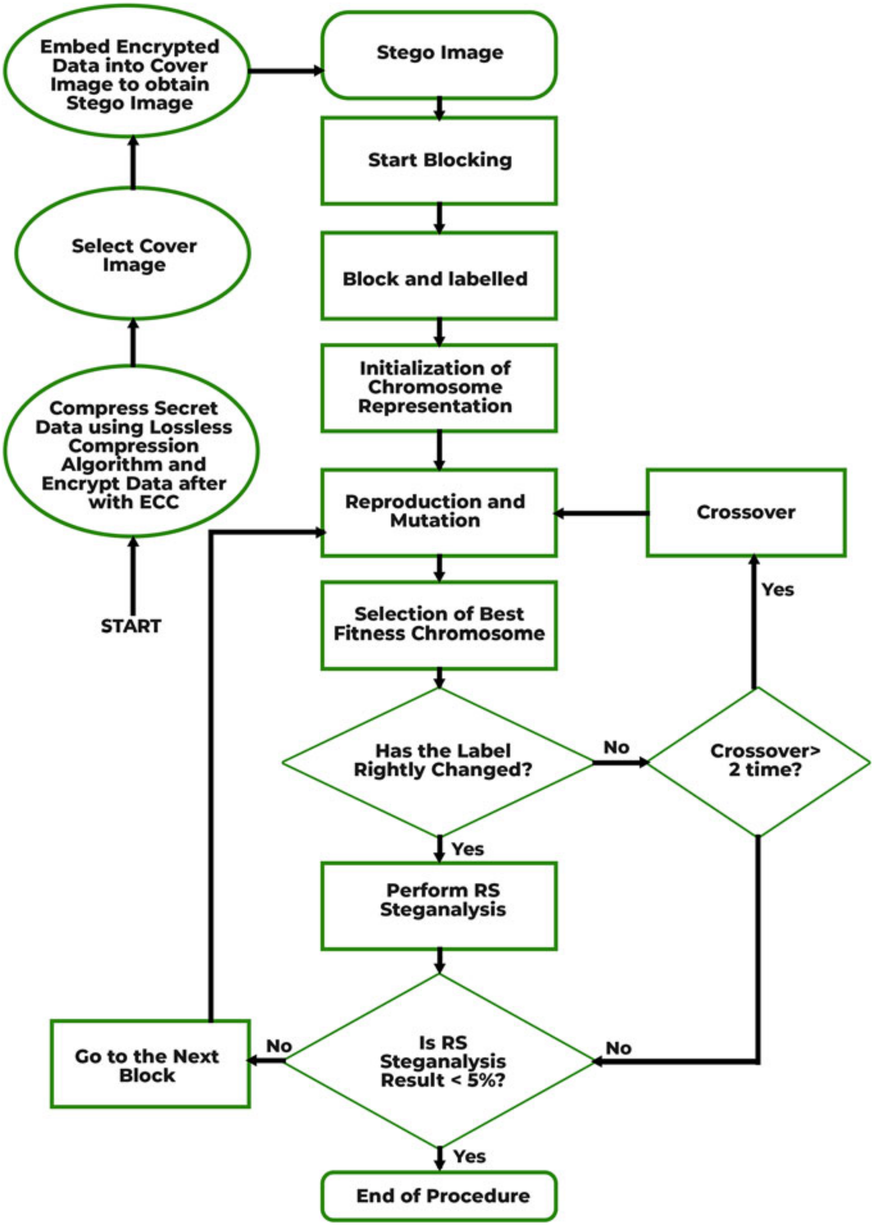


Fig. 1 Algorithm procedure followed for data hiding

number serves as the private key for decryption. The public key is obtained by multiplying the private key with a generator point **G** in the elliptic curve. Figure 2 shows the steps followed in the ECC secret key generation and implementation. An elliptic curve over a prime field is described as follows:

$$y^2 \bmod p = \left(x^3 + ax + b\right) \bmod p \tag{1}$$

Given the general vulnerabilities associated with spatial domain data hiding schemes, the LSB insertion method is modified to further harden the security of the secret message. One key problem inherent in many LSB-based data hiding scheme is non-accurate detection vulnerability which oftentimes results in false positive or negative RS analysis results. Estimation and detection of secret message become easy in such circumstances. This is because, in LSB a concept of raster order where the scanning of host pixels to hide data is executed row by row and scanned from the left-most to the right-most in a zigzag manner. Prediction and estimation become much easier.

The technique proposed improves the pixel substitution method. Let's assume ***k-rightmost*** LSBs of the image **X** are to be hidden by the **n-bit** secret message of **S.** This means that a conceptual virtual image of **k-bit** denoted as **S′** is formed by rearrangement of **S. S′** is therefore described as:

$$S' = \{s_i' S_i | 0 \le i < n', s\prime_i \in \{0, 1, 2, 3, ..., 2^k - 1\}\} \tag{2}$$

Equation 2 must satisfy a condition, thus **n < M * N**. A mapping relationship exists between **S′** and the **n-bit** secret message S. The mapping can be calculated using the formula described as:

$$s_i^{\iota} = \sum_{j=0}^{k-1} si \times k + j \times 2^{k-1-j}. \tag{3}$$

The process of embedding is completed after ***k*** LSBs of xl_i by ***S'$_i$*** is replaced. In this paper, the RS analysis step and procedure adopted by Wang et al. [15] is followed. RS steganalysis is defined over three kinds of block flipping. The block flipping are positive flippings ($\mathbf{F_1}$), negative flippings ($\mathbf{F_{-1}}$), and zero (0) flippings ($\mathbf{F_0}$). $\mathbf{F_1}$, $\mathbf{F_{-1}}$, and $\mathbf{F_0}$, become flipping functions and form what is termed flipped group. The flipped group results from applying the flipping functions on each divide image block pixel values. The flipped group is written as:

$$F(G) = (F_{M(1)(X1)}, F_{M(2)(X2)}, ..., F_{M(n)(Xn)}) \tag{4}$$

where **M = M (1), M (2), …, M (n)** represents the flipped mask, and **M (*i*)** has values indicating either **1, 0, or −1**. G is regular if $\boldsymbol{f}\ (\boldsymbol{G}) < \boldsymbol{f}\ (\boldsymbol{F}(\boldsymbol{G}))$ otherwise G is singular when $\boldsymbol{f}\ (\boldsymbol{G}) > \boldsymbol{f}\ (\boldsymbol{F}(\boldsymbol{G}))$**.** The implementation requires first dividing image into non-overlapping blocks and re-arranging each one of them into a vector

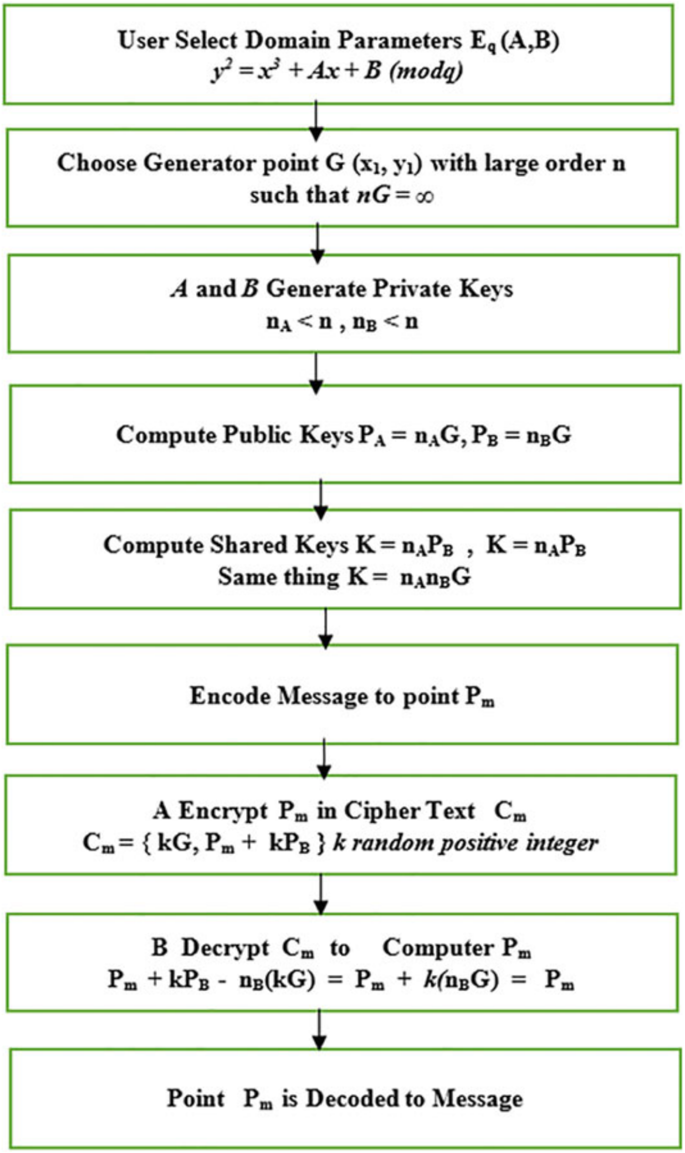


Fig. 2 ECC cryptographic procedure for implementation

$\mathbf{G = (x_1, x_2, x_3, \ldots x_n)}$. The blocks are arranged in a zigzag scan order. The discrimination function of the pixel's correlation is estimated by using the mathematical formula:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n-1} |x_i - x_{i+1}| \tag{5}$$

The pixel values are represented by $\boldsymbol{x}$ and $\boldsymbol{n}$ is used to represent the number of pixels. Also, $\boldsymbol{f}$ represents partial correlation between the adjacent pixels**.** Smaller $\boldsymbol{f}$ value means stronger the correlation exists between adjacent pixel values. The $\boldsymbol{f(F(G))}$ of each divide block is determined using Eq. (2). The relative numbers of each regular block and singular block after positive flipping are represented by $\boldsymbol{R_m}$ and $\boldsymbol{S_m}$ respectively, while the numbers of regular and singular blocks after negative flipping are represented by $\boldsymbol{R_{-m}}$ and $\boldsymbol{S_{-m}}$ respectively. The observed relationship between the block values can be defined as $\boldsymbol{R_m \approx R_{-m}}$, $\boldsymbol{S_m \approx S_{-m}}$, and $\boldsymbol{R_m > R_{-m}}$, $\boldsymbol{S_m > S_{-m}}$.

As indicated, the genetic algorithm is applied and implemented after LSB embedding to adjust the resulting stego-image pixel values by ensuring $\boldsymbol{R_m \approx S_m}$, $\boldsymbol{R_{-m} \approx S_{-m}}$. This paper adopts the GA to search for the best adjustment matrix, given that GA is optimization algorithm in general. The GA is implemented as a process of chromosome evolution. After several generations, the optimum solution is found when the best individual with high fitness level is selected. The most relevant processes of the GA are reproduction, crossover, and mutation. Basically, the individual within the given population which shows larger fitness values is determined to have higher possibilities of selection to breed the next generation. The GA steps implemented on the stego image after LSB embedding is demonstrated in Fig. 1.

4 Experimental Results and Discussion

The security of any genetic algorithm-based steganography technique including the proposed method is dependent on finding the best fittest chromosome for the pixel adjustment. The chromosome representation is therefore important. In this proposed method, the chromosome contains seven unique genes, represented by a total of 27 bits length. Table 1 gives further details. To obtain good fittest chromosomes, there is the need to select appropriate genetic parameters. The genetic algorithm (GA) parameters selected for the proposed system's experiment are given in Table 2. As already explained, the fitness of each chromosome is mathematically determined, and the chromosome that shows the fittest level is selected to replace the initial chromosomes and used for pixel adjustment.

A 256 × 256 Lena image (see Fig. 3a) converted to an 8-bit greyscale was used as cover image for the experiment. The embedding rates used for the experiment are 5, 10, 30, 40, 50, 70, and 90%. Figure 3b represents histogram generated for the original cover image, Fig. 3c is the generated histogram for stego image after embedding secret message using modified LSB without GA, and Fig. 3d represents GA pixel adjustment histogram for stego image. As the histograms can be observed, the visual quality differences between the three histograms can be seen as having

Table 1 Selection of appropriate representation of chromosome

Name of gene	Value range	Bits length	Descriptions
Direction	0–15	4 Bits	This represents the scan order direction for the host image pixel values. 16 different directions exist, hence the need for 4 bits to represent the directions
X-offsets	0–255	8 Bits	The starting point for x direction. This represents the offsets in rows, which is the coordinate starting point in the x direction. For an image of size 256 × 256, 2^8 values are available for x
Y-offsets	0–255	8 Bits	The starting point for the y direction. This represents the offsets in columns, which is the coordinate starting point in the y direction. For image of size 256 × 256, 2^8 values are available for y
Bit-Planes	0–15	4 Bits	This represents the least significant bits (LSBs) used for the secret bits embedding, which keeps records of LSB planes that have been utilized for message embedding
SB-poles	0 or 1	1 Bit	The pole of secret bits has two possibilities, either used in original form (bit 0) or used in flipped form (bit 1)
SB-dire	0 or 1	1 Bit	The secret bits direction is represented as 0, if the secret bit strings are used from left to right, else it is represented as 1
BP-dire	0 or 1	1 Bit	The direction bits plane demonstrates the order in which bits planes (BP) are filled, where 0 is used if BP is filled from most significant bit (MSB) to LSB, otherwise 1 is used for LSB to MSB filling

Table 2 Genetic algorithm parameters used for the experiment

Genetic algorithm parameters selected	Parameter values
Size of population	300
Rate of crossover	0.7
Rate of mutation	0.05
Replacement rate	0.8

different visual quality. However, there is no visual noticeable detection between the stego image with GA pixel value adjustment and GA pixel adjustment stego image. There is visual quality similarity between Fig. 3b and d, and the changes to the pixel values are not easily detectable by the human eye. The genetic algorithm therefore improved the security of the secret message and retained stego image visual quality.

The two most important factors in determining the performance of steganography methods are embedding payload capacity and stego image visual quality. In this study, different embedding ratios were experimented as given in Table 3, with 90% being the maximum embedding rate. Image steganography performance evaluation criteria were used to determine the visual quality of the stego image, the robustness, and imperceptibility measures. In this paper, the peak signal-to-noise ratio (PSNR) is used to measure the visual quality between the original cover image and the stego

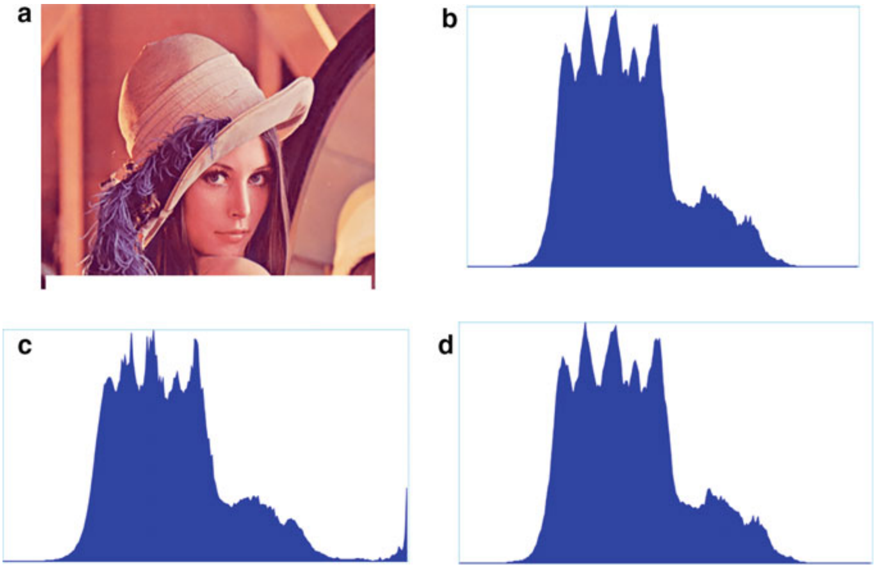


Fig. 3 **a** Lena cover image. **b** Histogram of cover image. **c** Hist. of initial stego image. **d** Hist. of stego image with GA

image. The mathematical equation for determining PSNR is written as:

$$\text{PSNR} = 20 \cdot \log_{10}(\text{MAX}_I) - 10 \cdot \log_{10}(\text{MSE}) \tag{6}$$

The maximum pixel value is represented by $\mathbf{MAX_I}$. The **MSE** is mean square error, which is a measure of noticeable distortion between the original cover image and the stego image. The mathematical equation for MSE is written as:

Table 3 Experimental results for various imperceptibility measures

Embedding rate (%)	BPP	MSE	PSNR	NC	Stego image RS analysis	
					$\lvert R_m - R_{-m} \rvert$	$\lvert S_m - S_{-m} \rvert$
5	0.5	0.0298	58.15	0.9999	*0.091*	*0.090*
10	0.70	0.0423	56.21	0.9998	0.093	0.091
30	1.51	0.0652	54.25	0.9996	0.095	0.093
40	1.89	0.0655	48.91	0.9989	0.099	0.096
50	2.25	0.0845	46.61	0.9988	0.124	0.111
70	3.01	0.0988	44.52	0.9985	0.151	0.140
90	3.39	0.0999	43.71	0.9983	0.272	0.245

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{7}$$

The height and width sizes of the image are represented by **M** and **N,** respectively. The Table 3 gives the values for PSNR and the MSE for the various embedding rates. The results show the method achieved high average values of PSNR and low values of MSE, an indication of the robustness and security of the technique proposed in this paper. The PSNR values were compared with few previous famous works, and the proposed method performed better than majority of the existing methods. The comparison of the proposed method with previous works is shown in Fig. 4. The proposed method is able to take as much as 3 bits per pixel (BPP). Another image quality measure was performed. The Pearson correlation coefficient (NC) is another good technique to check the imperceptibility between the original cover image and stego image visual quality. The NC is determined using the mathematical equation:

$$NC = \frac{\sum^{M}\sum^{N}\left(X_{MN} - \overline{X}\right)\left(Y_{MN} - \overline{Y}\right)}{\sqrt{\sum^{M}\sum^{N}\left(X_{MN} - \overline{X}\right)^2 \sum^{M}\sum^{N}\left(Y_{MN} - \overline{Y}\right)^2}} \tag{8}$$

where X is the cover image, Y is the stego image, $\dot{X}$ is the mean pixel intensity values for the cover image, and $\bar{Y}$ is the mean pixel intensity values for the stego image. The values obtained for NC at the various embedding rate and BPP are given in Table 3. Finally, the RS analysis is performed using the Eq. (5) shown above. The cover

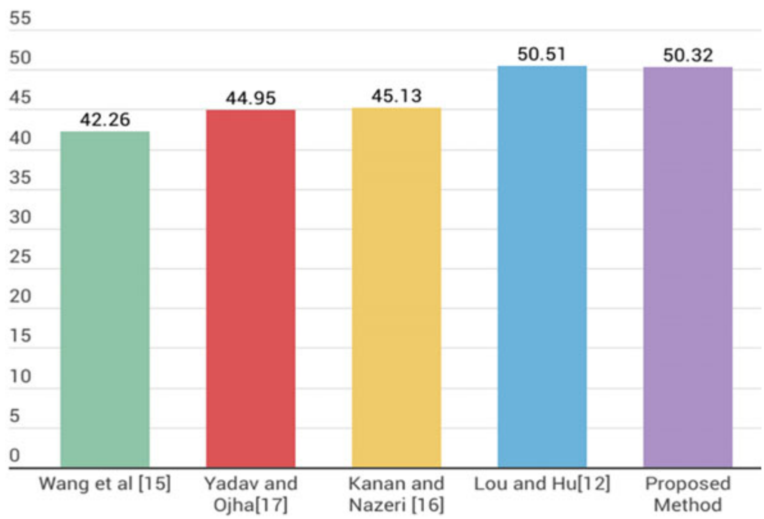


Fig. 4 Comparison of proposed method performance with previous works

image RS analysis $|R_m - R_{-m}|$ is 0.0078 and $|S_m - S_{-m}|$ is 0.0089, the stego image RS.

The results of analysis for $|R_m - R_{-m}$ and $|S_m - S_{-m}|$ are given in Table 3.

5 Conclusion

In this proposed method, during the implementation of the genetic algorithm, the blocks are labelled before adjustment is made. This was aimed at reducing the overheads of computational complexity and avoid the exhausting searching problem to make the algorithm easy to implement and more secured. The proposed technique presented in this paper has proved to be more robust to RS steganalytic attacks, improve payload capacity, as well as preserve and retain the visual quality to avoid data detection and modification as well as image distortion through statistical steganalysis. The average PSNR value of 50.3 achieved is higher than some famous previous studies. This paper has therefore provided the required foundation for the technique proposed to be further interrogated in an extended paper. The future work of this study will seek to improve the efficiency of the algorithm by enhancing the performance of the GA adjusted pixel value, increasing the population sample space and the bits per pixel (BPP) values.

References

1. Matted S, Shankar G, Jain BB (2021) Enhanced image security using stenography and cryptography. In: Smys S, Palanisamy R, Rocha Á, Beligiannis GN (eds) Computer networks and inventive communication technologies, vol 58. Lecture notes on data engineering and communications technologies. Springer, Singapore, pp 1171–1182
2. Pramanik S, Bandyopadhyay SK, Ghosh R (2020) Signature image hiding in color image using steganography and cryptography based on digital signature concepts. In: 2020 2nd international conference on innovative mechanisms for industry applications (ICIMIA). IEEE, Bangalore, India, pp 665–669
3. Rasras RJ, AlQadi ZA, Sara MRA (2019) A methodology based on steganography and cryptography to protect highly secure messages. Eng Technol Appl Sci Res 9(1):3681–3684
4. Kusuma EJ, Sari CA, Rachmawanto EH, Moses Setiadi DRI (2018) A combination of inverted LSB, RSA, and Arnold transformation to get secure and imperceptible image steganography. J ICT Res Appl 12(2):103–122
5. Ahmed DE, Khalifa OO (2014) Robust and secure image steganography based on elliptic curve cryptography. In: 2014 international conference on computer and communication engineering. IEEE, Kuala Lumpur, Malaysia, pp 288–291
6. Al-Juaid NA, Gutub AA, Khan EA (2018) Enhancing PC data security via combining RSA cryptography and video based steganography. J Inf Secur Cybercrimes Res 1(1):5–13
7. Mahto D, Yadav DK (2017) RSA and ECC: a comparative analysis. Int J Appl Eng Res 12(19):9053–9061
8. Poornima R, Iswarya RJ (2013) An overview of digital image steganography. Int J Comput Sci Eng Surv 4(1):23

9. Apau R, Hayfron-Acquah JB, Twum F (2016) Enhancing data security using video steganography, RSA and Huffman code algorithms with LSB insertion. Int J Comput Appl 143(4):28–36
10. Fatnassi A, Gharsellaoui H, Bouamama S (2016) A novel approach for embedding information in image cover media with security. In: 2016 world symposium on computer applications & research (WSCAR). IEEE, Cairo, Egypt, pp 104–109
11. Lou DC, Hu CH (2012) LSB steganographic method based on reversible histogram transformation function for resisting statistical steganalysis. Inf Sci 188:346–358
12. Baagyere EY, Agbedemnab PAN, Qin Z, Daabo MI, Qin Z (2020) A multi-layered data encryption and decryption scheme based on genetic algorithm and residual numbers. IEEE Access 8:100438–100447
13. Fridrich J, Goljan M, Du R (2001) Reliable detection of LSB steganography in color and grayscale images. In: Proceedings of the 2001 workshop on Multimedia and security: new challenges. ACM, pp 27–30
14. Wang S, Yang B, Niu X (2010) A secure steganography method based on genetic algorithm. J Inf Hiding Multim Signal Process 1(1):28–35
15. Kanan HR, Nazeri B (2014) A novel image steganography scheme with high embedding capacity and tunable visual image quality based on a genetic algorithm. Expert Syst Appl 41(14):6123–6130
16. Yadav GS, Mangal P, Parmar G, Soliya S (2023) Genetic algorithm and Hamilton path based data hiding scheme including embedding cost optimization. Multimedia Tools Appl 1–17
17. Yadav GS, Ojha A (2018) Hamiltonian path based image steganography scheme with improved imperceptibility and undetectability. Appl Soft Comput 73:497–507

Content Recommender System Based on Users Reviews

Puneshkumar U. Tembhare, Ritesh Hiware, Shrey Ojha, Abhisheik Nimpure, and Faiz raza

Abstract The usage of recommender systems is widespread across many areas and has been demonstrated to be extremely important in several fields. The majority of conventional recommender systems rely on a user's numerical rating of a consumed item to reflect that user's opinion; however, these ratings are not always available. As a result, to make up for the absence of these evaluations, a new source of information represented by user-generated reviews is added to the recommendation process. This saves the user time from searching the Internet for movies among the thousands that are already available. The items that might be suggested to the user are described by content-based recommendation systems. It makes predictions about what content a user will enjoy based on a dataset and takes into account the qualities of the content they have already liked. The sentiment analysis field can be used to gather rich and extensive information from reviews about the entire item or a specific aspect. This publication provides a thorough introduction to assist researchers who wish to work with sentiment analysis and recommender systems. It provides background information on recommender systems, including their phases, techniques, and performance measures. After that, it talks about the idea of sentiment analysis and highlights its key components, such as level, approaches, and aspect-based sentiment analysis.

Keywords Recommendation system · Collaborative filtering (CF) · Content-based filtering (CB) · Hybrid recommendation · Data mining · Recommender system (RS)

1 Introduction

The amount of information on the web has increased dramatically recently and is still growing, giving users and customers access to a variety of resources concerning services including goods, lodging, and dining options. Despite such data's benefits,

P. U. Tembhare · R. Hiware (✉) · S. Ojha · A. Nimpure · F. raza
Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India
e-mail: riteshhiware28@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_40

the vast flow of information causes challenges for users to deal with and choose from a vast number of available options. This results in a problem of information overload and makes decision-making more difficult. In order to help the user or client in this situation make the best choice, it is crucial to filter the information down to a manageable amount based on their current preferences. Recommender systems (RSs) intended to address the issue of information overload by making individualized service (i.e., item) suggestions to particular clients in accordance with their preferences, and are commonly used to carry out this filtering process. Since the field of RS first emerged more than a few decades ago, it has played a significant role in academics, business, and industry. This gave the RS a wide study scope and inspired additional research projects to identify useful methods for enhancing the RS's performance.

A unique kind of information filtering system are recommendation systems. These are computer programmers designed to assist users in making decisions as they engage with a wealth of information. They base their recommendations on the consumers' preferences, which might be communicated openly or implicitly. Such systems are becoming crucial tools for users in a number of applications, typically including information searching or e-commerce activities, because of the continuously growing volume and increasing complexity of information on the web. By exposing users to the most intriguing content and by providing innovation, insight, and relevance, recommendation systems aid users in overcoming the problem of information overload. Hence, recommendation technology plays a crucial role in the information seeking issue that has arisen alongside the World Wide Web.

There are three main recommendation approaches which are content-based, collaborative-based, and hybrid [4]. The content-based (CB) strategy finds the best recommendations for a user based on his most recent actions, such as what he recently liked, purchased, or watched [1]. Figure 1 illustrates the use of CB strategy.

The recommendation for a user is generated using the collaborative filtering approach (CF) based on the similarities among users who have previously expressed similar preferences or interests to him. Figure 2 illustrates the use of CF approach.

The last approach is the hybrid which integrates two or more recommendation components or algorithms implementations into a single recommendation system [5] (Fig. 3).

This provided the RS with a broad study scope and sparked other research efforts to find practical ways to improve the RS's performance.

2 Related Work

Over the years, several different recommendation systems have been created. Numerous approaches are used by these systems, such as collaborative, content-based, utility-based, hybrid, etc. The content-based (CB) strategy finds the best recommendations for a user based on his most recent actions, such as what he recently liked, purchased, or watched. The recommendation for a user is generated using the

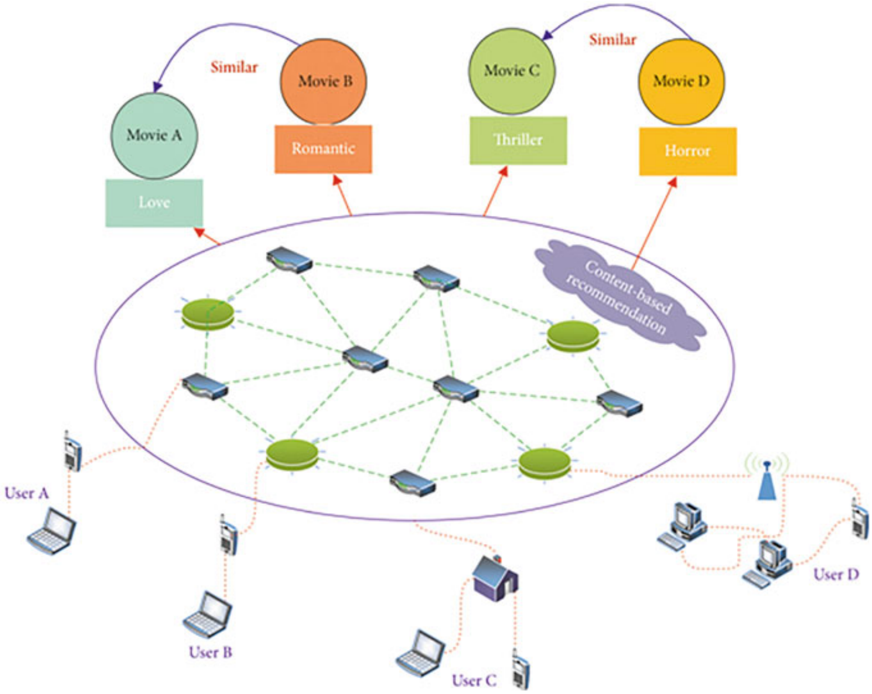


Fig. 1 Content-based filtering

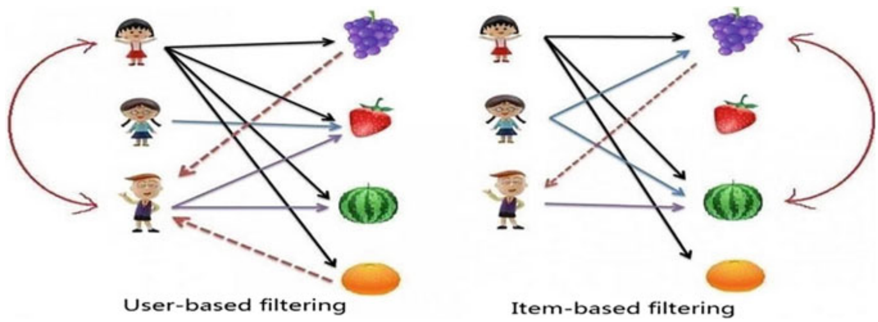


Fig. 2 Collaborative filtering

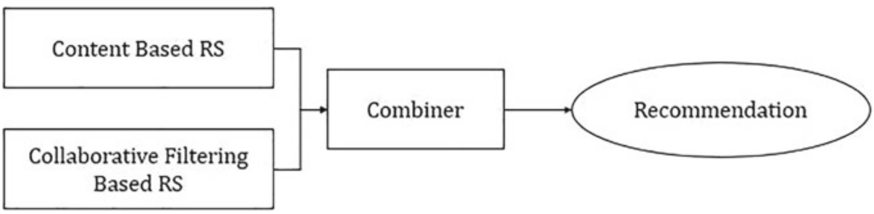


Fig. 3 Hybrid filtering

collaborative filtering approach (CF) based on the similarities among users who have previously expressed similar preferences or interests to him. The last strategy is a hybrid that combines two or more algorithms or recommendation component implementations into a single recommendation system. However, the major source for the recommendation process in most traditional RS techniques is a single-criterion rating (overall rating).

Yeole Madhavi et al. [1] use content-based recommendation using count vectorizer and cosine similarity and content-based recommendation using tfidf vectorizer and cosine similarity to get the most popular movies from both of the recommendations after receiving suggestions from algorithms 1 and 2, respectively, and add the remaining films in a different way later, after the common films.

Shruti et al. [2] in this, an algorithm known as the feature-based product ranking and recommendation algorithm (FBPRRA) has been proposed. It is developed for recommending products on the basis of an analysis report of consumer feedback. By taking into account information about customers who have purchased and to whom it was purchased, they make product recommendations based on online reviews.

Pradeep et al. [3] use a content-based recommendation approach. When making recommendations, the content-based recommendation method totally ignores other user profiles. Input from the user will be used to generate personalized suggestions.

3 Proposed Methodology

The proposed strategy combines memory-based and model-based algorithms with the content recommendation system. Analyses of the parameters and their results are carried out. This method looks up rating predictions using user history that is kept in a database. The time needed to find the rating predictions is lowered if additional algorithms are applied after utilizing the domain recommendation. As a result, users will receive predictions rapidly. These predictions take into account both the users' past behavior and similar users' viewpoints. The image and text information are used to match with the information of contents stored in databases to provide content recommendations that are taken directly from a product. There is a lot of information in the content. Here, a content recommendation system is implemented primarily using image and description data. To predict users' interest in information, products, and services among the amazing variety of options available, recommender systems use data mining techniques and prediction algorithms. The user prototypical, which contains details about personal preferences that govern an individual's behavior in a complex environment of web-based systems is the core element of all recommender systems.

3.1 Memory Based

Memory-based methods execute recommendations by directly interacting with the database, whereas model-based methods use transaction data to build models that can produce recommendations. Memory-based methods are adaptable to data changes since they access databases directly, but depending on the magnitude of the data, they take a long time to process. Regarding the model-based approach, it is not adaptable to changes in the data and has a constant computation time regardless of the size of the data.

3.2 Apriori Algorithm

Making accurate and effective recommender systems has been the subject of numerous algorithmic techniques. The first "recommender systems" were content filtering programs created to handle textual domains' information overload. These frequently used conventional information filtering and information retrieval systems as their foundation. A well-known technique for understanding association rules is the recommendation system employing apriori algorithms. Apriori is made to work with databases that contain transactions. For the purpose of identifying association rules in data lacking relationships, additional algorithms are calculated.

3.3 Content-Based Filtering

The content-based strategy finds the best recommendations for a user based on his most recent actions, such as what he recently liked, purchased, or watched. Based on the user's prior experiences, this filtering makes recommendations for products. For example, if a user exclusively enjoys action movies, the algorithm will recommend to him just more action movies that are comparable with those he has highly rated. User attributes are used to inform content-based filtering techniques.

3.4 Model-Based Method

In the model-based approach, a user model is created using the ratings of each user to assess the expected worth of unrated items. To create a model, this technique typically uses a data mining or machine learning algorithm. The utility matrix, which is constructed using user ratings for any item, is used in the model's development.

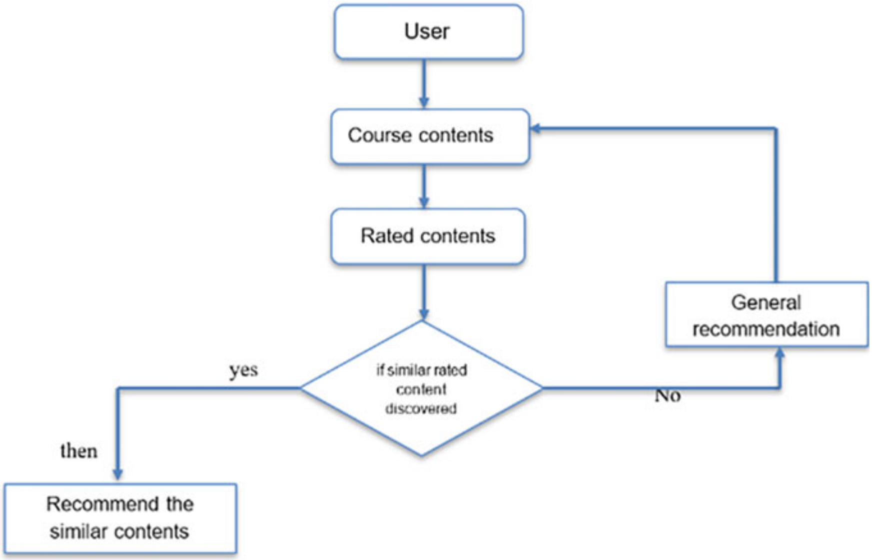


Fig. 4 Recommendation of contents

3.5 User-Oriented Reviews

In the suggested approach, we suggest materials based on user reviews of such products or items. They are fully aware of the contents' qualities. Figure 4 shows how the content is recommended to the particular user.

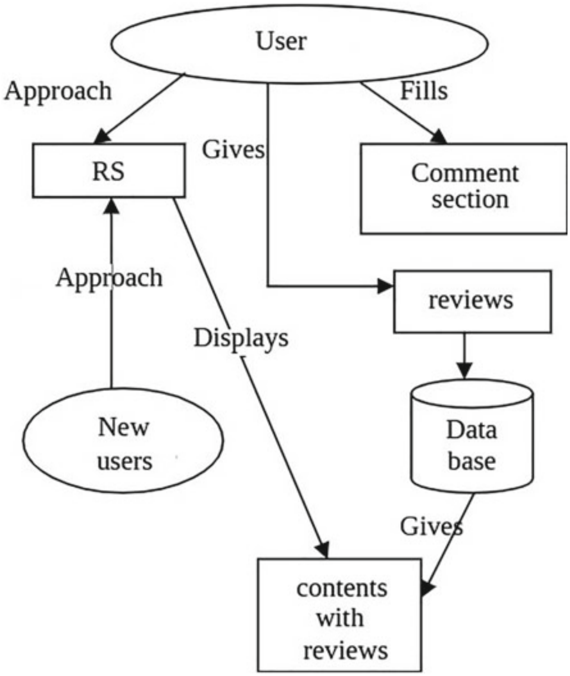
Users are selected in a way that ensures they have viewed the material. Long-being in the sense that they would only view a certain piece of content repeatedly if they were truly interested in it.

3.6 Recommendation Process

The recommender system compiles user reviews and provides them to other users looking for recommendations. The recommender system uses the following procedure to receive reviews and make recommendation.

1. User opinions are initially gathered. If a user visits RS frequently (let's say more than a certain number of times; this is depending on specific recommender systems), RS will choose their reviews.
2. These users visit RS and look over the contents. They were provided a comment section box to provide feedback after viewing the material. The purpose of this comment area is to learn how interesting the information is to other people.

Fig. 5 Users-oriented reviews recommendation



3. From the comment section, reviews can be viewed and it will helpful for the users.
4. Along with comments, the user's name is also gathered. Their personal profile is not stored.
5. When new users approach RS and search for contents, they were given the list of reviews about the contents (Fig. 5).

4 Results and Discussion

1. User opinions are initially gathered, Fig. 6 shows that user has commented on the movie KGF part 1 movie. It is stored in database. It has a comment section box to provide feedback after viewing the material. The purpose of this comment area is to learn how interesting the information is to other people.
2. Along with comments, the user's name is also gathered. In future, if movie with the same genre is released then the movie will be recommended to the particular user by this website. Figure 7 shows that the movie with the same genre is recommended successfully.

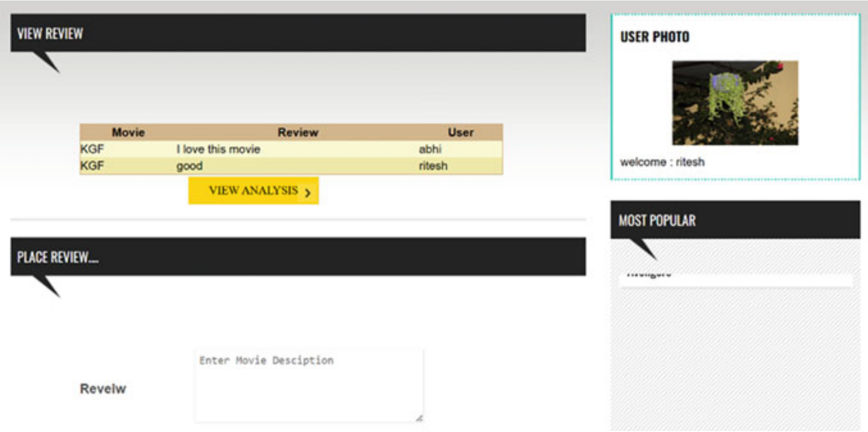


Fig. 6 User reviews the movie

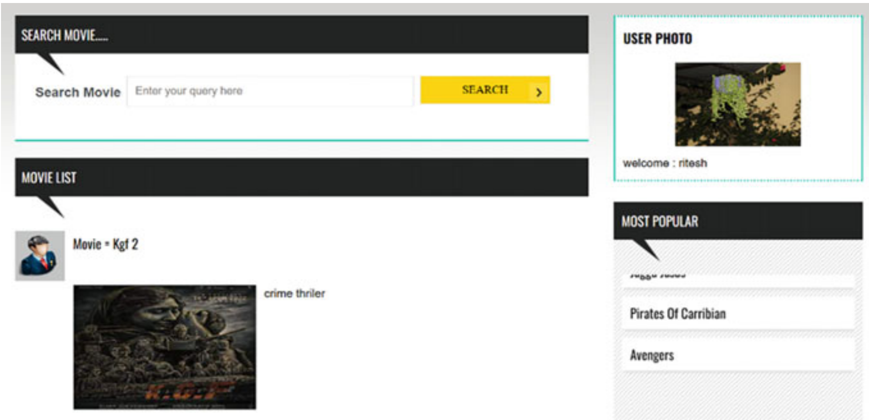


Fig. 7 Recommended the movie with the same genre

3. When new users approach website and search for contents, they were given the list of reviews about the contents. The website also has the graph which shows the analysis of the negative, positive, and neutral comment on the particular movie (Fig. 8).

4.1 User Information Table

This table contains data about users, primarily the user ID, user name, age, and occupation. The main key is the user ID. Figure 9 displays the specific layout of the table.

Fig. 8 Analysis of negative, positive, and neutral comment

Fig. 9 User information database table

Title	Data type	Notes
ID	Int (12)	User ID
NM	Varchar (60)	User name
PW	Varchar (60)	Password
AG	Int (12)	Age
GD	Bool	Gender
OP	Varchar (30)	Occupation

Fig. 10 Project information database table

Title	Data type	Notes
ID	Int (16)	Movie ID
TL	Varchar (255)	Movie title
RD	Date	Issue date
IU	Text	Viewing address
PC	Text	Picture path
TP	Text	Film genre
DT	Varchar (30)	Director
AT	Varchar (30)	Performer

4.2 Project Information Table

This table contains all pertinent information about the movie, primarily the title, genre, release date, and picture path. The main key is the movie ID. Figure 10 displays the specific layout of the table.

4.3 Scoring Table

This table keeps track of user scoring data for movies, primarily user ID, movie ID, score, and scoring time. The two main keys are user ID and movie ID. Figure 11 displays the specific layout of table.

Fig. 11 Rating information data sheet

Title	Data type	Notes
UID	Int (15)	User ID
MID	Int (15)	Movie ID
RT	Int (15)	Score
TM	Int (15)	Scoring time

5 Conclusion

Due to the overabundance of data today, information retrieval has become exceedingly challenging, making it difficult for users to acquire the information that most closely matches their tastes. The suggestion mechanism is used in this situation. For many users, it is useful to retrieve personalized information. The popularity of movies as a form of entertainment has increased, but so has the anxiety over what to watch. Other user profiles are not taken into account when creating recommendations using the content-based recommendation approach. The user will benefit from customized suggestions for their input thanks to this. The proposed method offers an effective means of locating the relevant rating predictions. In other words, it gives individuals with similar interests better recommendations. This method determines the ratings of the users' searched items. The history of the products is examined to see which ones are effective and popular. The product reviews are helpful during the analysis stage. Utilizing fewer recommendations maximizes how quickly ratings can be calculated while considering all recommendations. The analysis stage is made simple by it. Additionally, finding trustworthy suggestions goes more quickly.

References

1. Yeole Madhavi B, Rokade Monika D, Khatal Sunil S (2021) Movie recommendation system using content-based filtering, vol 7, issue 4
2. Shruti M, Gayatri P, Sojwala G, Prathamesh Y (2021) A review paper on product recommendation system using online reviews. vol 12, issue 1
3. Pradeep N, Rao Mangalore KK, Rajpal B, Prasad N, Shastri R (2020) Content based movie recommendation system. Int J Res Ind Eng 9(4):337–348
4. Bhonde MM, Sawarkar CH, Mulkalwar PN (2020) Data mining and recommender system: a review. JETIR 7(2)
5. Kaur H, Bathla G (2019) Techniques of recommender system, vol 8, issue 9S. ISSN: 2278-3075
6. Tembhare PU, Balpande S, Bargade J, Prasad M, Kumari S, Lanjewar V (2016) Review analyzer analyzing consumer product, vol 5, issue 3, pp 693–698
7. Patel A, Thakkar A, Bhatt N, Prajapati P (2019) Survey and evolution study focusing comparative analysis and future research direction in the field of recommendation system specific to collaborative filtering approach. Springer, Singapore, pp 155–163

Devanagari Handwritten Character Recognition Using Deep Learning and Neural Network

Dhanraj Jadhav, Sanket Ujgare, Vaibhav Chaudhari, Saurabh Dhavale, and Rohit Dadhel

Abstract Manually processing handwritten documents is a very difficult task due to different people having different handwriting styles and a lack of currently present language dictionaries to interpret documents. There are millions of people who are disconnected from fundamental technologies due to the language gap. The majority of people don't know English. Recognizing character becomes more important for automation in various fields. The paper proposes a methodology for the recognition of handwritten English and Hindi characters. In this paper, for character recognition, we apply deep learning technology with convolutional neural network (CNN).

Keywords Convolutional neural network · Keras · Tensorflow · Softmax · ReLU · XGBoost · Support vector machine (SVM) · Deep learning · Machine learning · Handwritten character recognition

1 Introduction

Aim of this project should be an implement system that will be helpful in recognizing the characters in pictures, texts, documents and different sources of English language and translate them in machine understandable and readable form. It can further develop to recognize characters of different language. The handwritten character recognition is very useful in various sectors like banking and healthcare for making automation. Accurate recognition of handwriting remains a major problem, the trending researches in CNN have major impact on character identification by investigating discriminating characteristics from massive volumes of data. The HCR is widely used in "computer vision" and "pattern recognition", it is one of the part of this technique. The CNN automatically recognizes and extracts the unique and differentiating features from the images of samples input. CNNs work better on bigger sample datasets. HCR has many applications such as checking official documents, checking bank cheques and helping the blind persons, reading bills or street

D. Jadhav (✉) · S. Ujgare · V. Chaudhari · S. Dhavale · R. Dadhel
JSPM's Rajarshi Shahu College of Engineering, Pune, India
e-mail: dhanraj.acem12@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_41

signs. Although various application is available but there are still many challenges are present which is different writers have different handwriting style and quality of image in handwritten character recognition.

2 Preparation of Paper

2.1 Related Works

In 2022, Hirugade and his team [1] proposed a model for recognizing a handwritten Devanagari character. The aim of this paper to study different techniques is used in this field. The CNN model with use of dropout layers will increase performance and accuracy of model, it detects the important features.

In 2021, Villupuram [2] proposed an approach of recurrent neural network that is used with OCR and it uses bottom-up approach, using OCR it effectively identifies and recognizes the characters. RNN comprehended the challenge with the assistance of hidden layer.

In 2021, Tabassum [3] in this stroke rotation and parallel shift with bidirectional LSTM model is used, when the samples of characters are taken from different datasets the SRP method is used which is best technique. If these datasets are expanded it can be used for data expansion, for this having various styles of handwriting will increase the samples of data. It develops corpus dataset and, in this paper, system is optimized by reduced errors.

In 2021, Heo [4] tried to explore factors necessary for perception performance of handwritten English and Korean character recognition through OCR, and this paper tried to improve the cursive recognition rate using deep learning technology, the steps included in this paper are input image information, segmentation of background and text area using ANN, separate the initial, median and final areas using CNN and then recognition performance is evaluated.

In 2021, Narayan [5] used a Chars74K dataset of English language, further data is split into 80:20 ratio for testing and training and validation. Also, the image preprocessing techniques are used for distortion which improves the unique features of images and allows efficient performance on model. They used CNN model created using layers like convolutional, pooling, dense or fully connected layer to identify characters using techniques like contour recognition. They try to increase the performance of model in future.

In 2021, Alrobah and Albhali [6] proposed a model using CNN for both classification and feature extraction of Arabic characters, the accuracy is about 97%for AHCD dataset. They identified that the Hijja dataset contains irregularities, some blurry letters and distorted unclear symbols and they use a hybrid model of machine learning and deep learning.

In 2020, Gupta and his team [7] proposed a model for recognizing the characters in Devanagari with the help of CNN with accuracy of 95.6%, they trained CNN with 378,951 images and optimized using three epochs.

In 2020, Yadav and his team [8] proposed a model for "Devanagari character recognition" by using Shirorekha dataset, they use CNN architecture with consecutive convolutional layers for extracting high-level features. The model has 99.65% accuracy which is very high for character recognition.

2.2 *Major Techniques Referenced in the Literature Survey*

2.2.1 Convolutional Neural Network (CNN)

CNN is the part of deep learning technology, CNN is the algorithm of deep learning which accepts and does a processing on inputs and is important for many factors, and can classify and identify images through feature extraction. The smallest and important components of a CNN are neurons. Convnet structures exist in neurons present in the human brain and are facilitated by neuronal structures dictated by structures in the visual cortex. Convolution is one of the important operations in CNN, using CNN we can extract the detail info from images, a matrix traverses through the screen over a pixel and feature is computed from values of every window. ConvNet architectures are same as neural network which can efficiently classify the data, it is very useful in classification and recognition of image technique.

1. **Convolutional layers** Convolutional layers are the main building blocks of CNNs, where most of the computations are performed. It requires several components: inputs, filters and feature maps. Suppose we have taken an input as an image which has colour and contains a matrix of 3D pixels. This 3D matrix contains information about three dimensions such as height, depth, width with respect to RGB colour of image. It also has feature detectors, called kernels or filters, that traverses through the respective field of the images that identifies unique feature and performs operation, this method is called convolution.
2. **Pooling layer** Pooling layer is similar to convolutional layer. In pooling layer, it reduces the feature of the input, the difference in this is that filter does not have weights. In pooling layer, we are taking the input and reducing features present in samples of data, using dimensionality reduction. The kernel in the pooling layer uses aggregate function on the value in which we get in the output array with the data in acceptor field. Pooling is of two types mainly.
3. **Max pooling** In max pooling, there is filter which traverses through a sample, we select highest value of pixel which is stored for an output of same members, this max pooling method is mainly used for intermediate combinations.
4. **Average pooling** In this method of average pooling, the filter which traverses across the samples of data, and for receptive area of image we get the output in the form of array by calculating the average value of input.

5. **Fully Connected layer** Fully connected layer uses a SoftMax activation function where inputs are classified correctly by giving probability between 0 and 1, similarly ReLu activation function is used in the convolutional and pooling layers. In other layers, the output layer is not connected directly with an input of image but, every node in the output layer is connected to a node of previous layer in the fully connected layer. The various filters and extracted features are supported by the layers which implement classification work. The fully connected layer is layer as name suggests in this layer that all the nodes are connected with all other nodes, this layer performs a very important role in the CNN.

2.2.2 CNN Algorithms

1. **LetNet** This LetNet algorithm is very simple, small and easy to use. At beginner level, we use this algorithm. The fundamentals of CNN get clear by studying and understanding this LetNet algorithm. In this algorithm, there is convolutional neural network with one layer consisting set of two layers, average pooling and convolutional, and further convolutional layer has got flattened. The last layer is dense layer in which the softmax classifier and layers of two fully connected are utilized.
2. **AlexNet** AlexNet is one of the algorithms of CNN, in deep learning (DL) and machine learning technology, AlexNet has done some contributions. There are eight total layers in the AlexNet, the starting five layers are layers of fully connected. In AlexNet, the ReLu activation function is used which gives better result and accuracy than other activation function like sigmoid, tan.
3. **VGGNet** VGGNet is used for large project, so it is known as CNN with very deep, the image classification and their identification is performed by VGGNet in bigger scale. VGGNet is a network direct by stacking 3*3 convolutional layers to deepen the model, and the volume reduction with increasing layers is covered by the largest pooling layer. Finally after this in the model, softmax classifier is implemented in dense layer. There are two strongly connected layers. The VGG model has its own drawbacks, it is very slow to train and slows down.

2.3 CNN Architecture

See Fig. 1.

2.4 System Architecture of HCR

See Fig. 2.

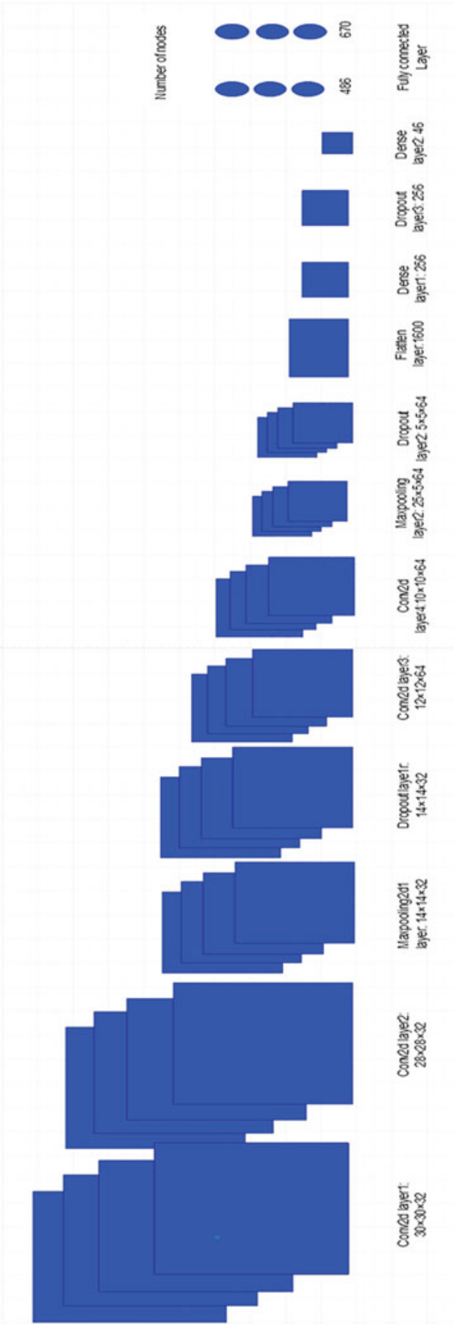


Fig. 1 CNN architecture of system methodology

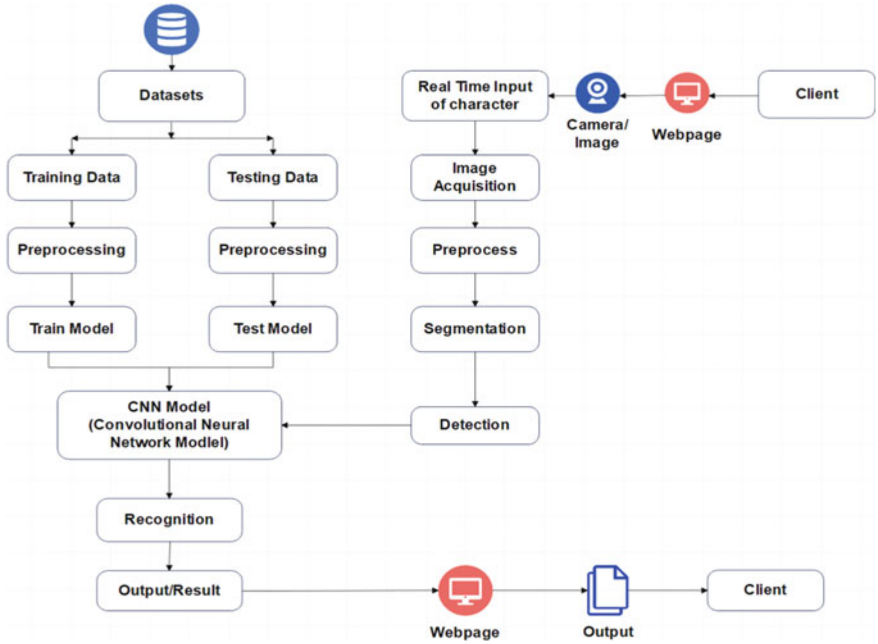


Fig. 2 System architecture of proposed methodology

2.5 *Summarization of Literature Survey*

See Table 1.

2.6 *Observation on Literature Survey*

Many activities and a large amount of work is undertaken in handwritten character recognition using the technology like optical character recognition, neural network and machine learning. The handwritten character recognition should be classified in two types online and offline, using techniques like image processing, feature extraction and image normalization for recognizing the characters. For handwritten character recognition many algorithms and different techniques are used giving good accuracy, using convolutional neural network and deep learning we get higher accuracy and accurate result but there are also various challenges in handwritten character recognition because of variety of handwriting style, as every person has its own handwriting style. After understanding and studying all the techniques and methods used for handwritten character recognition from all previous research papers, we understand the CNN as this is a technique of deep learning which is best for our system, it gives a best performance with a great accuracy.

Table 1 Summarization of Literature Review

Year and method	Title	Summary
2022 Survey paper	"Survey on optical character recognition for handwritten Devanagari script using deep learning"	In this paper, the author has done research of various papers which works in the area of recognition of character in Devanagari script using deep learning models. As the accuracy of CNN models depend on the number of layers used in architecture, author observed that use of the dropout layers helps to prevent overfitting, and it improves accuracy and performance
2021 Convolutional neural network (CNN)	"Optical character recognition for English handwritten text using recurrent neural network"	This "network of recurrent neural" uses OCR techniques that use a bottom-up approach. OCR effectively distinguishes characters. RNN used hidden layer to understand the problem (no need to remember past words), having accuracy 90%
2021 SRP with LSTM	"Recognition of doctors cursive handwritten medical words by using bidirectional LSTM and SRP data augmentation"	This SRP (using stroke rotation and parallel transformations with bidirectional LSTM models) allows SRP methods to perform for different samples of input when we get the input of character for recognition. Extending these datasets can be used for data expansion, and various writers having different writing styles will increase and expand a dataset's capacity size. It develops a handwritten corpus dataset of medical terms which minimizes the defects and errors, with 89.5% accuracy
2021 CNN with OCR	"Comparative study on the perception performance of handwriting in Korean and English using machine learning"	In this paper, system is built using CNN and ANN which is important in the OCR use, it improves the cursive recognition rate by using optical character recognition and conversion of neural network algorithm but CNN requires a large dataset for training. Cursive writing is difficult to understand so developed system is very useful for understanding the script, with 98% accuracy
2021 CNN	"Image character recognition using convolutional neural network"	In this work, there is research done for building system using CNNs for more accurate recognition and recognition of handwritten script in the pictures. There is CNN system which is implemented and tested with English handwriting and performance verified. The system is developed to recognize character using extraction of features from the input of picture across different levels and then system is trained for recognition, with 97.59% accuracy

(continued)

Table 1 (continued)

Year and method	Title	Summary
2021 CNN	"Hybrid deep model for recognizing Arabic handwritten characters"	This paper contains a work of SVM, CNN and XGBoost detectors are used. It detects malformations such as some blur letters and other distorted and indistinct signals and diminishing. MLP is also used to reduce visual risks that reduces errors in the training set and increases area of margin for training samples of classes having 96.30% accuracy
2020 CNN	"Convolutional neural network-based handwritten Devanagari character recognition"	In this paper, process of character recognition in which how actually character is recognized in online and offline mode is given. The scanner is used to detect and recognize a character in offline method. To identify the character, identifying feature is very important task. In online method of character recognition, the touchpad or stylus is used to get pixel values based on its movement to recognize the character. The CNN works as a regular neural network by using methods such as back propagation or gradient descent. In this, tensorflow and library of keras are used in the backend for implementation of CNN networks, with 95.6% accuracy
2020 CNN	"Devanagari handwritten character recognition using convolutional networks"	In this paper, proposed system for recognizing a character uses a technique like image processing and deep learning, the dataset available for the author is with shirorekha, so author created his own dataset without shirorekha (Shirorekha is the horizontal line in the Hindi Character). He trained his model with approximately 1000 images for each character, and that's why it gives 99.65% of accuracy. Algorithm used in this paper consists of steps, converts image to greyscale and then into binary, edge detection using canny edge operator, removing noise from image, feature extraction is done by neural networks using ReLu activation function, maxpooling layer pools the feature, then 3d matrix is flattened and then softmax calculates class probabilities and Argmax outputs character, with 99.65% accuracy

2.7 Challenges/Issues

The handwritten character recognition is present in the market from 70s but there are still many challenges present in this field, if the quality of image is not clear and blurry and low-quality image is present. Different writers have different handwriting style, this will also create problem, the content in the background will call a noise which reduced the accuracy, if all the issues are removed then model gives high accuracy and good performance. We reduced this issue using black and white image with good quality and keeping data in margin will increase overall performance of model.

2.8 Conclusion

In India, many documents and research related work are present in Devanagari form to save this work and make a proper use of them, if a handwritten document is converted into digital format it saves the all-knowledge present in that book. Using handwritten character recognition, we convert a handwritten document into digital format, we are able to translate the content, which helps people to understand other language. In this paper, we study and analyze many papers of traditional neural networks for recognition character which is handwritten. Handwriting recognition is a process of recognizing characters from images, text and other formats, converting them into machine-readable formats, and further processing them. A category of "neural network" also called as conventional neural network. We study many neural network models for character recognition, the use of hybrid model which is a neural network model like CNN with techniques like SVM, LSTM and SRP gives a better accuracy and high performance.

2.9 Future Scope

Further this system will develop with more advanced datasets to get more accurate result, and this developed system will be available as an API for others usage. This system will develop so that it will be available on both web and smartphone devices.

References

1. Hirugade P, Suryavanshi N, Bhagwat R, Rajput S, Phadke R (2022) A survey on optical character recognition for handwritten Devanagari script using deep learning. In: Proceedings of the international conference on innovative computing & communication (ICICC), Available at SSRN: https://ssrn.com/abstract=4031738 or https://doi.org/10.2139/ssrn.4031738
2. Parthiban R, Ezhilarasi R, Saravanan D (2020) Optical character recognition for english handwritten text using recurrent neural network. In: 2020 international conference on system, computation, automation and networking (ICSCAN), Pondicherry, India, pp 1–5. https://doi.org/10.1109/ICSCAN49426.2020.9262379
3. Tabassum S et al (2021) Recognition of doctors' cursive handwritten medical words by using bidirectional LSTM and SRP data augmentation. In: 2021 IEEE technology and engineering

management conference—Europe (TEMSCON-EUR), Dubrovnik, Croatia, pp 1–6. https://doi.org/10.1109/TEMSCON-EUR52034.2021.9488622

4. Heo JH, Lee SW, Lee HW (2021) A comparative study on the perception performance of handwriting in Korean and English using machine learning. In: 2021 21st ACIS international winter conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD-Winter), Ho Chi Minh City, Vietnam, pp 274–275. https://doi.org/10.1109/SNPDWinter52325.2021.00070
5. Narayan A, Muthalagu R (2021) Image character recognition using convolutional neural networks. In: 2021 seventh international conference on bio signals, images, and instrumentation (ICBSII), Chennai, India, pp 1–5. https://doi.org/10.1109/ICBSII51839.2021.9445136
6. Alrobah N, Albahli S (2021) A hybrid deep model for recognizing Arabic handwritten characters. IEEE Access 9:87058–87069. https://doi.org/10.1109/ACCESS.2021.3087647
7. Gupta P, Deshmukh S, Pandey S, Tonge K, Urkunde V, Kide S (2020) Convolutional neural network based handwritten Devanagari character recognition. In: 2020 international conference on smart technologies in computing, electrical and electronics (ICSTCEE), Bengaluru, India, pp 322–326. https://doi.org/10.1109/ICSTCEE49637.2020.9277222
8. Gurav Y, Bhagat P, Jadhav R, Sinha S (2020) Devanagari handwritten character recognition using convolutional neural networks. In: 2020 international conference on electrical, communication, and computer engineering (ICECCE), Istanbul, Turkey, pp 1–6. https://doi.org/10.1109/ICECCE49384.2020.9179193

Measuring the Behavioral Intention Toward the Implementation of Super Artificial Intelligence (Super-AI) in Healthcare Sector: An Empirical Analysis with Structural Equation Modeling (SEM)

Soumya Kanti Dhara, Arunangshu Giri, Adrinil Santra, and Dipanwita Chakrabarty

Abstract The rationale of the study is to measure the behavioral intention toward the implementation of super artificial intelligence in healthcare sector—a model designed with the conglomeration of 'technology acceptance model (TAM)' and 'theory of planned behavior (TPB)' with their respective theoretical constructs. The authors intend to examine the effect of the constructs viz. 'perceived usefulness' and 'perceived ease of use' on 'attitude toward use' which further influences 'behavioral intention toward use'. The study also examines the influence of 'subjective norms' on 'behavioral intention toward use' and the effect of moderating construct 'perceived behavioral control' to gauge the influence of 'attitude toward use' on 'behavioral intention'. To test the proposed hypothesis, data were collected through structured questionnaire from 285 respondents. Exploratory factor analysis and structural equation modeling using AMOS were conducted for the establishment of the hypothesized research model. Moderation analysis was conducted through SPSS process macros. The result proposes the positive significant influence of all the factors on behavioral intention toward use with the exception of subjective norms that have negative influence on behavioral intention. The moderation effect of perceived behavioral control stated in this research is also justified. This paper contributes to test the behavioral intention toward the implementation of Super AI, an emerging field of computer science and engineering in healthcare sector and no such existing literature of previous research conducted has been found to provide any secondary data. Hence, the analysis is based on primary data collected, which justifies originality of the study.

Keywords Super artificial intelligence · ICT · TAM · TPB · Healthcare sector

S. K. Dhara · A. Giri (✉) · A. Santra · D. Chakrabarty
Haldia Institute of Technology, Haldia, West Bengal 721657, India
e-mail: arunangshu.giri@hithaldia.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_42

1 Introduction

The growing population in the world has been posing a grueling challenge in healthcare services. To tackle such challenges, information and communication technology (ICT) played a significant role in enhancing the capacity of healthcare sector. Nevertheless, with increased efforts, ICT has been attempting to play a crucial role in specialized healthcare service delivery to every remote corner of the globe. Technologies that are gearing up the ICT industry are machine learning, natural language processing, deep learning, artificial intelligence, etc. Super artificial intelligence (Super AI) is a more advanced, powerful and intelligent type of artificial intelligence which can surpass human intelligence by manifesting cognitive skills and developing thinking skills of its own. Super AI transcends the intelligence of some of the extraordinary minds of the people. The thinking ability of human brain is limited to a set of a few billion neurons. Machines with Super AI are self-conscious and possess the ability to think about abstract patterns, complex relationships and meaningful interpretations that are beyond the capacity of human beings. Apart from replicating multidimensional behavioral intelligence of human being, Super AI can also comprehend and interpret human emotions and experiences. Super AI currently, is a theoretical concept rather than a practical reality. The study investigates the behavioral intention toward the practical implication of Super AI in healthcare sector. The research model uses the combination of 'TAM' and 'TPB' to explore the study.

2 Background of the Study

2.1 Super Artificial Intelligence

Super intelligence has been referred as an intellectual ability that poses very highly smarter compared with the exceptionally human brains in every practical domain that includes scientific innovation, general wisdom and social skills [1]. The likelihood of super artificial intelligence superseding the current human intelligence can be understood from several prevailing aspects. Super AI may be thought to deliver a general reasoning system that can overcome the cognitive barriers of human beings. AI is a field of computer science and engineering that covers sub-fields such as machine learning, machine perception, natural language processing, neural networks, knowledge engineering, robotics, etc. that enables the machines to learn, plan, recognize and solve. Super AI goes beyond the limitations of general AI. Apart from this, Super AI can develop algorithms by itself with little or no human intervention to solve complex problems.

2.2 TAM and TPB

Technology acceptance model (TAM) proposed by Davis in 1989 is employed to study one's behavioral intention from the attitude toward use **(ATU)**, determined by two factors—perceived ease of use **(PEU)** and perceived usefulness **(PU)** of the individual [2]. The model is used to study the effectiveness of Super AI on healthcare sector in conjunction with the TPB model. Theory of planned behavior (TPB) proposed by Ajzen in 1985 suggests that one's behavioral intention to use **(BIU)** is influenced by the individual's attitude toward a behavior and subjective norms (**SN**), moderated by another factor—perceived behavioral control (**PBC**) [3].

3 Literature Review

A study was conducted on TPB model to investigate the influence of attitude components (affective and instrumental), subjective norm (SM) components (injunctive and descriptive) and perceived behavioral control components (self-efficacy and controllability) on the intention to exercise among undergraduates and cancer survivors. The study also aimed to examine whether any general common factor or any specific factor within the subcomponents does influence the intention and behavior. The findings revealed in the conceptual models developed that injunctive and descriptive components of subjective norm best predict intention as a general factor, and self-efficacy component alone of perceived behavioral control predicts exercise intention for both undergraduate students and cancer survivors. In contrast, the findings of the optimal models developed revealed that instrumental and affective components of attitude best predict exercise intention with a general factor among the undergraduates but only affective component of attitude as a specific factor to influence the same among cancer survivors. The three independent components of SN considered in this study are (i) healthcare professionals, (ii) patients and (iii) family and friends. The components considered for BIU in our study are (i) usage frequency, (ii) passionateness and (iii) multi-domain activity [4]. From the meta-analysis conducted on 26 selected empirical research articles based on TAM to amalgamate the empirical evidence, the findings propose strong correlations between (1) 'PU' and technology acceptance and (2) 'PU' and 'PEU' but weak correlation between 'PEU' and technology acceptance [5, 6]. The study based on TPB model investigated 'descriptive role' and 'social role' under 'SN' toward the purchase intention of green food by the household primary shoppers from a transitional country in the Southeast Europe region, which depicted a significant positive correlation between purchase intention with the three predictors, i.e., personal attitude, 'PBC' and 'SN' [5]. The quantitative survey validated and ranked 10 perceived benefits of hospital information systems (HIS) and electronic medical records (EMR) to different categories of healthcare professionals, which are (i) information access improvement, (ii) increased productivity of healthcare professionals, (iii) improvement in efficiency and accuracy of coding and billing, (iv) healthcare quality improvement, (v) improvement in clinical diagnosis and treatment, (vi) cost curtailment in paper-based medical records, (vii) minimal medical errors, (viii) improvement in patient safety, (ix) improvement in patient outcomes and (x) patient satisfaction improvement. Four vital components of PU that have been considered for measuring the behavioral intention of super AI in healthcare sector in our study are (i) improved efficiency of healthcare professionals, (ii) improved clinical management, (iii) cost reduction and (iv) improved patient satisfaction [7]. A study was conducted on TAM-based model to explore the effect of information technology in multimedia usage by the school teachers on their perceived ease of use, 'PU' and attitudes on behavioral intention. The study revealed that the behavioral intention was influenced significantly by 'PU' and 'usage attitude'. The 'usage attitude' was further influenced by 'PEU' and 'PU'. The 'PEU' also impacted 'PU'. Four components that have been considered for PEU in our study are (i) comprehensibility, (ii) proficiency development, (iii) applicability and (iv) flexibility. Three components that have been derived for ATU in our study are (i) acceptability, (ii) valuableness and (iii) affordability [8]. The empirical study examined the diffusion pattern of three antecedent technologies, namely robotics, automation and artificial intelligence in manufacturing, communication and energy sectors to address both evolutionary and control mechanism in managing artificial super intelligence. The research outcomes revealed that the diffusion model best fits with the robotics, telecom and solar installation data which suggest the diffusion process to be gradual and any future threat could be predominated by some social strategies [4, 9]. From the above literature reviews based on TAM and TPB models, the following hypotheses can be framed:

H_1**:** **'PU'** has a positive significant influence on the **'ATU'**.
H_2**:** **'PEU'** has a positive significant influence on the **'ATU'**.
H_3**:** **'ATU'** has a positive significant influence on **'BIU'**.
H_4**:** **'SM'** has a positive significant influence on **'BIU'**.
H_5**:** **'PBC'** is acting as a moderating variable to alter the degree of positive influence on the relationship between **'ATU'** and **'BIU'**.

4 Research Methodology

In this study, a framework of cross-sectional descriptive research design has been used for collecting and analyzing the data [10, 11]. The methodology used in this research is detailed in the following sections.

4.1 Sample and Survey Instrument

After developing the research model (Fig. 1) and proposed hypotheses, a structured questionnaire was constructed on the basis of research objective. The target population for survey was the employees of healthcare sector having the knowledge on artificial super intelligence. A total of 300 survey questionnaires were circulated to the respondents and 285 finished responses were received. Table 1 depicts the demography under different categories of the sample.

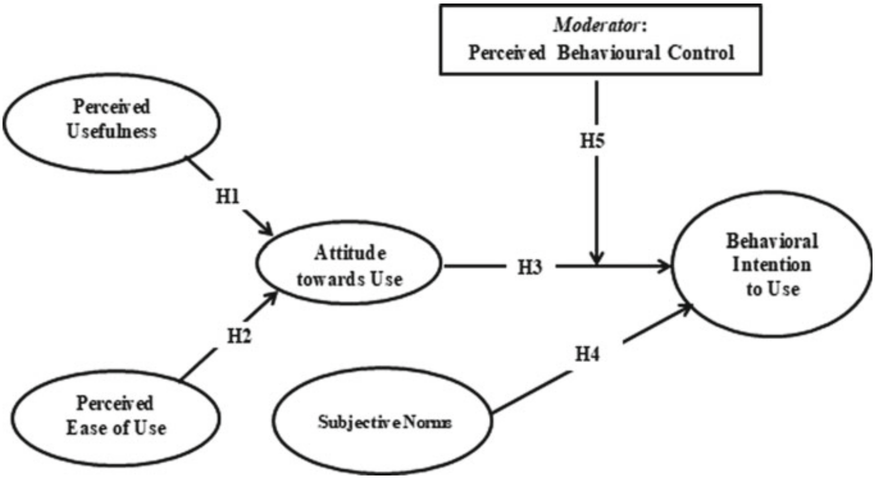


Fig. 1 Research model and proposed hypotheses

Table 1 Demographic details of the sample survey

Category	Items	No. of respondents	Percentage (%)
Sex	Male	161	56.49
	Female	124	43.51
Age	Less than 25 years	43	15.09
	26–35 years	105	36.84
	36–45 years	68	23.86
	46–55 years	51	17.89
	More than 55 years	18	6.32
Education	Graduate	83	29.12
	Post graduate	131	45.96
	Doctors	71	24.91

4.2 Sampling Technique, Survey Method and Measures

Data was collected from different healthcare centers in India. The questionnaire was administered through online and offline modes. Convenience sampling has been used for the current study [12]. The survey was conducted for a period of three months (July 2022–September 2022). Five study measures (factors) and also associated measurement items that have been considered in the research model were drawn from the existing literature. A total of 17 variables were identified in this case. Here, in the first stage face and content validity of the questionnaire were executed. Also reliability was checked.

5 Data Analysis and Results

The two statistical tools, SPSS 28.0 and AMOS 28.0 were used for conducting data analysis. Confirmatory factor analysis (CFA) was performed to test the concerned measurement model and structural equation modeling (SEM) was used to check the effect of the independent variables on dependent variables.

5.1 Common-Method-Bias (CMB)

Harman's one factor test was used to examine the CMB [13]. The maximum variance that is described by the single factor was 29.16% (< 50%), representing that CMB was not a potential threat.

5.2 Measurement Model: Examining Reliability and Validity

Goodness-of-fits (GOF) for measurement model were checked and found suitable which indicated the evidence of construct validity [14]. The standardized factor loading for individual variable should be more than 0.70 [15]. It was observed that all factor loadings are more than 0.70 and the total variation explained was 84.66%. Also, five factors from 17 variables were extracted as given below (Table 2).

To test the internal consistency of the dataset, reliability test was executed at factor level. The Cronbach's alpha coefficients for all the factors were more than the threshold value of 0.70 [16]. Here, data validity was measured by using construct validity. Here, construct reliability (CR) for the individual construct was more than 0.70 and average variance extracted (AVE) values were more than 0.50. It was observed that CR values were higher than AVE values. Also it has been found that maximum shared variance (MSV) < AVE, average shared variance (ASV) < AVE

Table 2 Exploratory factor analysis

Factors ←	Perceived usefulness (PU)	Perceived ease of use (PEU)	Attitude toward use (ATU)	Subjective norms (SN)	Behavioral intention to use (BIU)
Variables with factor loading (> 0.70)	q_1 (0.931)	q_5 (0.974)	q_9 (0.942)	q_{12} (0.945)	q_{15} (0.914)
	q_2 (0.939)	q_6 (0.961)	q_{10} (0.913)	q_{13} (0.937)	q_{16} (0.881)
	q_3 (0.868)	q_7 (0.956)	q_{11} (0.851)	q_{14} (0.926)	q_{17} (0.873)
	q_4 (0.853)	q_8 (0.922)			
% of variance explained	18.852	18.831	16.171	15.128	15.684
Cronbach's α	0.941	0.921	0.907	0.889	0.866

Table 3 Measurement model evaluation

	CR	AVE	MSV	ASV	*PU*	*PEU*	*ATU*	*SN*	*BIU*
PU	0.975	0.938	0.215	0.131	**0.96**				
PEU	0.954	0.867	0.121	0.062	0.43	**0.89**			
ATU	0.976	0.917	0.072	0.038	0.21	0.08	**0.94**		
SN	0.912	0.721	0.046	0.013	− 0.023	− 0.025	0.06	**0.85**	
BIU	0.969	0.893	0.144	0.047	0.27	0.13	0.04	0.03	**0.91**

NB Bold diagonal elements are square root of AVE (average variance extracted)

and AVE > corresponding squared inter construct correlation (SIC) [17, 18]. Thus, in Table 3, the measurement model adequately gives an evidence of construct validity.

5.3 *Structural Model*

After the reliability and validity of the data was ensured, the proposed structural model was tested. A structural model was evaluated to test the hypothesis in this current research. The standardized path coefficients are depicted in Fig. 2. The path analysis results indicated an acceptable model [19, 20] fit as reflected in Table 4.

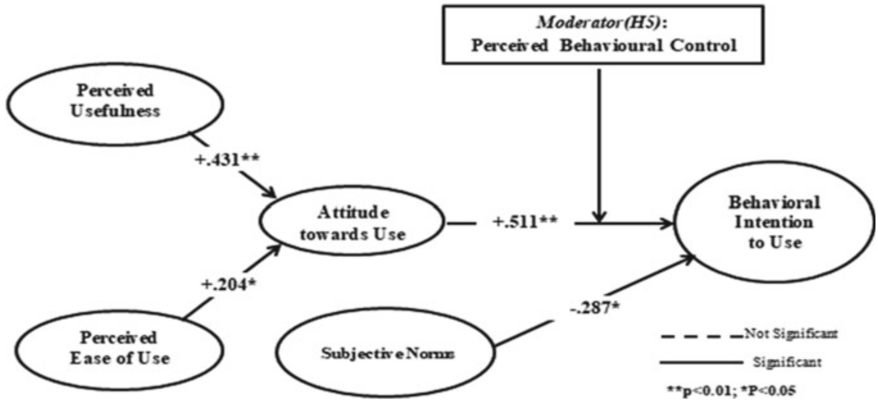


Fig. 2 Path diagram (SEM)

Table 4 Goodness-of-fit measures through SEM

Fit indices with different measures		Recommended values	Fitness values of measurement model
χ^2/df		< 3.00	1.145
Absolute fit indices	GFI	> 0.90	0.974
Incremental fit indices	NFI	> 0.90	0.966
Goodness-of-fit index	CFI	> 0.90	0.881
Badness-of-fit index	RMSEA	< 0.06	0.031
Parsimony fit indices	PCFI	> 0.90	0.913

NB: 1. χ^2/df (Chi-square/degree of freedom); GFI (goodness-of-fit index); NFI (normed fit index); CFI (comparative fit index); RMSEA (root mean square error of approximation); PCFI (parsimonious comparative fit index)

5.4 *Path Analysis Results*

Table 5 gives the path analysis output that reflects the hypotheses assessment. To test the hypotheses, we have calculated standardized regression estimates (β values), and *p*-values. Here, all hypotheses were accepted. 'Subjective norm' has negative significant impact on 'BIU' (H4: $\beta = -0.287$, $p = 0.032$).

5.5 *Moderation Analysis*

See Table 6.

Table 5 Hypothesis testing by SEM

Measurement path			Hypothesis	Regression coefficients (β)	S.E	C.R	*p*-Value	Inference
ATU	←	PU	H1	+ 0.431	0.044	4.925	$< 0.001^{**}$	Accepted
ATU	←	PEU	H2	+ 0.204	0.058	5.031	0.014^{*}	Accepted
BIU	←	ATU	H3	+ 0.511	0.061	3.818	$< 0.001^{**}$	Accepted
BIU	←	SN	H4	– 0.287	0.052	– 4.672	$< 0.032^{*}$	Partly accepted

NB ** and * indicate regression estimate with $p < 0.01$ and $p < 0.05$

Table 6 Moderation analysis

Moderator: *Perceived behavioral control*						
	Regression estimate (β)	*t*-Value	*p*-Value	LLCI	ULCI	Moderation effect
ATU → BIU	**0.513**	16.78	**0.000***	0.384	0.472	Present

NB *represents 1% significance level

6 Discussion

From the above analysis, the following observations have been revealed.

(i) Regression estimate ($\beta = +\,0.431$ with *p*-value < 0.01) indicates that 'PU' has a positive significant influence on 'ATU'. The improved clinical management and cost reduction gear up the efficiency of healthcare professionals which results better satisfaction to the patients; hypothesis H1 is accepted.

(ii) Regression estimate ($\beta = +\,0.204$ with *p*-value < 0.05) indicates that 'PEU' has a positive significant influence on 'ATU'. The comprehensibility and flexibility of the Super AI-based system in healthcare sector may develop more proficiency by the end users and extend its applicability; hypothesis H2 is accepted.

(iii) Regression estimate ($\beta = +\,0.511$ with *p*-value < 0.01) indicates that 'ATU' has a positive significant influence on 'BIU'. The valuableness and affordability of Super AI-based system would result acceptability of the same which in turn increases BIU; hypothesis H3 is accepted.

(iv) The literature review reveals that 'SN' has a positive impact on 'BIU'. But as per our analysis, regression estimate (β = – 0.287 with p-value < 0.05) indicates that the influence is negatively significant. There may be partial or total disagreement among the healthcare professionals, family and friends and patients to experiment the use of super AI-based system, fearing some possibility of unfavorable outcomes on the patients' health, which leads to negative intention toward use; H4 is partially accepted.

(v) 'PBC' is normally guided by user's barriers and user's confidence that initiate user's personal controllability of super AI-based system. Regression estimate ($\beta = +0.513$ with p-value < 0.01) in the moderation analysis indicates presence of moderation effect of positive significant influence of PBC on the relationship between BIU and ATU. Hence, hypothesis H5 is accepted (Table 6).

7 Conclusion and Further Scope of Research

From the findings of the study, it can be concluded that 'PU' and 'PEU' will impact 'ATU' which would further impact 'BIU' moderated by 'PBC'. 'SN' may be considered feebly effective in influencing the 'BIU'. Since Super AI has been proposed to behave smarter than the human brains in every field viz. scientific research, general insight and social skillset, its use will enhance the degree of relationship between the constructs of the model developed for healthcare information system and as per the principle of Super AI, the relationship may be self-modified to make the model more optimized. As super artificial intelligence is still under development phase in the ICT industry, there is limited availability literature in its application. Moreover, respondents selected under study are not users of Super AI to provide sufficient evidence. So further researches may be conducted to exemplify the relationships between the constructs used for TAM- and TPB-based model from the future evidences collected through the deployment of Super AI.

References

1. Bostrom N (2006) How long before super intelligence? Linguist Philos Invest 5(1):11–30
2. Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 13:319–340
3. Ajzen I (1985) From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckman J (eds) Action-control: from cognition to behavior. Springer, Berlin, pp 11–39
4. Narain K, Swami A, Srivastava A, Swami S (2019) Evolution and control of artificial superintelligence (ASI): a management perspective. J Adv Manage Res 16(5):698–714. https://doi.org/10.1108/JAMR-01-2019-0006
5. Ham M, Jeger M, Ivković AF (2015) The role of subjective norms in forming the intention to purchase green food. Econ Res-Ekonomska Istraživanja 28(1):738–748. https://doi.org/10.1080/1331677X.2015.1083875
6. Paul P, Giri A, Chatterjee S, Biswas S (2019) Determining the effectiveness of 'cloud computing' on human resource management by structural equation modeling (SEM) in manufacturing sector of West Bengal, India. Int J Innov Technol Exploring Eng 8(10):1937–1942. https://doi.org/10.35940/ijitee.J9276.0881019
7. Ma Q, Liu L (2004) The technology acceptance model: a meta-analysis of empirical findings. J Organ End User Comput 16(1):59–72. https://doi.org/10.4018/9781591404743.ch006.ch000
8. Weng F, Yang RJ, Ho HJ, Su HM (2018) A TAM-based study of the attitude towards use intention of multimedia among school teachers. Appl Syst Innov (MDPI) 1(3):1–9
9. Chakraborty S, Giri A, Aich A, Biswas S (2020) Evaluating influence of artificial intelligence on human resource management using PLS-SEM (Partial least squares-structural equation modeling). Int J Sci Technol Res 9(3):5876–5880

10. Creswell JW, Creswell J (2003) Research design. Sage, Thousand Oaks
11. Giri A, Biswas W, Biswas D (2018) The impact of social networking sites on college students: a survey study in West Bengal. Indian J Mark 48(8):7–23
12. Giri A, Gangopadhyay S, Majumder J, Paul P (2019) Model development for employee retention in Indian construction industry using structural equation modeling (SEM). Int J Manage 10(4):196–204. https://doi.org/10.34218/IJM.10.4.2019.019
13. Podsakoff PM, Organ DW (1986) Self-reports in organizational research: problems and prospects. J Manag 12(4):531–544
14. Giri A, Chatterjee S, Biswas S, Aich A (2020) Factors influencing consumer purchase intention of daily groceries through b2c websites in metro-cities of India. Int J Sci Technol Res 9(1):719–722
15. Giri A, Pandey M (2016) Relationship marketing as an effective promotional tool of yoga marketing in the urban Indian market: an empirical study. Indian J Mark 46(5):42–54
16. Nunnally JC, Bernstein IH (1994) Psychometric theory, 3rd edn. McGraw Hill, NY
17. Chatterjee S, Giri A, Paul P, Bag M (2019) Impact of 'customer relationship management (CRM) software' on patient satisfaction in public hospitals of urban West Bengal, India: an empirical analysis. Int J Eng Adv Technol 8(5):521–526
18. Giri A, Chatterjee S, Paul P, Chakraborty S (2019) Determining the impact of artificial intelligence on 'developing marketing strategies' in organized retail sector of West Bengal, India. Int J Eng Adv Technol 8(6):3031–3036
19. Giri A, Chatterjee S (2020) Impact of fluid team performance on strategic hrm: An empirical study in the organized retail sector of West Bengal. Prabandhan Indian J Manage 13(4):25–42
20. Chakraborty S, Giri A, Biswas S, Bag M (2020) Measuring the impact of celebrity endorsement on consumer purchase intention of beauty soap in Indian context. Int J Sci Technol Res 9(2):1019–1022

Real-Time Face Recognition and Counting

Archana Chaudhari, Riya Raut, Divyesh Thakur, Ravi Rathod, and Shlok Agarwal

Abstract Facial recognition is an important task in computer vision and is used in security, surveillance, human–computer interaction, etc. It has many applications in fields. Siamese neural networks have emerged as a good method for detection as they can learn good representations of faces and resolve the difference between illumination, illumination and instruction. In this article, we propose a new face detection method based on Siamese neural networks. Our method uses a pair of identical subnets to assign weights and learn to extract features from two face images. The results of the two-web links are then compared with distance measurements to determine if the faces in the input images are the same. Evaluation of proposed method is compared on several datasets and exhibits a good performance in terms of accuracy and speed. Results of the proposed method demonstrate the performance of Siamese neural networks in face recognition and their potential for practical applications.

Keywords Counting · Face · Recognition · Siamese · TensorFlow

A. Chaudhari · R. Raut · D. Thakur · R. Rathod (✉) · S. Agarwal
Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India
e-mail: ravi.rathod20@vit.edu

A. Chaudhari
e-mail: archana.chaudhari@vit.edu

R. Raut
e-mail: riya.raut20@vit.edu

D. Thakur
e-mail: divyesh.thakur20@vit.edu

S. Agarwal
e-mail: shlok.agarwal20@vit.edu

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_43

1 Introduction

Face recognition in crowded environments is a challenging task that has important applications in areas such as security, surveillance, and crowd management. Deep neural networks have emerged as a powerful tool for face recognition due to their ability to learn discriminative features from large amounts of data. In recent years, there has been a significant amount of research focused on developing deep neural network-based approaches for face recognition in crowded environments. The goal of face recognition in crowded environments is to accurately identify individuals in real-world scenarios where there are large numbers of people in a scene. This task is complicated by the fact that faces can vary significantly in terms of pose, illumination, occlusion, and expression, and that individuals in crowded scenes can have similar appearances. Face recognition has recently proven effective in user authentication, replacing older authentication methods such as password protection or one-time passwords. Over the last decade, we've seen a huge increase in the use of face authentication in smart marketing, and there are many apps that can add biological filters to the face, like Snapchat or Instagram. Continuous research in this area has shown many advantages in various applications such as disease surveillance or law enforcement. This article is about exploring the popular algorithm that works behind facial recognition technology. The Siamese network is explored in the proposed study.

2 Related Work

The work presents a literature review on face recognition algorithms as well as face counting algorithms.

2.1 Literature Review on Face Recognition

In this paper, an occlusion-proof face recognition method based on Bidirectional Differential Siamese Networks (PDSN) clearly produces correspondences between occluded face patches and corrupted feature elements. Competitive results of electronic equipment and real occluded face data suggest that the best of the plan is the ability to perform, especially the wide face [1]. In this paper, they utilize a Siamese organize engineering comprising of two parallel CNNs and utilize exchange learning of the VGG-16 demonstrate to perform confront acknowledgment. The comes about appear that the proposed show can progress exactness compared to other comparable strategies prepared on information with less tests. The creator prescribes other CNN systems within the Siamese arrange engineering, particularly communication systems that can extricate high-level highlights in expansion to low-level highlights [2]. This article presents a modern confront acknowledgment and profound learning approach with Siamese Convolutional Neural Systems in our framework. It moreover appears that the integration of neural systems can be done for multi-ideal and multi-classification issues [3]. In this paper, we ask for a deep Siamese mesh architecture to detect deformed faces. This is because the proposed framework learns to distinguish images with slight differences when examining the WVU Twins Day dataset and learns the characteristics of the corresponding data deformed by training in the training part of the dataset [4]. The proposed L-CNN aims to extract local features from areas of the face such as the nose, mouth, and eyes. G-CNN aims to capture the features of the entire face. Complementing each other, global and local proxies are created within the framework of social filtering to distinguish the target from their content and other faces. They also collected and identified the largest face tracking dataset, far from saturation, facilitating further research in this area. Extensive testing results on the collected data show that the proposed method performs well compared to many modern trackers [5]. In rundown, the qualification of low-level-based strategies is restricted, and the few profound learning strategies accessible need an understanding of how Halfway bottlenecks really influence the CNN demonstrate. Differentiating between two faces with diverse impediment conditions was not taken under consideration. The arrange fills within the lost parts of the coordinate and can confirm reasonable comparison, disposing of negative substance for prepared and disposed of CNN models. Therefore, our method is comparable to the above research.

2.2 Literature Review on Face Counting

In this article, the cloud platform provided by Google is used to manage and store the information of various colleges/universities. The detect Multiscale() method in Cascade Classifier is used to identify faces (i.e. the number of students) in the video [6]. Social behaviours such as shopping-to-sell or shopping-to-buy are common in the urban part of Indonesia on days off and holidays. It is Friday, Saturday, and Sunday. This behaviour is similar to the data provided by the method proposed in this study. In fact, the actual number of visitors in the store data is 65.08%, and the actual rate of passers-by is 66%0.12%, but the car model is the same as the behaviour data says. This model shows that this demand can also be used as a counter [7].

3 Proposed Methodology for Face Recognition and Face Counting

Unlike traditional CNNs, Siamese network does not divide images into groups or tags, but simply shows the distance between two given images. Here's how the Siamese network works. A pair of images is selected from the data to train the Siamese network. Since the networks have the same structure, the same processing will be done for each image. The final neural network has a fully connected layer, and the final layer has 128 nodes. This layer is the last feature created by the network when applied to the image. It is called embedded layer representation. Thus, two images in a pair of images processed by the Siamese network form two different embedding layer representations, as shown in the system architecture in Fig. 1. If the images are the same person, the embeddings will be similar, so the distance should be smaller. However, if the images of people are different, the distance should be a higher value. The Sigmoid function is used to bring the distances to the range 0–1. Put a loss on the resulting sigmoid to punish the network for adjusting its weights and biases.

Weight and update rate are the same on both networks. Repeat this process for all image pairs created from the file. We use TensorFlow API to build the network architecture.

We are using the Olivetti dataset obtained from the sklearn dataset API for my facial recognition system. For evaluation, we want to use the image we want to know, which is not in the dataset. To do this, we need to convert these images to the same format before using them for testing.

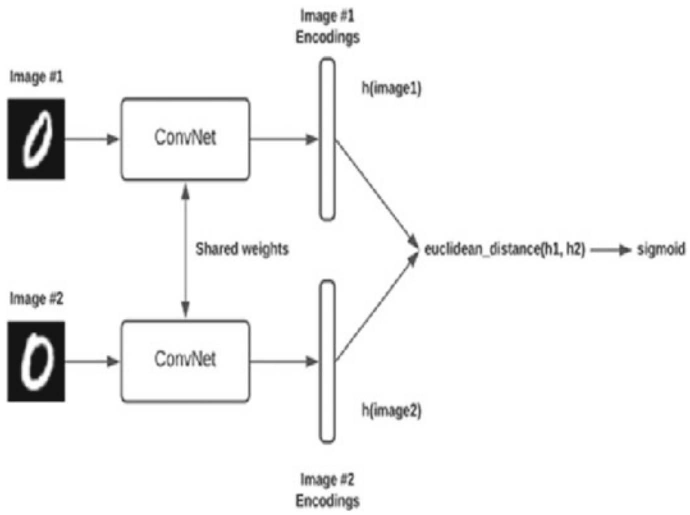


Fig. 1 Architecture of system

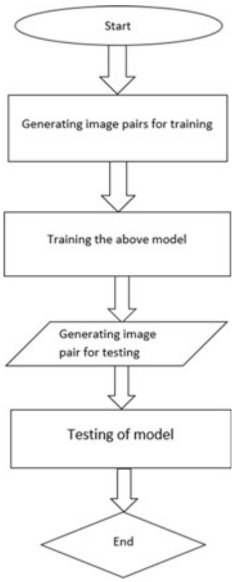


Fig. 2 Flow chart of proposed model

Figure 2 describes the flow of working to build the proposed model of Siamese network for face recognition and counting.

Overall, the key idea behind the working principle of a Siamese network is to learn a shared representation for input data points that captures their similarity or dissimilarity, enabling tasks such as similarity matching, verification, or retrieval.

4 Results and Discussions

4.1 Proposed Model Using Siamese Network Gives Us Following Results

Proposed model is counting faces in the frame using OpenCV as shown in Fig. 3.

In Fig. 4, proposed model gives value true as it recognizes the face1 (as seen in Fig. 3), which is given to proposed model to identify using Siamese network.

4.2 Performance of Proposed Model

In deep learning, the learning model uses failure. The aim of the model is to minimize the cost loss. The truth is much simpler. It measures how well our prediction model works by comparing the model prediction to the true value as a percentage. Less

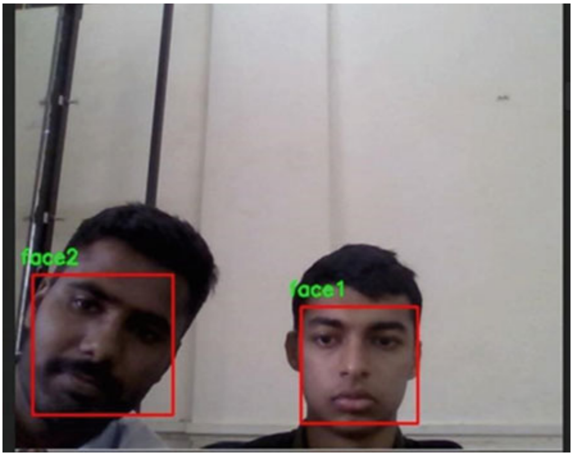


Fig. 3 Counting of faces in frame

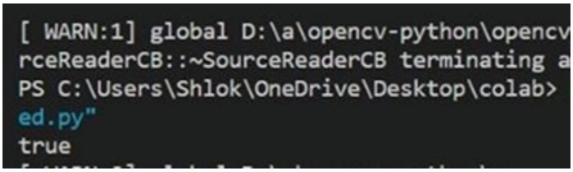


Fig. 4 Recognition of face from frame

accurate but disappointing means that the model may not be fair for most of the data. However, if both are declining and accuracy is low, there is a slight error in most of the information in the sample. However, if both are high, some data may have serious errors. Finally, it would be ideal to have real ups and downs so that the model has only a small error in some data.

The plot in Fig. 5 demonstrates the history of model accuracy and failure during training and validation. After training 100 times it ids observed from Fig. 5 that, the efficiency seems to drop by as much as 80%, it hasn't changed much here. Again, no significant improvement in accuracy after the first few iterations. The training loss should decrease over time, indicating that the network is learning to correctly classify similar and dissimilar pairs. From the plot in Fig. 5, we did not get no noticeable change because the exact shape of the plots can depend on a various factors, such as the size of the training dataset, the complexity of the network architecture, and the hyperparameters chosen for training.

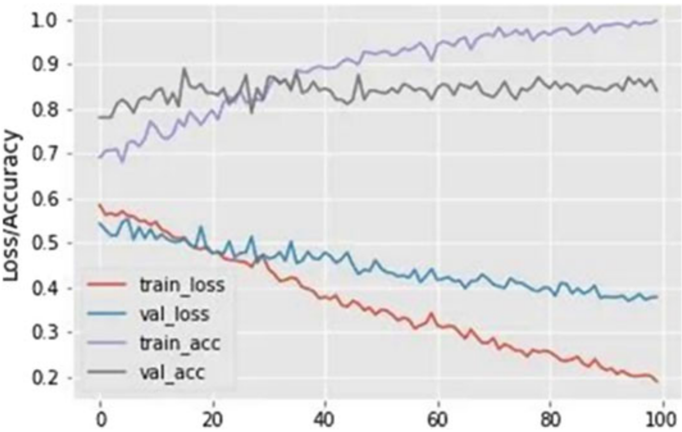


Fig. 5 Training loss and accuracy plot of proposed model of Siamese network

5 Conclusion

Siamese neural network is a cutting-edge algorithm that quickly and efficiently identifies differences between images. Algorithms that learn important features quickly and effectively open up new opportunities to distinguish between these features, especially for their use in areas such as video surveillance, facial recognition, and self-learning. During the implementation of the Siamese neural network, we found that it was very expensive and time consuming; For example, training the model takes time and equipment. It's also not the most important thing to try to fit each face into one of the characters/images. If a new face appears on the screen, it is assigned to one person or another. This problem can be solved by carefully selecting the threshold to know the similarity. Integrating Siamese with OpenCV can count the number of faces on the frame.

References

1. Song L et al (2019) Occlusion robust face recognition based on mask learning with pairwise differential siamese network. In: Proceedings of the IEEE/CVF international conference on computer vision
2. Heidari M, Kazim F-G (2020) Using siamese networks with transfer learning for face recognition on small-samples datasets. In: 2020 international conference on machine vision and image processing (MVIP). IEEE
3. Wu H et al (2017) Face recognition based on convolution siamese networks. In: 2017 10th international congress on image and signal processing, BioMedical engineering and informatics (CISP-BMEI). IEEE
4. Soleymani S et al (2021) Differential morphed face detection using deep siamese networks. In: Pattern recognition: ICPR international workshops and challenges: virtual event, January 10–15,

2021, proceedings, part VI. Springer, Cham (Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999))
5. Qi Y et al (2020) Siamese local and global networks for robust face tracking. IEEE Trans Image Process 29:9152–9164
6. Mehariya J et al (2020) Counting students using OpenCV and integration with firebase for classroom allocation. In: 2020 international conference on electronics and sustainable communication systems (ICESC). IEEE, 2020
7. Herviana A, Sudiharto DW, Yulianto FA (2020) The prototype of in-store visitor and people passing counters using single shot detector performed by OpenCV. In: 2020 1st international conference on information technology, advanced mechanical and electrical engineering (ICITAMEE). IEEE

Information System for Neuropathy Prediction Ensembling Ranking and Ordered Clustering for Diabetic Healthcare Monitoring

K. Poorani and M. Karuppasamy

Abstract Role of information system in health care is to provide a detailed view and easier aspect on major factors causing the defect. Predicting the major factors can lessen the severity of the disease. Diabetic neuropathy is found when a diabetic does not maintain the blood glucose level properly. This happens not only with the improper maintenance of blood glucose levels but also with the other factors like high cholesterol, smoking, alcohol consumption, less physical activity, obesity, and so on. This approach estimates the possible factors for the onset of neuropathy with diabetes health indicators and builds a model to diagnose it earlier. The K-means has been performed to separate the features and build a predictive model that helps in analyzing the complication factors of diabetes. Building an information system for healthcare application can be done effectively with the help of machine learning techniques.

Keywords Diabetes mellitus · Neuropathy · Ranking system · K-means clustering

1 Introduction

Diabetic neuropathy is the most prevaling complication of diabetes found with diabetic patients when the duration of diabetes is higher with improper maintenance of blood glucose levels [1]. It affects different parts of the nervous system. According to American Diabetes Association (ADA), almost half of the diabetic people have some form of nerve damage which may lead to amputation. Peripheral nervous system is the most affected part of the nervous system for diabetic people [2]. Symptoms of diabetic neuropathy include loss of sensation, muscle loss, tingling, and burning sensation. According to International Diabetes Foundation (IDF), 40–60 million people are affected with diabetic foot and lower limb complication in global

K. Poorani (✉) · M. Karuppasamy
Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India
e-mail: pooranikhari@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_44

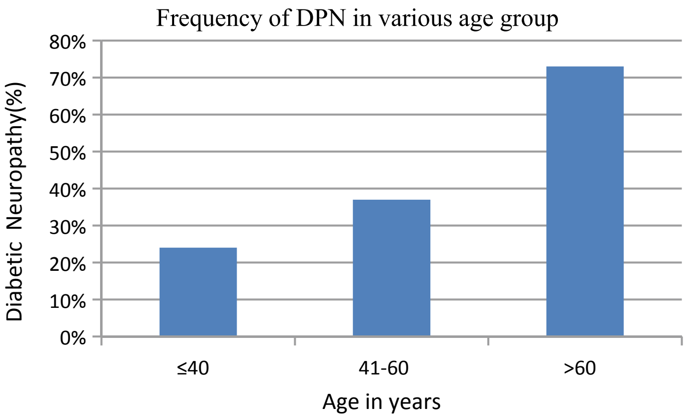


Fig. 1 Incidence of diabetic neuropathy with age

level [3]. It was very difficult for the physicians to recognize the disease at earlier stage which leads to high mortality. The prevalence of diabetic peripheral neuropathy has been described in Fig. 1. This represents that patient with age greater than 60 are more affected with peripheral neuropathy. Thus, they should be taken with great care in case of neuropathy.

Tankova et al. [4] in their studies explained that diabetic sensorimotor polyneuropathy (DSPN) is a very common form of neuropathy. It usually affects lower limb nerves the most. Long-term DSPN leads to limb amputation which increases the chances for early death.

Neuropathy is caused by several factors in which diabetes is the one of prominent cause of peripheral neuropathy in most cases. Peripheral neuropathy affects the hand and leg extremities the most. Whenever the diabetic patient is not concerned with the blood glucose monitoring, it affects the small blood vessels which lead to many microvascular complications. Even the daily habits like walking, physical health, mental health, cholesterol level, BMI affects the blood glucose levels, and it is good to maintain them at proper levels. Thus, many factors responsible for neuropathy are predicted in this work.

Iqbal et al. [5] in their work demonstrated the epidemiology, diagnosis methods, and pharmacotherapy on diabetes peripheral neuropathy. They found that the clinical system to find or misdiagnosis leads to high mortality in diabetic neuropathy cases. Glycemic control has been the main constituent to maintain the diabetic levels. Cardiovascular risk factors are highly dependent on neuropathy.

2 Related Work

Diabetic neuropathy is one of the major complications of diabetic. Many researchers concentrated only on the neurodegeneration for the onset of neuropathy in diabetes. But the normal common features with daily habits which may be responsible for the neuropathy onset has been analyzed in this work.

Callaghan et al. [6] described central obesity and its association with diabetic neuropathy progression. They employed obese group and lean group people and performed multivariate logistic regression to find the neuropathy progression. Obesity is the major problem of diabetic patient which causes increase in BMI range. Obesity is the underlying factor for many kinds of diseases. Their findings prove that normoglycemic patients have high prevalence of neuropathy. Not only obesity but waist circumference also plays a major role. Naranjo et al. [7] pointed out the most important features like anxiety, sleep disorders and depression factors in people with diabetic neuropathy. Thus people suffer lot of pain in case of neuropathy which should be predicted earlier.

Xu et al. [8] in their work investigated the predictive model for neuropathy detection and found the most common factors used for prediction are age, duration of diabetes, gender, BP, lipid levels, and mean HbA1c levels for Type 1 diabetic patients. The most commonly used methods include classification and regression algorithm (CART) random forest, support vector machine, logistic regression, and Neural Networks. Malik et al. [9] described the prevalence of peripheral neuropathy in South East Asia. The reason behind this is lack of guidance and less chance for diagnosis. Thus, neuropathy has to be found earlier to reduce its impact on diabetic patient. Excessive intake of alcohol causes severe diabetic complications including heart disease and diabetic neuropathy. Other complications like heart disease and its incidence on neuropathy due to medication have to be analyzed.

Sambyal et al. [10] examined various microvascular complications associated with diabetes and reviewed on statistical, machine learning and deep learning models for disease prediction and found machine learning models has been found good for disease prediction.They concluded that both the models can be combined to find the disease at very early stage. There were various factors influencing onset of neuropathy in diabetic patients. Thus, self-monitoring and control over blood glucose levels are the major ways to keep off neuropathy.

Jende et al. [11] defines the association of serum cholesterol level and its effect for neuropathy. Their findings prove that the lowering serum cholesterol levels have been associated with DPN. This helps to find more serum cholesterol lowering methods. Lou et al. [12] in their findings classified a tree model to analyze the risk factors of neuropathy in diabetic patients. They employed logistic regression analysis to find the factors of higher importance. There were many factors like duration of diabetes, HbA1c, BMI, haemoglobin, total cholesterol, hypertension, Fasting Blood Glucose (FBG) and retinopathy which are responsible for neuropathy. Among all factors HbA1c, FBG and hypertension has strong association with neuropathy in their findings.

3 Proposed Methodology

This is a novel method which focuses on the predictive data analytics and factors that accompany the diabetes complications and finding the onset of diabetes. The aim is to figure out the underlying most complicating factors of diabetes and present a model to predict the onset of neuropathy. The diabhealth dataset has been taken for analysis. It consists of 21 various columns and 253,680 rows. Among all the columns most important and factors responsible for neuropathy has been considered for analysis.

The Diabetes Health Indicators dataset is collected from kaggle. Since neuropathy is a complication of diabetes, it is a must to extract the features that leads to neuropathy progression. The features in the dataset include diabetes status, high BP, high cholesterol, BMI, smoking, stroke, heart disease, physical activity, Fruits, Veggies, heavy alcohol, Genhealth, and so on. The features are finely tuned to predict the major reasons for the progression of diabetes neuropathy. Features with less importance like income and education have been eliminated as they do not play a significant role in neuropathy progression. Preprocessing has been done and further process is carried out.

The high incidence of cholesterol is found with the described dataset. Likewise the incidence of every feature has been build and modeled for predictive analytics. The other factors with higher incidence have been considered for model prediction.

Figure 2 highlights that high cholesterol incidence may be predictor for neuropathy onset. The higher incidence factors are the predicting factors of diabetes neuropathy. Likewise only major factors contributing to neuropathy has been clustered using K-means. Certain factors are clustered on the predictive framed model.

Analyzing the underlying causes that affect neuropathy, features importance is found and clustering has been performed. The factors of high feature importance include high BP, High Chol, Phys Activity, heavy alcohol, Stroke, Age, BMI, smoke are clustered as the major features responsible for neuropathy pain in the diabetes patients.

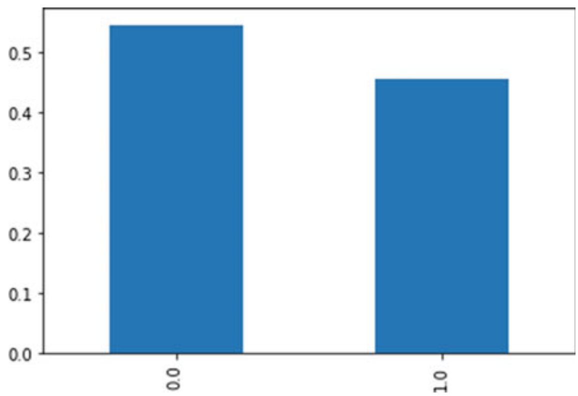


Fig. 2 Incidence of cholesterol in the dataset

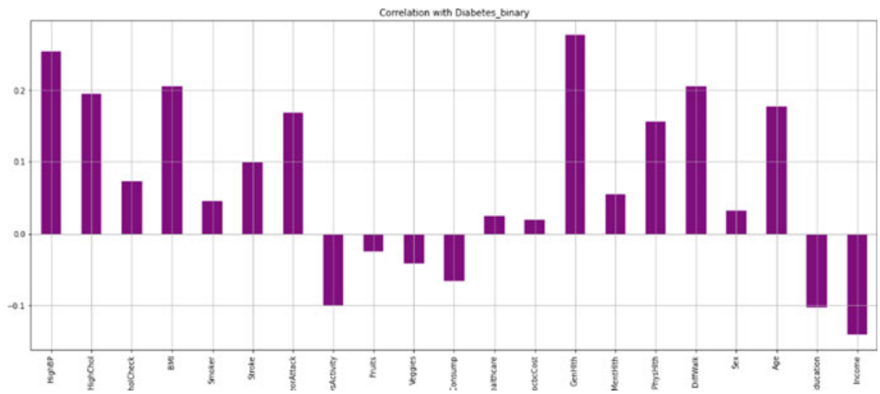


Fig. 3 Correlation of factors from diabetes binary

The major factors responsible for onset of neuropathy are analyzed, and clustering has been performed to form the group of factors that may cause neuropathy in diabetes patients. Thus, features with less importance have been left out. Diabetes binary has been predicted with all the features, and their correlation has been analyzed. The strongly correlated factors has been classified accordingly and processed.

Figure 3 represents correlation of each factor. Depending on the correlation factor, features has been extracted and predictive model has been build. Features with high correlation are the factors to be considered for the health of diabetic. K-means has been implemented and found important cluster with factors highly responsible for neuropathy. Totally, five features have been removed from the list, and process of clustering has been carried out with the other 16 attributes.

F-score is one of the validation metrics to find the model evaluation. F-score is also measured with the selected features, and their values are given above. Table 1 represents the f-score values of selected attributes for predictive model. The F-score value has been taken for clustering data accordingly. The cluster with higher to lower order is given higher and lower probability, respectively, and predictive model has been built. Since the data has been clustered according to the probability the cluster with high probability has higher chances of neuropathy onset and is classified accordingly. Figure 4 represents the confusion matrix obtained from important features selected for prediction.Table 1 explains the F-score of various features. Dataset has been taken form a repository [13].

Table 1 Representation of F-score

	Feature	Score
15	PhysHlth	133,424.406534
14	MenHlth	21,029.632228
3	BMI	18,355.166400
16	DiffWalk	10,059.506391
0	HighBP	10,029.013935
13	GenHlth	9938.507776
18	Age	9276.141199
6	HeartDisease or Attack	7221.975378
1	HighChol	5859.710582
20	Income	4829.816361
5	Stroke	2725.225194
7	PhyActivity	861.887532
10	HvyAlcoholConsump	779.424807
19	Education	756.035496
4	Smoker	521.978858
12	NoDocbccost	229.542412

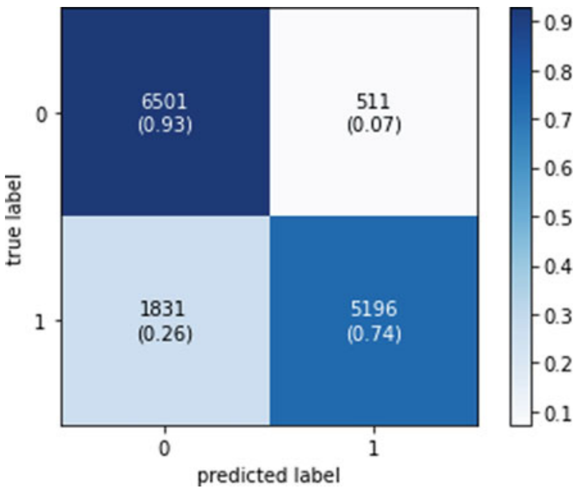


Fig. 4 Confusion matrix

4 Clustering and Ranking for Neuropathy Prediction

K-means clustering and ranking are used for partitioning clusters based on ranking system as there is no pre-specifying number of clusters to be created. The dataset is divided into clusters to create a tree-like structure. Ranking is based on correlation analysis.

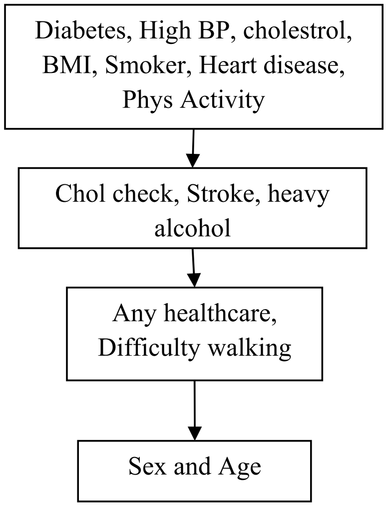


Fig. 5 Representation of ranking order

Spearman's Rank Correlation algorithm has been used to find the strength and distance of two ranked variables which helps in finding the various orders for building a predictive model.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n2 - 1)}$$

The classified higher order factors include diabetes, high BP, high Chol, BMI, Heart disease, Physical Activity, smoker that are considered as high ordered cluster and are considered as major factors for neuropathy onset. The secondly classified chol check, stroke, heavy alcohol are considered as second most important cluster, and the remaining factors are proceeded accordingly as the less important clusters (Fig. 5).

5 Discussion

Information system can be built in such a way that the factors at top level are to be given higher concentration and targeted therapeutic importance. Diabetic data are collected and managed with great care with essential drug management in order to avoid further complications and to reduce the mortality rate.

6 Conclusion

Diabetes is a chronic metabolic syndrome leading to many serious complications which has to be found earlier and maintained properly. Diabetic neuropathy which affects the hands and legs of the diabetic patients can even lead to foot ulcer and progressively to limb amputation. So various factors responsible for neuropathy onset have been discussed and found a predictive data model. The predictive model using k-means clustering gave a good result. Thus if the features are maintained at proper levels, then the progression of diabetic complication can be slowed down. This proposed method is being implemented on existing dataset for features concerning to diabetic neuropathy.

7 Future Work

Future work focuses on implementing the model in real-time healthcare dataset with electronic records and predicts various complications before its onset for better living of diabetic patients. HbA1c which is the most important factor for maintenance of all diabetic patients has to be taken into consideration.

References

1. Center for Disease Control and Prevention (CDC). https://www.cdc.gov/diabetes/basics/diabetes.html
2. American Diabetes Association (ADA)
3. International Diabetes Federation (IDF)
4. Ziegler D, Tesfaye S, Spallone V, Gurieva I, Al Kaabi J, Mankovsky B, Martinka E, Radulian G, Nguyen KT, Stirban AO, Tankova T (2022) Screening, diagnosis and management of diabetic sensorimotor polyneuropathy in clinical practice: international expert consensus recommendations. Diab Res Clin Pract 186
5. Iqbal Z, Azmi S, Yadav R, Ferdousi M, Kumar M, Cuthbertson DJ, Lim J, Malik RA, Alam U (2018) Diabetic peripheral neuropathy: epidemiology, diagnosis, and pharmacotherapy. Clin Ther 40(6):828–849
6. Callaghan BC, Reynolds E, Banerjee M, Chant E, Villegas-Umana E, Feldman EL (2020) Central obesity is associated with neuropathy in the severely obese. In: Mayo clinic proceedings, vol 95, no 7. Elsevier, Amsterdam, pp 1342–1353
7. Naranjo C, Del Reguero L, Moratalla G, Hercberg M, Valenzuela M, Failde I (2019) Anxiety, depression and sleep disorders in patients with diabetic neuropathic pain: a systematic review. Expert Rev Neurother 19(12):1201–1209
8. Xu Q, Wang L, Sansgiry SS (2020) A systematic literature review of predicting diabetic retinopathy, nephropathy and neuropathy in patients with type 1 diabetes using machine learning. J Med Artif Intell 3:6
9. Malik RA, Andag-Silva A, Dejthevaporn C, Hakim M, Koh JS, Pinzon R, Sukor N, Wong KS (2020) Diagnosing peripheral neuropathy in South-East Asia: a focus on diabetic neuropathy. J Diabetes Invest 11(5):1097–1103

10. Sambyal N, Saini P, Syal R (2020) Microvascular complications in type-2 diabetes: a review of statistical techniques and machine learning models. Wirel Pers Commun 115(1):1–26
11. Jende JM, Groener JB, Rother C, Kender Z, Hahn A, Hilgenfeld T, Juerchott A, Preisner F, Heiland S, Kopf S, Pham M (2019) Association of serum cholesterol levels with peripheral nerve damage in patients with type 2 diabetes. JAMA Network Open 2(5):e194798
12. Lou J, Jing L, Yang H, Qin F, Long W, Shi R (2019) Risk factors for diabetic nephropathy complications in community patients with type 2 diabetes mellitus in Shanghai: logistic regression and classification tree model analysis. Int J Health Plann Manage 34(3):1013–1024
13. https://www.kaggle.com/datasets/diabetes-health-indicators-dataset

Predicting Foreign Exchange Rates Using Machine Learning Techniques

Shubham Bhange, Kumbhar Vidya, and Sachin Naik

Abstract The forex market is a volatile and unpredictable financial market in the world, and it is the most challenging market to forecast accurately. Data for predicting the forex rate is time series data, which contains trends and seasonality. Data containing seasonality and trends are difficult to forecast. The ability to estimate this foreign exchange rate (FOREX) has therefore become crucial in the financial market. The risk of the trade investment can be reduced through accurate exchange rate forecasting, resulting in seamless global exchange and, most significantly, preserving investor returns. In this study, a machine learning models are built for predicting the forex rate. Support Vector Regressor (SVR), random forest regressor, and K nearest neighbor (KNN) these three models are tested on 5 years of historical data of the United State Dollar against the Indian Rupee, i.e., (USD/INR) to predict future closing prices. The prediction is made for the next 10 days closing forex prices. Furthermore, we increase the performance of models with the help of respective parameters. By changing the values of parameters, it is observed that model performance is increased. A comparative study of on performance of the proposed three models is assessed using various performance metrics, such as R^2, RMSE, MSE, MAE, explained variance, Mean Gamma Deviance (MGD), and Mean Poisson Deviance (MPD). SVR gives the better result with an optimal parameter such as linear kernel, $C = 1e2$, $\epsilon = 0.01$, where R^2 score, RMSE, and MSE are 0.9508, 0.2379, 0.0566, respectively, along with MAE, MGD, MPD are 0.1758, 1.02e−05, 0.0007, and explained variance is 0.9511.

Keywords Foreign exchange rate · Machine learning · SVR · Random forest · KNN

S. Bhange · K. Vidya (✉)
Symbiosis Institute of Geoinformatics, Symbiosis International (Deemed University), Pune, India
e-mail: kumbharvidya@gmail.com

S. Naik
Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Pune, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_45

1 Introduction

Money is the most significant and determining component, and it is crucial that we understand how to use it effectively. Nations and their citizens have begun investing in a variety of businesses in an effort to utilize the funds effectively and decided to utilize shares and other nations in addition to their own, which are primarily based on currency exchange rates. The organization with the best solution will always prevail and who prepare and invest by comprehending the monetary market's conditions (i.e., its ups and downs) will succeed [7].

Machine learning, in an instant, makes fewer subjective hypotheses and excels at nonlinear fitting. The commonly used conventional machine learning techniques are SVM [6] and random forest [16]. In this study, support vector regression (SVR), K nearest neighbor (KNN), and random forest regression to forecast USD/INR are used. One of the elements and K nearest neighbor of the model for foreign exchange proposed in this study is support vector regression (SVR). SVR is established on the structured risk minimization principle and is intimately associated with SVM [16]. Due to the infinite-dimensional feature vector that the Gaussian kernel creates, SVR with a Gaussian kernel gives the ability to create a nonlinear boundary. In SVM literature, the SVM algorithm is referred to as Support Vector Classification (SVC) whenever applied to classification tasks and Support Vector Regression whenever applied to regression tasks (SVR).

In a random forest regressor, with the aid of a group of decision trees, the ensemble learning method known as random forest can accomplish both regression and classification tasks. Ensemble learning is the name given to this approach since an ensemble is a collection of predictions [5]. Bootstrapping Aggregation or the bagging method is used for ensemble learning. Due to its randomness, the random forest's generalization error is smaller than that of a single decision tree, reducing the variance of the model [5]. The KNN approach assigns the new data point in the group that resembles the current categories the most, assuming that the new and previous cases are equivalent [1]. In all the supervised machine learning algorithms SVR, RF regressor, and KNN, this study estimates the next 10 days' closing forex rate using different parameters presented in the respective machine learning algorithm. In SVR, various parameters are used to evaluate the SVR model using different kernels and changing the values of C and ϵ. In this experiment, treating gamma as constant for all kernels. SVR contains different kernels such as linear, polynomial, RBF, and sigmoid. The performance evaluation is carried out for SVR with respect to kernels and changing values of C and ϵ. The effectiveness of RF deployment depends on figuring out how many trees and predictors each node needs to produce a decent response [15]. For random forest regressor, we are treating various parameters as default, but we have evaluated our RF model by taking a different number of trees. In the case of the KNN model, the effectiveness of KNN depends on k values. So we are taking the various k to find the optimal value of k, and this evaluation can also be calculated by plotting k values with the RMSE. For the performance assessment of our supervised machine learning algorithm SVR, RF regressor, and KNN, different metrics have been used. The metrics are R^2, RMSE, MSE, MAE, explained variance, Mean Gamma Deviance (MGD), and Mean Poisson Deviance (MPD). Many researchers used RSME, R^2, MSE, and MAE in their research papers and used RSME to evaluate model performance.

2 Method

The following Fig. 1 shows overall methodology for the prediction of foreign exchange rates.

2.1 *Data Collection*

To experiment, dataset is downloaded from Yahoo Finance (https://finance.yahoo.com/), in csv format. Data contains the historical foreign exchange rate of the United State dollar (USD) against the Indian rupee (INR), i.e., USD/the INR. This is the

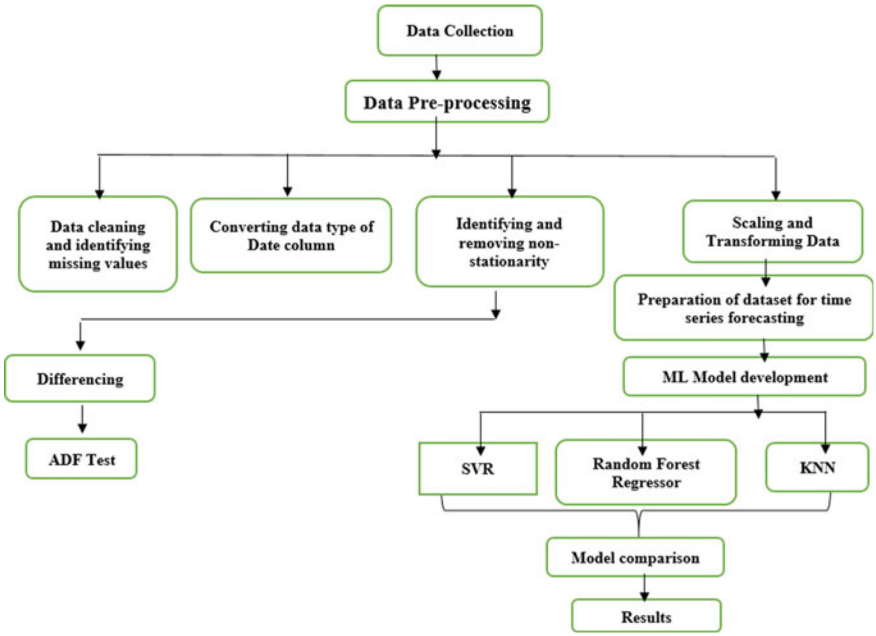


Fig. 1 Methodology for prediction of foreign exchange rates

Fig. 2 Predicting foreign exchange rates

previous five years of data from 2017-05-12 to 2022-05-12 and a compromise of 1826 days, containing 1302 records with 6 columns such as date, open, high, low, close, and adjusted close. To create accurate out-of-sample forecasts that can be compared to the actual observations, the first 846 data points served as the training set, and the final 456 data points served as the test set. Figure 2 represents all our data features (Open, High, Low, Close, and Adjusted Close) and gives an idea about the data.

2.2 Data Pre-processing

Prior to creating the prediction model, procedures that need to be completed as part of data pre-processing must be carried out. Data cleaning (the detection of null values, outliers, and their eradication or corrections) and data scaling and/or data transformations are basic jobs that fall under this category [10]. Time series forecasting also encompasses additional data pre-processing tasks. Making stationary time series out of non-stationary time series is one option. In this study, the dickey-fuller test was employed to complete the task, and another job is to create a new dataset according to the requirement of time series prediction. This study focuses on to forecast an accurate forex rate, and we have taken time steps as 15 to predict next 10 days forex rate. For predicting future forex values, we need to mold our dataset to create new data set. For this, we need to create a matrix with time steps and adding previous data to that matrix. Suppose we want to predict future $Y = t$ forex values, so we need previous data items. According to previously taken data item such as X_{t-1}, X_{t-2}, $X_{t-3}\ldots$, we forecast the next n items.

2.3 Support Vector Machine

Vapnik and fellow teammates at AT&T previously created the idea of a support vector machine (SVM) [14]. "SVM is an optimization" strategy that looks for a hyperplane in the initial input space to accurately divide a particular training set and allow much more space on each side of the hyperplane as possible [17]. The data points that achieve the maximum margin during regression estimation are referred to as support vectors. They are the number of observations, in other words, whose estimation losses are equivalent to or greater than the so-called SVM tube size [17]. A hyperplane that can divide the occurrences in the same feature space is therefore sought after in the feature space. A kernel function specifies the linkage between the input space and the feature space. By including a penalty factor C in the simulation study, the method also allows for misinterpretation, and the overall penalty is determined by adding the penalties for every misinterpretation or misclassification [2]. As a result, the method locates a hyperplane that aims to minimize penalty plus the inverse of the margin.

Various top-notch works provide a thorough understanding of SVM [14], $P = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, consequently, *yi* is the appropriate scalar real dependent variable and *xi* is a vector of truly independent factors. It is essential to estimate the regression equation in the feature space by (Fig. 3):

$$z(x, w) = (w.\phi(x) + c), \tag{1}$$

where

- $w =$ weight vectors
- $c =$ constant
- $\phi(x) =$ feature function
- $(w.\phi(x))$ is the dot product therein.

Minimize the below equation;

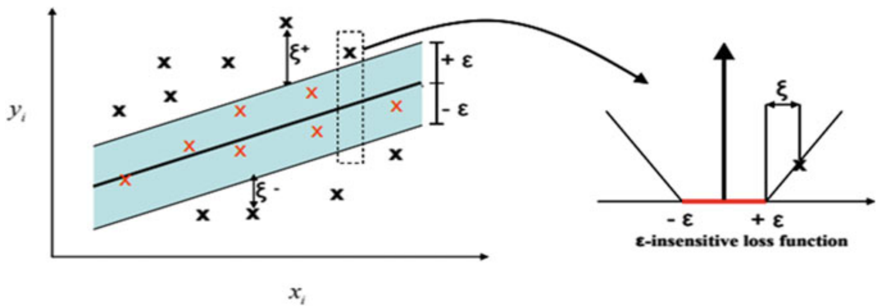


Fig. 3 Diagram of SVR

$$\text{Minimize: } Q(f) = C\frac{1}{N}L\epsilon(y, z(x, w)) + \frac{1}{2}||W2|| \tag{2}$$

and,

$$\begin{array}{ll} L\epsilon(y, z(x, w)) = 0 & \text{if } |y - z(x, w)| \leq \epsilon \\ |y - z(x, w)| - \epsilon & \text{otherwise} \end{array} \tag{3}$$

The actual loss is represented by the LHS part in Eq. (2), and the term C serves as an indicator of the optimization between the actual loss and the model complexity given by the equation's second term. A loss function named the insensitive loss function is defined by Eq. (3) [14]. By adding Lagrange multipliers and *, the optimization problem can be transformed into a dual problems. The support vectors are mostly the non-zero components and their input vectors, *xi*. The finished form appears as follows:

$$z(x, \beta i, \beta i*) = \sum_{i=1}^{N}(\beta i - \beta i*)(\phi(xi)0.\phi(xj)) + c \tag{4}$$

With the help of kernel function $K(xi, xj)$, we can obtain SVR equation as follows:

$$z(x, \beta i, \beta i*) = \sum_{i=1}^{N}(\beta i - \beta i*)K(xi, xj) + c \tag{5}$$

In SVR, there are popular kernels: Linear kernel, Polynomial kernel, radial basis kernel, and sigmoid kernel. In support vector machines [4] and approximation theory [17], the "radial basis function (RBF) kernels are frequently utilized." Due to its remarkable characteristics, the RBF kernel is typically a good first preference.

2.4 Random Forest Regressor

The extremely popular ensemble learning approach known as random forest (RF) entails the construction and combination of a large number of randomized decision trees to create a random forest, which is then applied to classification or regression [5]. Training such RFs is incredibly simple and effective [15]. We will frequently probably get better estimates if we combine the predictions of a number of predictors (such as classifiers or regressors) than we would with the best individual predictor. Ensemble learning is the name given to this approach since an ensemble is a collection of predictions. Bootstrapping, Aggregation, or the bagging approach are used for ensemble learning. Due to its randomness, the random forest's generalization error is smaller than that of a single decision tree, reducing the variance of the model [7].

Breiman first put forth RFs two decades ago [3], drawing inspiration from initial studies in the feature selection technique, the random subspace method, and the random split selection framework. Due to their capabilities in handling a wide range of data types, RFs have found great success across a wide range of industries. The input is a quantitative quantity; we concentrate on the regression component of a random forest RF in this instance. Formally, assume $PXY(X, Y)$ be an unknown joint distribution, and let Y denote a numerical response and $X = (X_1, X_2, X_3, \ldots, X_p)$T be a p-dimensional vector of predictors. The objective is to calculate a forecasting equation $f(x)$ for Y that minimizes a loss function's anticipated value, Loss $(Y, f(X))$

$$EXY(\text{Loss}(Y, f(X))) \tag{6}$$

The condition expectancy is created by minimizing the estimated value for squared error loss [3].

$$f(X) = E(Y|X = x) \tag{7}$$

Based on a set of so-called base learners, $l1(x), l2(x), \ldots lk(x)$, ensembles construct f. A tree with the formula $lk(X, \theta k)$ is a Kth base learner in a RF, wherein θk is a set of independent random variables for $K = 1, 2, \ldots, K$. In regression, the forecast function is determined by averaging the base learners. The expected Y is therefore given by

$$f(x) = \frac{1}{K} \sum_{K=1}^{K} lk(x) \tag{8}$$

The binary recursive partitioning approach which is described by BREIMAN [3] is used in the construction of the trees that make up a RF. A decision tree is built using recursive partitioning by dividing every vertex in the tree into two successors according to a splitting criterion. The final division of the predictor space is made up of terminal nodes, which are nodes that cannot be further divided [15]. The mean squared error at the node serves as a common splitting criterion when used to regression models (Fig. 4).

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y - \overline{y})2 \tag{9}$$

2.5 K Nearest Neighbor

One of K nearest Neighbor is one of the easiest machine learning algorithms, based on the supervised learning method. KNN algorithm assumes that the new case and

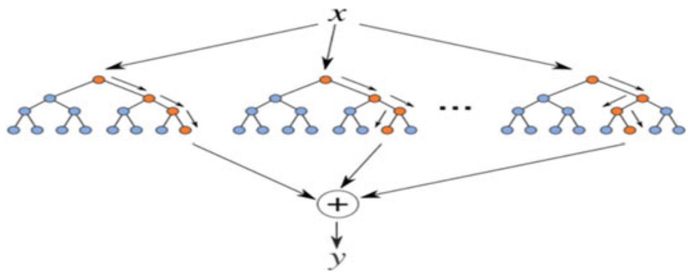


Fig. 4 Diagram of random forest

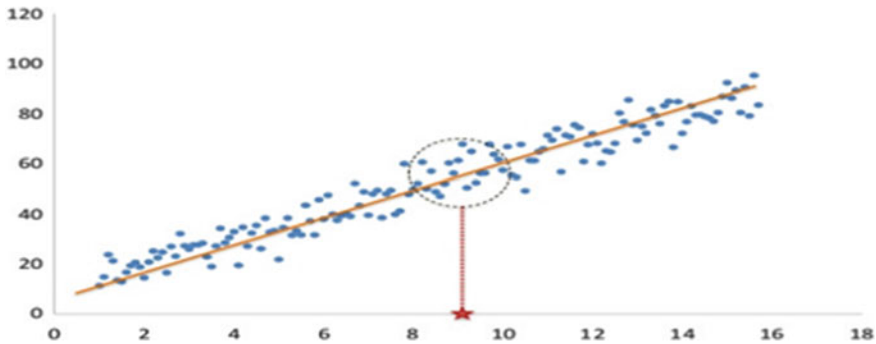


Fig. 5 Diagram of KNN

the existing cases are comparable, and it places the new instance in the category that is most like the existing categories [8] (Fig. 5).

A "new data item is classified using the KNN algorithm based on similarity after all the existing data has been stored." "This means that utilizing KNN method, fresh data can be quickly and accurately sorted into a suitable category." "Although the KNN approach is most frequently employed for classification problems, it can also be utilized for regression."

3 Results and Discussion

In our study, we have implemented SVR, RF regressor, and KNN that are the three machine learning models to forecast the forex rate of the next 10 days on the basis of past observations. A comparative study and analysis of these three models with the help of performance metric like RMSE, R^2 Score, MSE, MAE, Explained variance, Mean Gamma Deviance (MGD) and Mean Poisson Deviance (MPD) with visualization of each model have been conducted and discussed.

Fig. 6 Comparison of original close FX price versus predicted close FX price (SVR model)

One of the machine learning techniques is the SVR algorithm, which totally depends upon regularizing model parameters [10]. As a result, the best choice of the SVR parameters, C and ϵ, has a significant impact on the performance of SVR model. For regularizing the SVR models, grid searches and the cross-validation approach [11] showed the optimized parameters that provided low RMSE and R^2 values.

The model is trained and tested on USD/INR dataset, and SVR provides the best results. Results are plotted against predicted closed forex price versus actual closed forex price as shown in following Fig. 6.

In Table 1, the results of SVR are compared with various parameter like kernel, capacity, Gamma, Epsilon. In all of this comparison, we treat gamma as constant. And we evaluate SVR models different kernels say Linear, RBF, Polynomial, and Sigmoid with different performance metric. Considerably, different RMSE, R^2, and other metric obtained from linear, RBF and polynomial, sigmoid kernels. By changing values of respective parameters, the optimal hyperplane for each of the four kernel types was fitted with high values for C and ϵ [13].

Compared to RBF kernel, linear kernel is slightly showing better result than RBF. Because when we see at RMSE of both the kernel, RMSE of linear kernel is 0.2379 and RMSE of RBF kernel is 0.2635 which is higher than linear kernel. The results are obtained, as stated earlier with the optimal hyperplane for each of the four kernel types were fitted with high values for C and ϵ where ϵ value of RBF is 0.001 which is less than ϵ value of linear kernel. The SVR model with linear kernel at $C = 1e2$ and $\epsilon = 0.01$ gives RMSE 0.2379 and R^2 0.9508 and other metric performance also getting better results as compared to RBF, Polynomial, Sigmoid kernels (Fig. 7).

Table 1 SVR model performance assessment with various SVR parameter

Kernel	Capacity (C)	Gamma	Epsilon (ϵ)	RMSE	R^2	MSE	MAE	MGD	MPD	Explained variance
Linear	1e2	0.1	0.01	0.2379	0.9508	0.0566	0.1758	1.02e−05	0.0007	0.9511
RBF	1.0	0.1	0.001	0.2635	0.9397	0.0694	0.1968	1.2500e−05	0.0009	0.9408
Polynomial	0.5	0.1	0.0001	0.8217	0.4144	0.6753	0.6700	0.0001	0.0090	0.4477
Sigmoid	0.1	0.1	0.1	9.69966	80.5729	94.0835	8.8028	0.02125	1.4120	3.3863

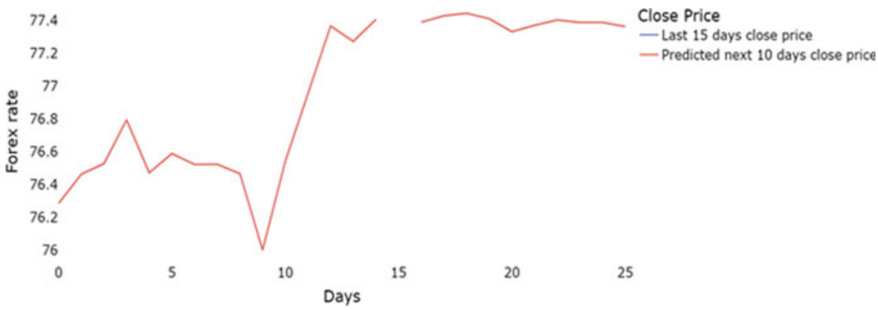


Fig. 7 Comparison of past 15 day's versus next 10 days (SVR)

3.1 Random Forest Regression

The effectiveness of RF deployment depends on figuring out how many trees and predictors each node needs to produce a decent response [5, 9]. In Fig. 8, for instance implementation closing forex price estimation, the graph shows actual closing forex prices against predicted closing forex prices over training and testing data.

Table 2 shows the RF regressor model performance evaluation using number of trees for estimation of USD/INR forex rate. Here, in this table we have compared various metric performance. While executing model, we treat various parameters of RF regressor as default except number of trees and treating random state as 0 because we are forecasting forex rate based on past data. While going through Table 4, we clearly see that as we increase number of trees of RF our metrics showing best result. When we take number of trees 10, RF R^2 is 0.8811, RMSE is 0.3718 and explained variance is 0.8923 as we know explained variance close to 1.0 is considered as good. And also MPD and MGD values close to 0.0 is considered as good, we see that MGD and MPD values for 10 trees shows close to zero. But when we increase number of trees, our model performance is increasing. The results for all

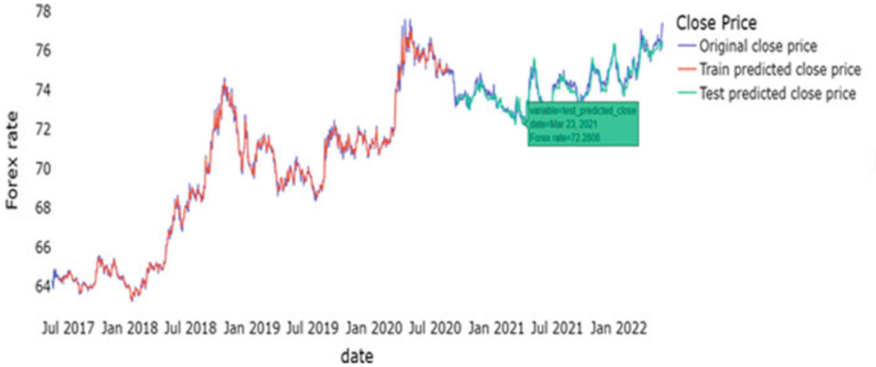


Fig. 8 Comparison of original close FX price versus predicted close FX price (RF model)

Table 2 Performance assessment of RF model using number of trees

Number of trees	R^2	RMSE	MSE	MAE	Explained variance	MGD	MPD
10	0.8811	0.37018	0.1303	0.2964	0.8923	2.4812e−05	0.0018
50	0.9147	0.3135	0.0982	0.2475	0.9195	1.7754e−05	0.00316
70	0.9175	0.3082	0.0950	0.2470	0.9219	1.7169e−05	0.00127
100	0.9141	0.3146	0.0989	0.2453	0.9193	1.7885e−05	0.00134
150	0.9131	0.3164	0.1001	0.2417	0.9188	1.8097e−05	0.00136

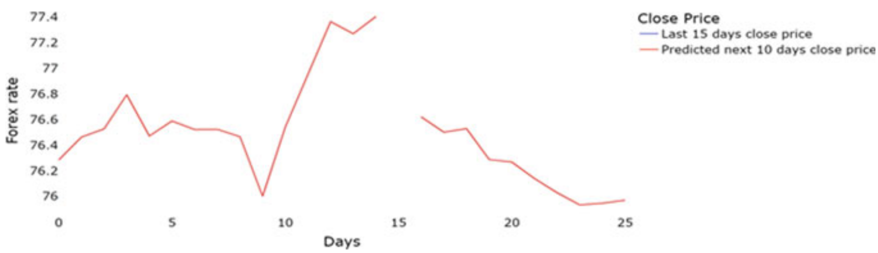


Fig. 9 Comparison of past 15 day's versus next 10 days (RF)

the metrics are significantly better with the model having 70 trees. Here, we clearly see that RMSE is 0.3082, R^2 is 0.9175, and MSE and MAE are 0.0950 and 0.2470, respectively. Explained variance, MGD, MPD scores are 0.9219, 1.7169e−05, and 0.00127, respectively. All these metrics for RF having number of optimal tree 70 show best results as compared to other optimal trees. Our RF regressor with 70 trees is giving best result. By referring to Table 2, when we increase number of trees to 100 and 150, results of RSME, R^2, and other metric were good but compared to performance of 70 trees is less (Fig. 9).

3.2 K Nearest Neighbor (KNN)

An accurate performance of KNN in the forecasting of forex rate of USD/INR generally depends on values of optimal k and distance measure. Here, distance measure is considered as Euclidean distance and checking model performance with the variation in k. In Table 3, we evaluate our KNN model with different values of k and conduct various metric performance.

We are estimating optimal k for our model, and to check this, we increase k values slowly. Firstly, we put $k = 1$ wherein R^2 is 0.6181 which is not best fit for KNN. Also, RMSE, MSE, MAE were not good. As we increase the k value our model performance with respect to given metrics is also increasing. After some time as k value increase performance of KNN degraded. When we go through Table 4, at k

Table 3 Performance assessment of KNN model with the variation in optimal k

K	R^2	RMSE	MSE	MAE	Explained variance	MGD	MPD
1	0.6181	0.6636	0.4404	0.5526	0.6514	7.9816e−05	0.005928
9	0.7427	0.5541	0.2967	0.4468	0.7635	5.3595e−05	0.003987
19	0.7786	0.5052	0.2552	0.4039	0.7933	4.5911e−05	0.003423
31	0.7898	0.4923	0.2423	0.3862	0.7938	4.3538e−05	0.003248
39	0.7815	0.5019	0.2519	0.3932	0.7827	4.5274e−05	0.003377
49	0.7671	0.5182	0.2685	0.4078	0.7677	4.8272e−05	0.003600

Table 4 Comparison table of SVR, RF, KNN with performance metric

ML models	Optimal parameters	R^2	RMSE	MSE	MAE	Explained variance	MGD	MPD
SVR	Kernel = linear $C = 1e2$, $\epsilon = 0.01$	0.9508	0.2379	0.0566	0.1778	0.9511	1.02e−05	0.0007
RF	Number of tree = 70	0.9175	0.3082	0.0950	0.2470	0.9219	1.7169e−05	0.00127
KNN	$K = 31$	0.7898	0.4923	0.2423	0.3862	0.7938	4.3538e−05	0.003248

= 31 KNN model performance is at best compared to other k values performance. At $k = 31$, our RMSE value is 0.4923 which is lowest RMSE as compared to other RSME at different values of k. In Fig. 8, we compare RMSE against optimal values of K. As we change values of k, our RMSE score is decreasing, and when value of k 31 onwards, then RMSE is increased. This graph gives the optimal k values which is 31, where RMSE is low. Now looking at table when k is 39, R^2 score is 0.7815 which is lower than R^2 score at $k = 31$, and RMSE, MSE, MAE scores are increasing. Explained variance is 0.7827, and it is decreased. MGD and MPD score is also increasing which is not good for the model. The same thing happens with $k = 49$, all of the metric gives lower performance than $k = 31$. So we conclude that $k = 31$ is an optimal k value for our KNN model.

In Fig. 10, for instance implementation closing forex price estimation with optimal k value, the graph shows actual closing forex prices against predicted closing forex prices over training and testing data.

Let's see the 10 days forecasted forex rate results of KNN model with optimal number of value. As we can see that in Fig. 10, we plotted next 10 days predicted close forex rate and we clearly observed that our KNN model with $K = 31$ is performing best.

And in Fig. 11, we plotted all of the three machine learning model (SVR, RF, KNN) which are predicting close forex rate of USD/INR (Figs. 12 and 13).

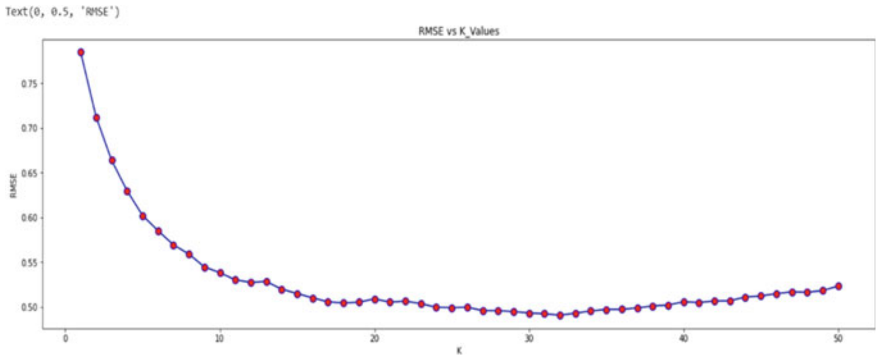


Fig. 10 Comparison of RMSE against values of K for optimal K

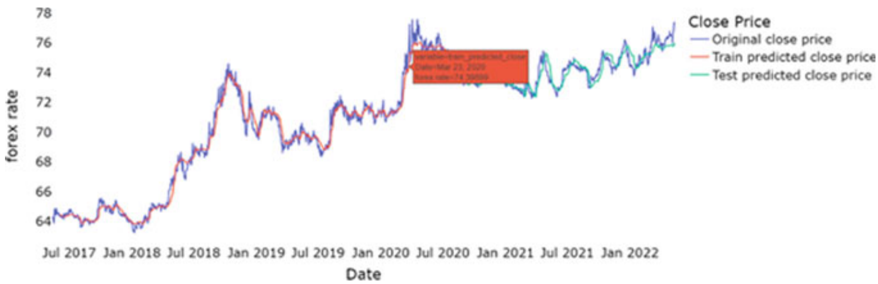


Fig. 11 Comparison of original close FX price versus predicted close FX price (KNN model)

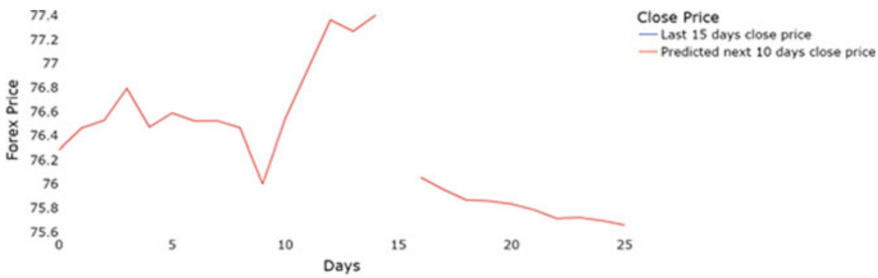


Fig. 12 Comparison of past 15 day's versus next 10 days (KNN)

4 Conclusion and Future Scope

The machine learning techniques (SVR, RF regressor, and KNN) were compared for forecasting next 10 days closing price of USD/INR forex rate. Prediction performance of these models depends on transformation and scaling that we have done during data pre-processing. The main aspect of the proposed study is the machine learning

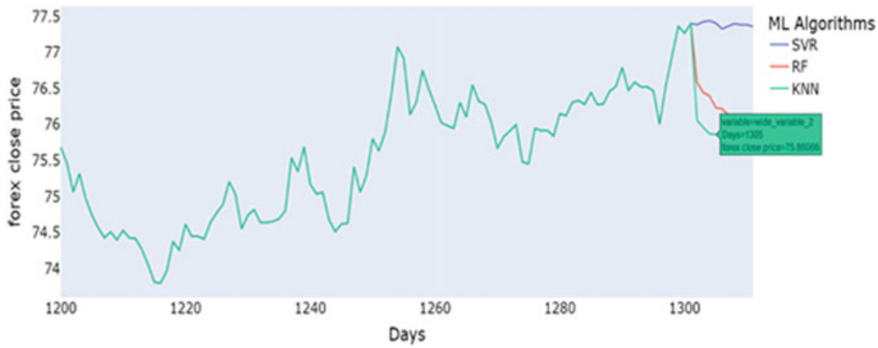


Fig. 13 Plot of all model with next 10 days prediction

model's ability to forecast rates as precisely as actual values when given both train and test data.

In this study, all of three machine learning models say SVR, RF regressor, and KNN with their optimal parameters were compared with performance metrics as given in Table 4.

From table we can clearly see the results of the SVR model against RF and KNN were best. R^2 score of SVR with optimal parameters like linear kernel, $C = 1e2$, $\epsilon = 0.01$ is 0.9508, and R^2 score of RF having 70 trees and as that of KNN with $k = 31$ is 0.9175 and 0.7898, respectively. As we discussed earlier, R^2 score ranges from 0 to 1, and R^2 score close to 0 says that our model is best fitted. In addition, RMSE of SVR is 0.2379 and that of RF and KNN are 0.3082 and 0.4923, respectively. The RMSE of SVR is lower than RF and KNN. While looking at other metric performances of each model, SVR defeated RF and KNN. Therefore, we conclude that among all machine learning model, SVR is better for predicting the forex rate.

This investigation utilized a time-consuming cross-validation method to find the best parameter values. We will look for a simple approach to get optimal parameters in further work. Also in this work, we predicted the forex rate using only one variable, in the future, we can work with multivariate forex time series. In addition, we will be looking for how deep learning techniques will affect forex prediction.

References

1. Ban T, Zhang R, Pang S, Sarrafzadeh A, Inoue D (2013) Referential kNN regression for financial time series forecasting. National Institute of Information and Communications Technology
2. Bou-Hamad I, Jamali I (2020) Forecasting financial time-series using data mining models: a simulation study. Res Int Bus Financ 51:101072
3. Breiman L (2001) Random forests. Kluwer Academic Publishers
4. de Almeida BJ, Neves RF, Horta N (2018) Combining support vector machine with genetic algorithms to optimize investments in forex markets with high leverage. Appl Soft Comput 64:596–613

5. Goehry B, Yan H, Goude Y, Massart P, Poggi JM (2021) Random forests for time series
6. Ince H, Trafalis TB (2006) A hybrid model for exchange rate prediction. Decis Support Syst 42(2):1054–1062
7. Kumar V (2016) Prediction of foreign exchange rate using data mining ensemble method
8. Martínez F, Frías MP, Pérez MD, Rivera AJ (2017) A methodology for applying k-nearest neighbor to time series forecasting
9. Nielsen LG (2017–18) Machine learning for foreign exchange rate forecasting. Scopus
10. Ojemakinde BT (2006) TRACE: Tennessee Research and Creative Exchange
11. Petropoulos A, Chatzis SP, Siakoulis V, Vlachogiannakis N (2017) A stacked generalization system for automated FOREX portfolio trading. Expert Syst Appl 90:290–302
12. Pevekar S (2021) Forecasting INR exchange rate against USD, GBP, JPY, SGD, EUR, AED using machine learning. EPRA Int J Res Dev (IJRD) 6(7)
13. Raimundo MS, Okamoto J Jr (2018) SVR-wavelet adaptive model for forecasting financial time series. In: International conference on information and computer technologies. IEEE
14. Vapnik V, Cortes C (1995) Support-vector networks. Kluwer Academic Publishers
15. Wang Y, Xia ST, Tang Q, Wu J, Zhu X (2017) A novel consistent random forest framework: bernoulli random forests. IEEE Trans Neural Netw Learn Syst 29(8):3510–3523
16. Wei W (2020) Multi-channel LSTM with different time scales for foreign exchange rate prediction. In: AISS ‘19. Association for Computing Machinery
17. Yaohao P, Albuquerque PHM (2019) Non-linear interactions and exchange rate prediction: empirical evidence using support vector regression. Appl Math Financ 26(1):69–100

Smart Bus Real-Time Tracking System Using GSM and GPS Module

Kapil Mundada, Sumedh Patti, Tejas Rajguru, Puskraj Savji, and Sayali Shambharkar

Abstract Real-time information regarding buses is made available to passengers via the GPS and GSM-based Real-Time Bus Monitoring system, which aims to enhance public transportation. The system uses GPS and GSM technologies to track the location and estimated arrival time of buses and transmit this information to commuters' mobile devices. This technology addresses the common problem of long waiting times at bus stops and the uncertainty of bus arrival times. The system's benefits are numerous, including the ability to reduce waiting times, optimize bus routes, and reduce traffic congestion. Commuters can track the arrival times of their buses and make informed decisions about their travel plans, which can save them time and effort. The system can also assist transportation authorities in making decisions about bus routes and improving the overall efficiency of public transportation. Moreover, the system's effectiveness can be enhanced through the use of data analytics and artificial intelligence techniques. By analyzing data on bus ridership and travel patterns, transportation authorities can identify areas of high demand and adjust bus schedules and routes accordingly. This approach can help reduce overcrowding on buses, improve travel times, and enhance the overall user experience.

Keywords Bus tracking · GSM · Google maps · GPS

K. Mundada (✉) · S. Patti · T. Rajguru · P. Savji · S. Shambharkar
Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, India
e-mail: kapil.mundada@vit.edu

S. Patti
e-mail: sumedh.patti21@vit.edu

T. Rajguru
e-mail: tejas.rajguru21@vit.edu

P. Savji
e-mail: pushkraj.savji21@vit.edu

S. Shambharkar
e-mail: sayali.shambharkar21@vit.edu

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_46

1 Introduction

With time becoming a valuable asset, an increasing number of people are relying on public transportation, leading to crowded bus stops. In these situations, passengers frequently wait in lengthy lines and miss their intended bus, while the next bus arrives at a point distance from where they are already waiting. Providing commuters with real-time bus tracking and arrival information based on their location can help them make informed decisions about whether to wait for the bus or look for alternative transportation. By allowing commuters to track the location and anticipated arrival time of the bus using a mobile device, the Real-Time Bus Monitoring system provides convenience for users.

Users may more effectively manage their time with the aid of the system, which shows them alternate bus routes to the same destination as well as the estimated time it will take them to arrive. This information allows users to decide whether to wait for the bus or explore other options, such as walking to the next stop, taking another bus, or hailing a taxi. Additionally, if there is only one bus running between the user's present position and their destination, the program displays the anticipated arrival time of the bus, saving the user from needlessly waiting at the bus stop.

Automatic vehicle location (AVL) systems, which are based on the global positioning system (GPS), have recently made advancements that have given transit sectors and public transport businesses capabilities to monitor and regulate the operations of their fleets efficiently and affordably. AVL technology allows bus operators to track the location and movement of their vehicles in real time, reducing the likelihood of bus bunching or unexpected delays. Furthermore, AVL systems can generate data on bus operations and performance metrics, which can be used to optimize route planning and improve service quality.

The Real-Time Bus Monitoring system offers a practical solution to address the challenges faced by passengers waiting at bus stops. The system enables users to make informed decisions, improving their entire commuting experience, by giving users real-time bus tracking, estimated arrival times, and alternative transportation options. The application of AVL technology in public transportation can lead to better operational efficiency and service quality, benefiting both passengers and transport operators.

2 Literature Review

For the literature review purpose, we have referred to the following papers:

1. **An Experimental Study on Real-Time Bus Arrival Time Prediction with GPS Data**

This study looks into how AVL techniques can be used to notify people about bus arrival times in rural places in real time. The study, conducted in Blacksburg, Virginia, focuses on developing arrival time estimation algorithms using GPS-based bus location data and other relevant information such as scheduled arrival time and delay correlation. Four algorithms are presented, and the study found that the dwell time at time-check stops is the most critical information for algorithm performance. The paper compares the algorithms' performance against three criteria: overall precision, robustness, and stability. The results indicate that incorporating additional information can improve algorithm accuracy and make the public transit system more user-friendly. The study highlights the importance of identifying delay patterns and data quality issues when using GPS data to estimate bus arrival times. The paper demonstrates the potential of AVL techniques in improving the competitiveness of public transit systems in rural areas through the provision of real-time bus arrival information. The findings provide valuable insights for the development of more robust and accurate algorithms for estimating bus arrival times in rural areas [1].

2. **Simple Bus Tracking System**

This paper presents a simple and low-cost vehicle tracking system designed for a small, closed-user community. Specifically, the system was developed to provide real-time tracking of shuttle buses that transport students between two campuses of UCSI University in Kuala Lumpur, Malaysia. By using WAN technology instead of GPS, the system was designed to fully utilize available resources and reduce operational costs. The system provides information about bus arrival and departure times at each stop along the fixed route, thus reducing wait times for students. The paper describes the system design approach, components, software considerations, message information, testing, and results. The proposed system can be implemented easily and at a low cost, making it a viable solution for similar small-scale transportation systems [2].

3. **Smart Bus Tracking and Management System Using IoT**

In order to track the whereabouts of buses and give students and staff updates in real time on their SMART phones, the Bus Tracking system is a software application that makes use of IoT and GPS technology. The installation of an electronic gadget with GPS tracking devices inside a bus is required for this system, which will periodically update a central server with the bus's whereabouts. The server then stores the data in a database and sends it to the client application to display the position of buses on a map. The system reduces waiting time for students and staff, and they can plan their travel more efficiently using the app. The Bus Tracking system uses an Arduino UNO microcontroller for software and hardware programming, and the application can be easily extended for a central tracking system to manage all buses efficiently. The key feature of the Bus Tracking system is its simple mode of communication, making it accessible and user-friendly for all. This paper describes the design, implementation, and testing of the Bus Tracking system using low-cost and mature system components. Overall, the Bus Tracking system provides an effective and cost-efficient solution for bus tracking and route management, making it a useful tool for both students and staff [3, 4].

4. **Real-Time Bus Tracker Application to Reduce Demand Supply of City Buses**

This paper discusses the issues faced by the public transportation system, particularly the bus service, in busy cities and towns due to the lack of a real-time tracking system. The proposed system aims to overcome these challenges by providing an Android application that offers up-to-date information about all buses traveling in Nagpur. The use of Android software is ideal since it is widely owned and user-friendly, and the number of applications made for the software is increasing significantly. The paper highlights the constraints that need to be met to ensure the efficient supply and demand of buses. The application was developed using Android Studio 3.3.2, ADT, and SDK. The proposed system is expected to enhance the user experience of the bus service and provide the bus supplier with valuable information to improve their operations [5].

5. **Online Bus Tracking and Ticketing System**

The efficient scheduling and tracking of public transportation are crucial for ensuring that commuters can make the most of their time. To meet this need, they have developed an Android application that allows users to track buses using GPS technology. Additionally, the app features an e-wallet for cashless transactions and a virtual bus pass with a QR code for easy verification by bus organization employees. The app also includes a Bluetooth printer for on-time ticketing, and all data is stored on a Firebase cloud, providing valuable insights for the future optimization of public transportation systems. Their solution aims to improve the efficiency and convenience of public transportation for both commuters and bus organizations [6, 7].

3 Tools Used

3.1 Hardware

1. Arduino Uno

The ATmega328P microprocessor serves as the foundation for the Arduino Uno microcontroller board. It contains a USB port, a power jack, an ICSP header, six analog inputs, fourteen digital input/output pins, a sixteen MHz quartz crystal, and a reset button. The board may be powered externally or via USB, and it can be programmed by using Arduino Software using a wiring-based programming language.

2. GSM SIM800L

GSM SIM800L is a small and low-power quad-band GSM/GPRS module that enables wireless communication capabilities for embedded systems. The module supports TCP/IP, HTTP, FTP, POP3/SMTP, and other communication protocols. It runs on frequencies of 850/900/1800/1900 MHz. It has a tiny form size and is simple to integrate with embedded systems or microcontrollers like Arduino or Raspberry Pi. The module may be managed with AT commands and needs an external antenna and power source.

3. GPS Neo 6

GPS Neo 6 is a Global Navigation Satellite System (GNSS) receiver module that is commonly used in various embedded applications such as vehicle tracking, UAV navigation, and handheld devices. It is an affordable, effective GPS module. It has a high degree of sensitivity, precision, and a quick time to first fix (TTFF). The module outputs the NMEA data in ASCII format and interfaces with the host microcontroller through a serial interface (UART). It is simple to connect the module to well-known development boards like Arduino and Raspberry Pi. It can be operated with only a few external components and can be powered by a 3.3 V or 5 V supply.

3.2 Software

1. HTML

Hypertext Markup Language is what HTML stands for. It is a common markup language used to create websites and web applications. A website's structure and content are determined by it, which also acts as the Internet's core.

2. CSS

An acronym for CSS stands for Cascading Style Sheets. It's a language that is used to make style sheets that specify how a document will look when displayed in a markup language like HTML. Web pages are styled and laid out using CSS to make them more appealing and responsive.

3. JavaScript

The programming language JavaScript is primarily employed to develop dynamic and interactive web content. It is a high-level, interpreted language that can be used both on the server and the client (in a web browser) (using Node.js). Web page behavior and interactivity can be added with JavaScript, including user input validation, dynamic content updates, and animations.

4. PHP

PHP is a well-liked server-side scripting language for building websites (Hypertext Preprocessor). Developing dynamic web pages and web apps is its main usage. On the server, PHP code is run, creating HTML that is transmitted to the client's browser. PHP is capable of working with databases, controlling server-side files, sending and receiving cookies, creating and altering cookies, and much more.

5. SQL

Relational database management is done using the programming language Structured Query Language (SQL). By employing different statements and instructions, databases can be created, modified, and queried. In order to communicate with databases, which are frequently used to store data, SQL is used. The ability to filter, sort, and group data as well as carry out calculations and transformations on it are just a few of the strong features that SQL offers for handling data. Relational database management is done using the programming language Structured Query Language (SQL). By employing different statements and instructions, databases can be created, modified, and queried. In order to communicate with databases, which are frequently used to store data, SQL is used. The ability to filter, sort, and group data as well as carry out calculations and transformations on it are just a few of the strong features that SQL offers for handling data.

6. Firebase

Google developed Firebase, a cloud-based platform that offers a number of capabilities for creating web and mobile applications. A real-time database, hosting, storage, authentication, cloud messaging, and other services are all part of Firebase. The real-time database, a NoSQL cloud-hosted database, synchronizes data in real time with all connected clients and stores data in JSON format. Building real-time apps can benefit from it.

Users can easily be verified with Firebase Authentication with their email address and password, phone number, Google, Facebook, Twitter, and other well-known identity sources. Its serverless architecture does away with the need to manage infrastructure, and its extensive feature set makes it simple for developers to create robust apps. Furthermore, Firebase offers a sizable free tier, making it available to developers and small businesses.

4 Methodology

A crucial component of the GPS and GSM-based Real-Time Bus Monitoring system is calculating the distance between users and bus stops. This system is designed to track the location of the user using the integrated GPS receiver, and the user provides

Fig. 1 Bus tracking system flowchart

the system with the source name, destination name, and bus number. Then, the system uses this information to identify the user's current location accurately (Fig. 1).

The GPS fitted on the bus gets the coordinates of the bus and then it is transmitted to the server/database with the use of GSM SIM800l. Now the system has the coordinates of the bus and the user's source location, then it determines the route between the two locations and estimates the time of arrival.

The user interface shows a Google Maps-based view that gives a visual depiction of the user's location and the bus's current location. The map provides basic control functions, such as zooming, panning, and mode changing, to enable the user to navigate the map easily. This information can help users plan their journey and make informed decisions about which bus stop to go to based on their destination.

To effectively improve public transportation, the system must be able to precisely calculate the distance between users and bus stops. With this information, commuters can make informed decisions about their travel plans, reducing the amount of time spent waiting at bus stops and increasing the efficiency of the public transportation system.

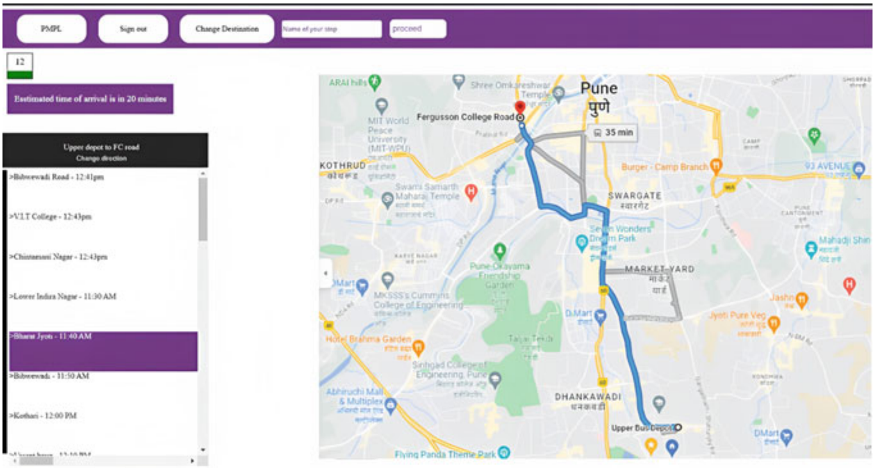


Fig. 2 Live location of the bus

The system makes public transit more effective, dependable, and user-friendly by combining GPS technology, message transmission, and Google Maps to give passengers real-time information on bus whereabouts, expected arrival times, and bus station specifics.

The following standard features are present in this application:

- A GPS-based tool that measures the separation between a bus and a user's bus stops.
- When the bus should arrive at that bus stop can be determined by a timing device.
- The prospective bus stops can be displayed on a map interface.
- A bus route advisor can provide consumers with options.
- A real-time map marker.

5 Results and Discussions

The visual representations presented above showcase the outcomes derived from the testing of our initial model. The illustrations demonstrate that the user has the ability to select their desired route and source location, while the system calculates the distance between the bus and the user's chosen starting bus stop. This information is then presented to the user through an interactive interface that utilizes the Google Maps API. The interface not only displays the distance but also estimates the time it will take for the bus to arrive at the user's chosen starting location (Figs. 2 and 3).

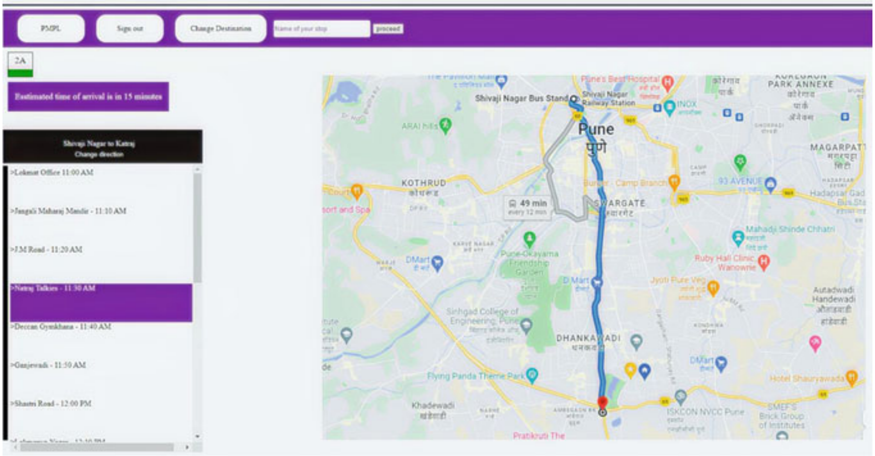


Fig. 3 Live location of the bus

6 Conclusion

The paper discusses the features and benefits of GPS and GSM-based Real-Time Bus Monitoring systems. These systems track the present location of buses and predict when they will arrive at certain stops using GPS technology. Passengers receive the information via text messages or other channels when the estimations are updated in real time as the buses give location updates. The Real-Time Bus Monitoring system has advantages for drivers and administrators of transportation as well as for passengers. By receiving exact information on the bus's location right now and how long it will take to get to their destination, the technology helps passengers save time. Additionally, passengers can also receive information on the occupancy of the bus, which can help them make informed decisions about whether to wait for a less crowded bus or board the current one.

This system provides many benefits to the transport industry as well. By monitoring the real-time location of buses, transport administrators can ensure that buses are on schedule and make adjustments to routes and schedules as necessary. The system can also help reduce fuel consumption and operating costs by optimizing routes and reducing idle time.

In conclusion, GPS and GSM-based Real-Time Bus Monitoring systems offer many advantages for both passengers and transport administrators. They provide real-time tracking and accurate arrival estimates, which can help passengers save time and make informed decisions. Additionally, transport administrators can use this system to optimize routes and schedules, leading to better overall transport management. Real-Time Bus Monitoring is now possible thanks to the development of GPS and cellphone networks, and it has the potential to completely change how we perceive public transportation.

Acknowledgements We express our gratitude to our supervisor, Prof. Kapil Mundada, for providing us with his invaluable guidance and critical evaluation throughout the duration of this project. Additionally, we extend our appreciation to the Department of Instrumentation Engineering at VIT Pune, for their generous support and provision of essential resources that enabled us to successfully undertake this endeavor.

References

1. Lin W-H, Zeng J (1999) Experimental study on real-time bus arrival time prediction with GPS data. Transp Res Rec J Transp Res Board (1666):1019
2. Lau EC-W (2013) Simple bus tracking system. J Adv Comput Sci Technol Res 3(1)
3. Shende P, Bhosale P, Khan S, Patil P (2016) Bus tracking and transportation safety using internet of things. Int Res J Eng Technol (IRJET) 03(02)
4. Selvapriya PR, Mundada MR (2015) IOT based bus transport system in Bangalore. Int J Eng Tech Res (IJETR) 3(2). ISSN: 2321-0869
5. Kumbhar M, Survase M, Mastud P, Salunke A (2016) Real time web based bus tracking system. Int Res J Eng Technol (IRJET) 03(02)
6. Nair V, Pawar A, Tidke DL, Pagar V, Wani N (2018) Online bus tracking and ticketing system. MVP J Eng Sci 1(1). http://doi.org/10.18311/mvpjes/2018/v1i1/18297
7. Chandurkar S, Mugade S, Sinha S, Borkar P (2013) Implementation of real-time bus monitoring and passenger information system. Int J Sci Res Publ 3(5)

Probing the Role of Information and Communication Technology (ICT) in Enhancing Research: An Epilogue of Accessible Research Tools

M. Vinay and J. Jayapriya

Abstract Information and Communication Technology (ICT) has revolutionized the way researchers conduct their work. It has enabled them to access a wealth of information through online databases, collaborate with colleagues across the globe, and analyze vast amounts of data quickly and accurately. This paper explores the role of ICT in enhancing research tools, highlighting the benefits it provides to researchers in terms of increased efficiency, improved accuracy, and greater access to resources. It also discusses some of the challenges associated with using ICT in research, such as data security and privacy concerns, and offers potential solutions. Overall, the paper concludes that ICT is an essential tool for researchers and will continue to play an increasingly important role in advancing scientific knowledge and innovation.

Keywords ICT · Research · Tools · Search engine · Literature review tool · Database tool

1 Introduction

ICT stands for Information and Communication Technology. It is a term used to describe the convergence of various technologies used for information processing and communication. These technologies include computers, software, telecommunications equipment, and the Internet. It has revolutionized the way we communicate and share information. It has enabled people to connect with each other globally and has made access to information easier and faster. It has also brought about new opportunities for business, education, health care, and entertainment. It encompasses a wide range of applications, such as e-mail, instant messaging, social media, video conferencing, e-commerce, online education, and telemedicine. It has become an essential tool for both individuals and organizations to operate efficiently and effec-

M. Vinay (✉) · J. Jayapriya
Christ University, Bangalore, Karnataka, India
e-mail: vinay.m@christuniversity.in

J. Jayapriya
e-mail: jayapriya.j@christuniversity.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_47

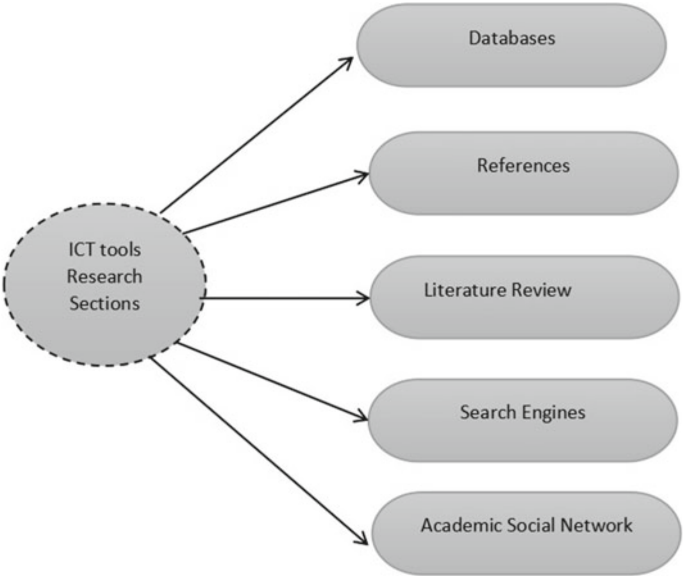


Fig. 1 ICT tools for research sections

tively in the modern world. ICT has significantly impacted the field of research, providing researchers with advanced tools to conduct research, process and analyze data, and communicate their findings with the scientific community and the public. One of the primary benefits of ICT in research is the ease of accessing and sharing information. With the Internet and online databases, researchers easily access a vast array of scientific literature, research reports, and datasets from all over the world. This has made it easier for researchers to stay updated on the latest research findings and to collaborate with colleagues worldwide. ICT has also enabled researchers to collect, process, and analyze large amounts of data more efficiently and accurately. Advanced software tools and algorithms allow researchers to process large datasets, identify patterns and correlations, and make more precise and reliable conclusions. Moreover, ICT has made it easier for researchers to communicate their findings with the scientific community and the public. With online publishing platforms and social media, researchers quickly and easily share their research with a global audience, receive feedback, and engage in discussions with other researchers (Fig. 1).

The objective of this study is to:

1. Data collection and storage: ICT tools enable researchers to collect and store vast amounts of data from various sources, including online surveys, social media, and other digital platforms.
2. Data analysis: With ICT tools, researchers perform complex data analysis, such as statistical analysis, data mining, and machine learning, to identify patterns and insights that would be impossible to detect manually.

3. Collaboration: ICT tools facilitate collaboration among researchers by allowing them to share data, communicate, and work together in real time, regardless of their geographical location.
4. Dissemination: ICT tools enable researchers to disseminate their findings widely and quickly through digital channels, such as websites, social media, and online journals. Efficiency: By automating many research processes, ICT tools help researchers save time, reduce costs, and improve the overall efficiency of their research.

The paper's structure is as follows: Sect. 1 presents the Introduction, followed by a depiction of previous research work in Sect. 2. Section 3 outlines the proposed study, which explores various ICT research tools for different purposes. In Sect. 4, the advantages and disadvantages of these tools are discussed. Lastly, Sect. 5 concludes the paper.

2 Background

Information and Communication Technology (ICT) tools have significantly impacted the research landscape. These tools have revolutionized the way researchers collect, store, analyze, and present data. They have enabled researchers to access large amounts of data, collaborate with other researchers globally, and disseminate their findings to a wider audience. The development of ICT tools for research dates back to the early 1960s when computers were first used for data processing. Since then, the ICT industry has continuously evolved, leading to the development of new tools and technologies that have transformed research practices. A significant amount of literature is available on ICT tools for research. Abdulkarim and Hasan [1] the book provides an overview of different research methods in education, including the use of ICT tools for data collection, analysis, and presentation. Jonassen et al. [5] the book covers a wide range of topics related to educational technology, including the use of ICT tools for research. Macfarlane [7] discusses the ethical considerations involved in using ICT tools for research, such as issues related to data privacy and security. Fisher et al. [3] aimed at teachers, it provides useful information on the use of ICT tools for research purposes in the context of education. Fricker [4] provides an overview of different online research methods, including the use of ICT tools for data collection and analysis. Due to limited empirical data on the necessary ICT tools for achieving specific goals in Laos, the present study aims to address this gap by examining the opinions of Laotian academics regarding the required ICT tools and platforms, their current level of ICT usage, and necessary ICT training. The study collected quantitative data through a survey of 11 participants who are actively involved in the ERASMUS project for Building Research Building Social Research Capacities in Higher Education Institutions in Lao PDR and Malaysia (BRECIL) [1, 10]. The analysis revealed that academics require tools such as SPSS, Survey Monkey, Google Scholar, and Grammarly to support them in their research. Additionally,

Fig. 2 Related work of ICT tools in education

they require platforms for plagiarism and grammar checking, dissemination and networking, and analysis and interpretation. Most respondents expressed the need for training on how to use ICT tools to analyze research data. These findings help educational decision-makers develop strategies for introducing ICT tools and organizing training for Laotian academics on research and development [12, 13]. Siddiqui's educational technology information bank defines online learning environments as virtual spaces that allow real time, collaborative interaction between teachers and students, along with access to various asynchronous learning materials. This type of learning environment also promotes a fundamental comprehension of computer-based learning. Online instruction leverages technology and the Internet to enable teaching and learning between educators and learners. The study aims to investigate the perspectives and difficulties that teachers and students encounter when utilizing ICT tools [2, 6, 8, 9, 11] (Fig. 2).

3 Proposed Study

There are various research tools and Information and Communication Technology (ICT) resources available to assist with title research. Online databases like JSTOR, Google Scholar, and Pubmed provide access to millions of scholarly articles, research papers, and other publications that are used to gather information related to a specific

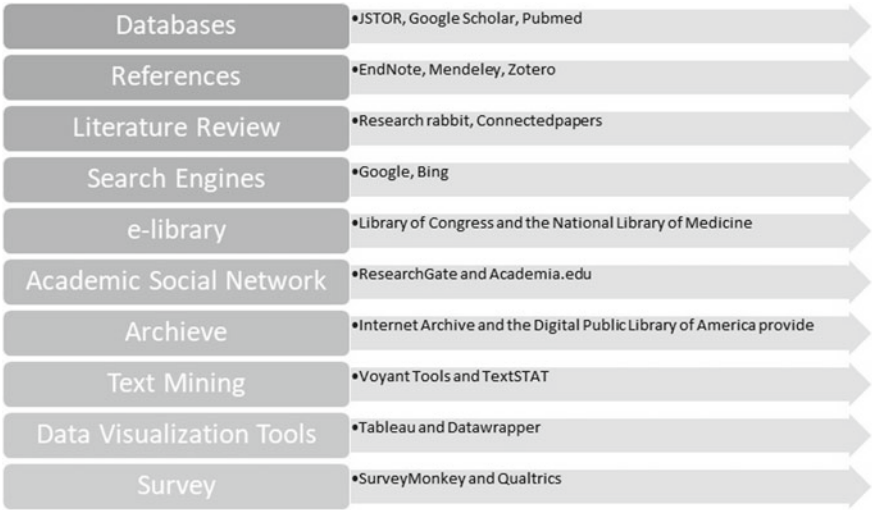


Fig. 3 Online tools for different tasks in research

title. Reference managers like EndNote, Mendeley, and Zotero used to manage references and citations while conducting research. These tools help to organize research materials, create bibliographies, and keep track of research sources. Search engines like Google and Bing are used to find relevant information on a particular title. They provide access to a vast range of information sources, including news articles, blogs, and social media posts. Electronic libraries like the Library of Congress and the National Library of Medicine provide access to vast collections of books, journals, and other materials that are used to conduct research. Academic social networks like ResearchGate and Academia.edu used to connect with other researchers and access research papers, articles, and other publications. Online archives like the Internet Archive and the Digital Public Library of America provide access to historical documents, images, and other materials that are used to gather information related to a particular title. Text mining tools like Voyant Tools and TextSTAT used to analyze and visualize large amounts of text data. These tools help to identify trends, patterns, and relationships in the data. Data visualization tools like Tableau and Datawrapper used to create visual representations of data, making it easier to understand and interpret. Online survey tools like SurveyMonkey and Qualtrics used to gather data from a large number of people. These tools are useful for conducting market research or collecting feedback from customers.

Figure 3 shows the basic sections and its respective ICT tools used for research. ICT database tools are commonly used in research to gather, organize, and analyze data. Excel is a widely used spreadsheet software that is used for organizing and analyzing data. It allows researchers to create tables, graphs, and charts that are used to visualize data. Access is a database management system that is often used to store and manage large amounts of data. Researchers use Access to create tables, forms, and queries to organize and analyze their data. SPSS (Statistical Package for the Social Sciences) is a software program that is commonly used in social science research. It allows researchers to conduct statistical analyses and generate graphs and charts to visualize their data. SAS (Statistical Analysis System) is a software suite used for data management, analysis, and reporting. It is widely used in both academic and commercial research. NVivo is a qualitative data analysis software that allows researchers to organize, analyze, and visualize unstructured data, such as text and audio. Tableau is a data visualization software that allows researchers to create interactive visualizations and dashboards to communicate their findings. Google Forms is a free online tool that allows researchers to create surveys and collect data. It is widely used in academic and market research. Literature review-based tool is Connected Papers. It works by analyzing citation data from research papers, and then using this data to create a visual network of related papers. Search for a paper or author, and Connected Papers will generate a network graph of related papers based on their citation relationships. Explore the network graph to identify important papers, key authors, and relevant topics within the literature. Some of the features of Connected Papers include:

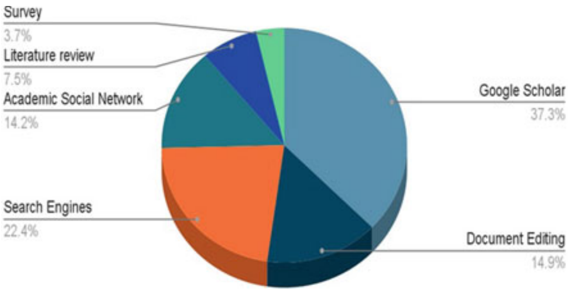
1. Interactive network graphs: Connected Papers allows you to explore the citation network graph in an interactive way. Zoom in and out, move nodes around, and click on individual papers to see their details.
2. Related papers: Connected Papers identifies related papers based on their citation relationships, and allow you to explore the network to discover new research.
3. Key papers: Connected Papers highlights key papers within the network, based on their citation relationships and their importance within the literature.
4. Paper details: Connected Papers provide details on individual papers, including abstracts, authors, and citation counts.
5. Connected Papers is a useful tool for researchers who are conducting literature reviews and want to explore the citation relationships between papers in their field. It helps you to identify important papers, key authors, and relevant topics within the literature, and save time in the literature review process.

In this paper also, the Connected Papers tool is used for the study of the previous related work.

4 Discussion

To understand the real scenario of ICT usage, a survey was conducted for a group of researchers. Surveys are a commonly used research method to collect data from a population or a sample of individuals. The purpose of a survey in research is to gather information about a particular topic or phenomenon of interest from a large number of people. Surveys are used to collect quantitative data (numerical data) or qualitative data (non-numerical data) depending on the type of questions asked. Some of the main purposes of using surveys in research include:

Fig. 4 Sample online tools for different tasks in research



1. Understanding attitudes and opinions: Surveys help researchers understand what people think, believe, or feel about a particular issue or topic.
2. Assessing behavior or practices: Surveys used to gather information about what people do, such as their buying habits, health behaviors, or social media use.
3. Evaluating program or product effectiveness: It is used to assess the effectiveness of a program or product by measuring outcomes or gathering feedback.
4. Generating new ideas: Surveys are used to generate new ideas or test hypotheses for further investigation.

For this survey, SurveyMonkey tool is used. SurveyMonkey is an online survey development tool that allows individuals and businesses to create and publish surveys and collect responses. It offers a user-friendly interface, a wide range of question types and design options, and various features for analyzing and sharing survey results. With this, users customize their surveys by selecting from over 100 survey templates, adding their own branding and logos, and tailoring questions to meet their specific needs. The platform also provides tools for sending out surveys via e-mail, social media, and other channels, as well as for monitoring response rates and analyzing data in real time. In addition, it offers advanced features such as custom scripting, data export, and integration with other software tools, allowing users to automate survey workflows and analyze data at scale. It also has a mobile app, making it easy to create and send surveys on the go. It is a powerful and versatile tool for conducting surveys, gathering feedback, and making data-driven decisions (Fig. 4).

5 Conclusion

This paper concludes that Information and Communication Technology (ICT) tools have greatly enhanced the research process, making it faster, easier, and more efficient. These tools provide researchers with access to vast amounts of data, allow for easy collaboration and communication, and offer a wide range of tools for data analysis and visualization. From search engines, reference managers, and data analysis software to online surveys, social media, and virtual research environments, it has become essential for researchers in all fields. However, it is important to keep in mind the limitations and ethical considerations when using ICT tools for research. Researchers must be mindful of data privacy, security, and integrity, and ensure that they are following ethical guidelines and standards. ICT tools have revolutionized the research process and will undoubtedly continue to play a crucial role in advancing knowledge and discovery in the future. This study can be extended by elaborating on all the essential ICT tools used for research. The ultimate aim of the usage of ICT tools is to reduce the time in different basic processes and give more importance to idea enhancement and implementation.

References

1. Abdulkarim SA, Hasan SR (2023) Teachers and students attitude and challenges on applications of ICT tools. QALAAI ZANIST J 8(2):1253–1265
2. Alkamel MAA, Chouthaiwale SS (2018) The use of ICT tools in English language teaching and learning: a literature review. Veda’s J Engl Lang Lit JOELL 5(2):29–33
3. Fisher A, Exley K, Ciobanu D (2014) Using technology to support learning and teaching. Routledge, London
4. Fricker RD (2016) Sampling methods for online surveys. The SAGE handbook of online research methods, pp 162–183
5. Jonassen D, Spector MJ, Driscoll M, Merrill MD, van Merrienboer J, Driscoll MP (2008) Handbook of research on educational communications and technology: a project of the association for educational communications and technology. Routledge, London
6. Lu D, Xie YN (2022) Critical thinking cultivation in TESOL with ICT tools: a systematic review. Comput Assist Lang Learn 1–21
7. Macfarlane B (2010) Researching with integrity: the ethics of academic enquiry. Routledge, London
8. Maharjan M, Dahal N, Pant BP (2022) ICTs into mathematical instructions for meaningful teaching and learning. Adv Mob Learn Educ Res 2(2):341–350
9. Morze N, Makhachashvili R, Smyrnova-Trybulska E (2016) Communication in education: ICT tools assessment. In: Proceedings from DIVAI, pp 351–354
10. Omar M, Ahmad M, Yasin A, Ibrahim H, Soulignavong L, Vongsouangtham B et al (2022) The needs of information and communication technology (ICT) tools for research development: a case in Laos. J Adv Res Appl Sci Eng Technol 28(3):49–56
11. Strutynska O, Umryk M (2017) ICT tools and trends in research, education and science: local survey, no 3, pp 150–160
12. Toma F, Ardelean A, Grădinaru C, Nedelea A, Diaconu DC (2023) Effects of ICT integration in teaching using learning activities. Sustainability 15(8):6885
13. Wirajing MAK, Nchofoung TN (2023) The role of education in modulating the effect of ICT on governance in Africa. Educ Inf Technol 1–34

Resourcing Small Indigenous Languages in the Field

Antria Panayiotou, Francis Saa-Dittoh, and Anna Bon

Abstract Small indigenous languages of the world are often spoken in regions where literacy rates are low and access to digital services is limited. Whereas voice information services could be useful for low literate speakers of these languages, written and audio data from these languages are hardly available. This paper presents a method for resourcing indigenous languages, with the aim to build targeted voice services based on an automatic speech recognition system. The proposed method for data collection has been co-developed with native speakers of an indigenous language, *dagbanli* and validated in the field, through a case study in northern Ghana.

Keywords Under-resourced indigenous languages · Small data · Low-resource environments · User-centered design · Agile development · ICT4D

1 The Importance of Resourcing Small Indigenous Languages

Indigenous languages are often spoken in regions where educational systems are lacking, and consequently, literacy rates are low. In these regions, generally, purchasing power is low and people are less connected to digital services. In many countries in the Global South, that have been colonized in the eighteenth and nineteenth centuries, the colonizer's language has been imposed as *lingua franca*. In Africa, English, French and Portuguese have been introduced and have rapidly spread through newly introduced, European educational systems. Adoption of the colonizer's language had many advantages, as a channel for communication and trade. However, this pushed away many indigenous languages, of which many have disappeared or are at the brink of extinction [1].

A. Panayiotou · A. Bon (✉)
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
e-mail: a.bon@vu.nl

F. Saa-Dittoh
University for Development Studies, Tamale, Ghana

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_48

There is a growing interest in the conservation of culture and indigenous knowledge and languages. Digitization has the potential to preserve these languages and connect large groups of still "unconnected" people. However, costs of voice services are generally high. Digital data such as text or audio from indigenous languages are not widely available.

In this paper we present a detailed case study of designing a method for data collection for small indigenous languages. By co-designing, testing, improving, it with native speakers of *dagbanli* language in northern Ghana, the method has been validated. In this paper we aim to generalize our findings and draw lessons how to resource small indigenous languages in general.

2 Previous Work

Various studies have investigated how online crowd-sourcing is used for data collection for automatic speech recognition.

Toshniwal et al. [2] present a method for multilingual speech recognition involving training of a singular end-to-end model on multiple languages. Compared to conventional models that are trained separately for each language, this method shows comparable/superior performance. The authors demonstrate that this method is computationally economical, as the model can be trained simultaneously on multiple languages without requiring additional resources.

An open-source platform named Woefzela, to collect speech data in low-resource environments was presented by De Vries et al. [3]. The platform is designed to operate on low-end Android devices and includes real-time quality control to ensure that the acquired data is of sufficient quality for training speech recognition models. The study describes how Woefzela was utilized to capture speech data for some South African languages to train speech recognition models.

According to a study by the Interaction Design Foundation,[1] prototypes play a crucial role in the design process, allowing designers to explore and test various ideas, collect feedback and make informed decisions about the design's direction. Also, according to the National Institute of Standards and Technology, prototypes are a powerful tool for evaluating and improving designs because they allow designers to find and fix problems early in the design process before it becomes more costly to do so.[2]

Due to limited Internet connectivity, crowd-sourcing platforms do not reach many people in less connected areas. To mitigate this problem, Stan [4] presented a lightweight method to resource small speech data using a machine learning (ML) algorithm. A small corpus speech recognition system was built as a proof-of-concept.

[1] https://www.interaction-design.org/literature/topics/prototyping.

[2] https://www.nist.gov/publications/prototyping-new-products-systems-and-processes.

In this paper, we intend to validate Stan's method [4] using speech data from an indigenous language. To accomplish this, a system for field-based data collection must be designed and built. The collected speech data will be used at a later stage to train an Artificial intelligence (AI) model on a comprehensive and diverse dataset. Sufficient variation in pronunciation and intonation of the collected speech are needed to improvement of the automated speech recognition model's accuracy.

3 Problem Statement

Resourcing indigenous languages presents a number of obstacles. In regions where indigenous languages are spoken, the deployment of digital technologies is frequently hampered by numerous infrastructural constraints, for example, limited Internet. Data must be collected with users in the field. Meanwhile, there is typically a language and cultural gap between the locals and the ICT system's technical developers. Given the complexity of the local context, the focus of this paper is on determining the optimal method for designing and constructing a data collection system to produce a digitized language corpus.

4 Methodology and Approach

The general approach to designing digital technologies in low-resource environments is to co-design with local users in a living lab setting and learn from the experience and the feedback. The first action is a stakeholder analysis, including needs assessment. Next, a use case and requirements analysis is carried out. We choose agile development with short iterative cycles of rapid prototyping [5]. This breaks the development process into small, manageable stages of one or a few days, incorporating user feedback at every stage, facilitating the creation of a user-friendly interface, that meets the users needs [6, 7].

For the design, we adhere to the guidelines from the 10 Nielsen rules and NASA design, developed in the 1990s by Nielsen [8], and of the Human Interaction Design Protocol by the "NASA's International Space Station Program at Johnson Space Center in Houston, Texas" [9]. These are guidelines to design for safety, accessibility and adaptability. This includes communicating with users in language and concepts that are familiar for them, while giving them autonomy and control. It stresses the importance of simplicity and clarity of the design, and of consistency and predictability. This results in an intuitive and enhanced experience for users with different levels of skills.

The following two sections describe a case study in Ghana, in which a data collection application for resourcing of the small indigenous language *dagbanli* was developed. This was done iteratively, in a living lab setting, in which the users were able to test the application in the local context, at every stage of its development. From the findings of this field-based case study we draw lessons how to resource small indigenous languages.

5 Tiballi, A Field-Based Case Study in Ghana

This research is carried out in northern Ghana, with the objective to develop ASR-based voice services to provide weather and rainfall information to rural communities in their own native language. Developing a data collection application to generate a speech corpus is the first activity in this research project, dubbed Tiballi, which means “our language” in local *dagbanli*.

The data collection tool targets online communities that speak the chosen language and are already involved in community-centered work. The launch of the project involves a workshop that brings together stakeholders in the field: farmers from the communities, AI, linguistic and ICT4D experts from various organizations. These stakeholders form a group of knowledgeable individuals who will provide data and facilitate the spread of crowd-sourcing at the appropriate time. The process also employs individuals with experience in data collection. They are compensated to visit the communities and collect voice fragments from native speakers. Data collection starts in the selected Nyankpala and Tingoli, communities, in the Tolon District, and may be expanded to other regions.

5.1 Use Case Analysis

The project starts with a use case analysis, to find out the objectives and requirements for an app to collect utterances of dagbanli language. The next question is how the data collection app will be used and who the users are. Two of the most common use cases that are done though the app are displayed in Fig. 1:

Use case 1: User perspective (dagbanli speaker).
User persona 1: A dagbanli speaker who wants to contribute to the community by recording their voice and helping to create a knowledge base for a ML program.
Use case 2: Developer perspective.
User persona 2: The developer of the data collection app wants to monitor the contributions and performance of the app.

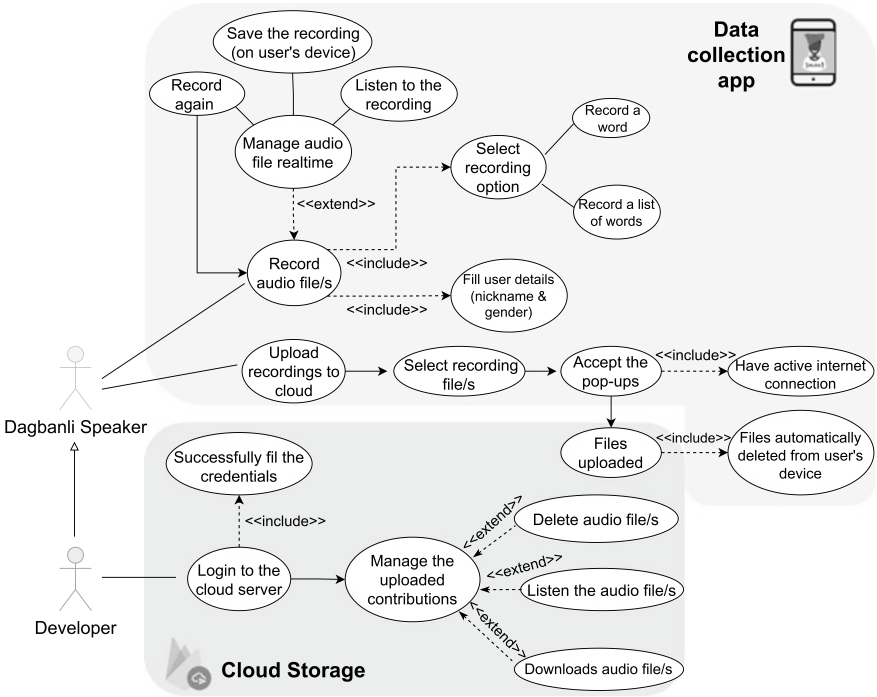


Fig. 1 Use case scenarios for the personas

5.2 *Stakeholder Analysis*

Identifying stakeholders and their specific needs with respect to the app is the 1st stage of the analysis. This project's stakeholders include app end-users, developers and the originators of the project concept. Secondly, given the poor local (digital) infrastructure, and the technical requirements for data collection, a digital platform should be available which is scalable, maintainable and secure for the collection of a language corpus of sufficient quality and size for voice data for model training. Thirdly, attention must be given to the human-computer interaction aspects of the app. The app's functional and non-functional requirements are described in detail in Sect. 6.1.

The 1st round of user feedback is mapped using the MoSCoW method, a technique for ranking importance of user requirements against ease/complexity/cost of development, and prioritizing which *must* be in the app, which *should* be, *could* be or *would not* be built in the app [10].

6 Designing a Data Collection App

6.1 Initial Requirements

The most basic technical and general requirements for the data collection app are specified and listed as (i) functional and (ii) non-functional requirements. The functional requirements:

- An audio recording's data storage must be arranged.
- Quality of the recorded speech must meet the machine leaning models.
- If the recorded audio quality is below quality, there must be an option for quality enhancement.
- A human quality control function must be in place to ensure the accuracy and quality of user-submitted transcriptions of the utterances.

Non-functional requirements relate in general to the performance, compatibility and usability, including performance, security and usability of an app. The following are collected for the envisaged app:

- The app must be able to quickly and efficiently process a certain volume of voice recordings.
- The app must function on multiple platforms: Smartphone and Desktop computers, to facilitate user accessibility in the field and at the office.
- An intuitive human-computer interface for recording and submitting voice data is required, that meets the local users' level of computer literacy.
- The audio must be of sufficient quality and volume, that in a subsequent phase, a specified ML model must be able to accurately recognize and transcribe the recorded utterances for the given language.

6.2 User Requirements Refined—1st Iteration

To ensure that the project's user requirements are precisely captured and refined during the 1st iteration of user-tests, prototyping techniques are used, such as the creation of a video and a website prototype, to communicate the project vision with all stakeholders. This process yields feedback that assists updating and refining the end-users' requirements. Rapid prototyping is a vital component of the design process, ensuring that design is rapidly adapted to the emerging requirements and expectations of test users.

In the prototyping stage, the developer employes resources such as proto.io[3] to construct interactive and operational prototypes. Proto.io provides a diverse range

[3] https://proto.io/.

of functionalities such as audio recording, cloud-based file uploading and cautionary pop-up notifications that enhance the design's realism and facilitate users in evaluating the design as if it were an actual product.

The interactive feature facilitates the evaluation of design functionality and the identification of potential impediments, enabling designers to make informed decisions regarding the design's trajectory and solicit feedback from relevant parties. The user emphasizes that Proto.io's interface is intuitive and suitable for designers of varying skill levels. In addition, the tool is renowned for its capacity to improve design quality and streamline the prototyping process.

In addition, a pre-recorded video prototype[4] of the design is generated. This elucidates functionalities and characteristics of the design in a non-interactive static format. Despite limitations in comparison to real-time prototypes, this approach is commonly employed in conjunction with other prototyping techniques to offer the users comprehension of the design and its potential impact.

The 1st iteration took place in Ghana, during a field visit to the local community of Tingoli, in February 2023. Deliberations led to the decision to develop two prototypes. Following the workshop, several modifications were implemented, including revisions to the text to enhance clarity and redesigning. Moreover, the data categories for the collected utterances were, for the 1st stage deployment, reduced to two. The two categories under consideration comprise numerical values 0–10 and a category for collection of "Yes" and "No" in dagbanli language. Additional word categories for collecting for example for weather and rainfall (rain, storm, drizzle) or relevant local crops (maize, rice, groundnuts) are planned for a later phase of data collection.

6.3 User Feedback—2nd Iteration

Upon implementation of the new requirements, the initial iteration of the mobile version of the app was deployed and shown to the users in Ghana. The results of the 1st cycle's requirements are shown below:

The main screens from the data collection app[5]: To record one single word, the user selects the "one word" option in the *ratio category* as displayed in Fig. 2. Once selected, the user can view available words, choose a desired word, and record it. "Record again", "Save" and "Play", to save or re-record it. As the user enters a category of words, additional options such as "Skip the word" and "Save and continue" will become accessible. By selecting the "Skip the word" icon, users can bypass unfamiliar terms. Additionally, users can save their progress by selecting the "Save and Continue" button, which displays the next word to be recorded. Once the last word in a category has been recorded, the "Finish and Submit" option becomes available.

---

[4] https://www.youtube.com/watch?v=U6P6FmC2_4s.

[5] The source code can be made available upon request: https://github.com/AntriaPan/TIBaLLi-project-voice-services.git.

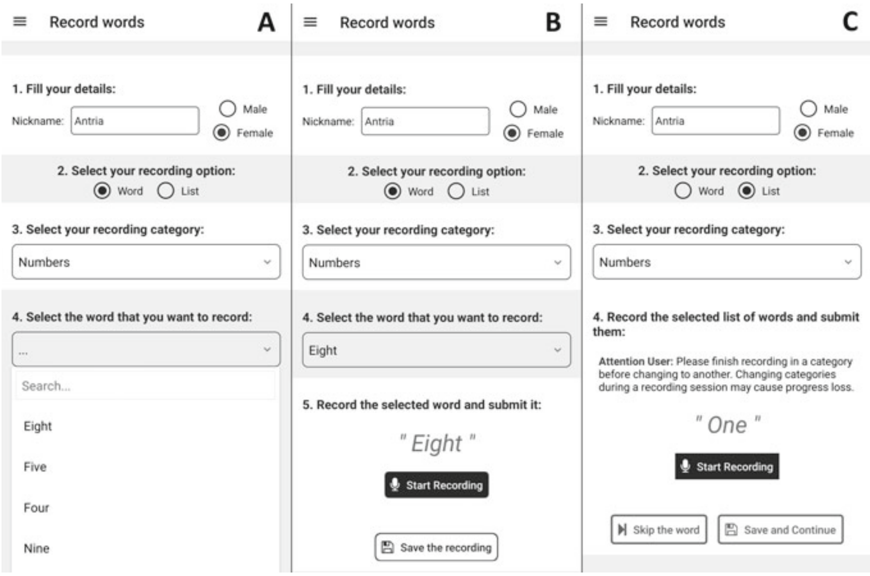


Fig. 2 Screenshots from the mobile app during the recording process [A: Selecting the recording word, B: Available "word" options, C: Available "list" options]

Before uploading, it is necessary to verify the stability of the Internet connection. An unstable connection prompts the user to retry the upload, whereas a stable connection initiates the upload process automatically, see Fig. 3. A loading indicator indicates upload status and progression. A message is displayed when the upload status reaches 100%. The app requires access to the device's microphone and storage for respectively audio capture and local storage of data.

The following recommendations were made during 1st user evaluation:

- incorporating a "simplified" mode to facilitate independent app navigation.
- providing the option to display the interface in the local language improves accessibility and usability for individuals who are fluent in the local tongue.
- recommendations were given to enhance the procedure of documenting and classifying vocabulary and expressions.
- restricting the frequency of pop-up displays during the uploading process, for a more streamlined and effortless process was proposed.
- substituting the "category" parameter with "list" to avoid user confusion.
- restoring the previously selected category and word option after word recording, and categorizing contributions by recording time or speaker.
- incorporation of both full-sentence recordings and single-word recordings.
- classification of recordings based on variables such as the speaker's gender, age or ambient noise. The aforementioned improvements have the potential to offer users a more thorough and genuine educational encounter.

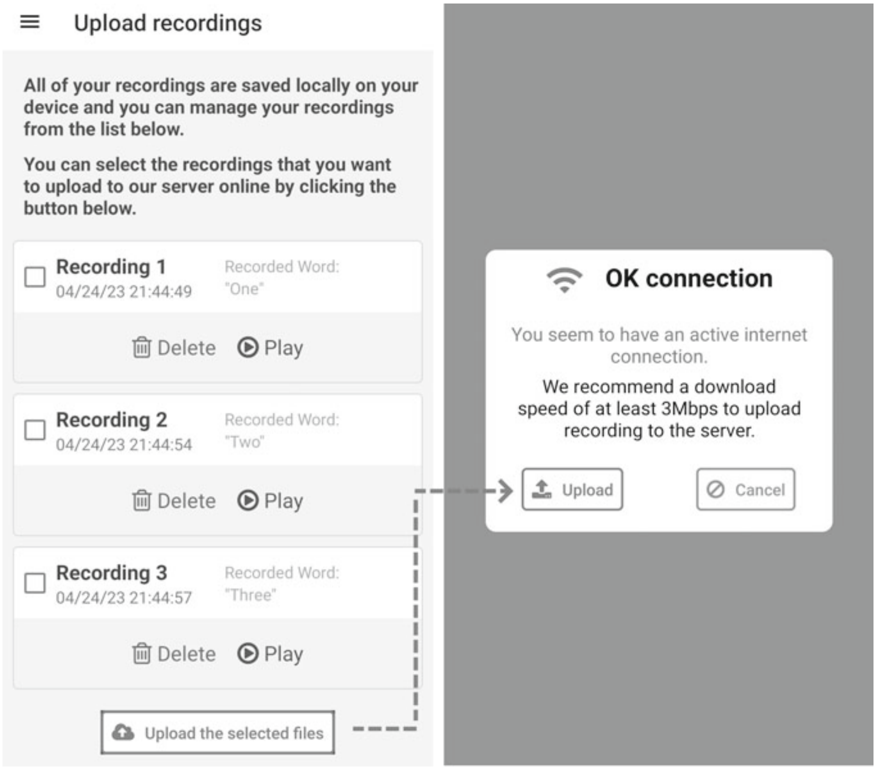


Fig. 3 Screenshots from the mobile app from the uploading process

- load words and categories to enhance efficiency and scalability of the process of updating and managing the app's content.

The aforementioned requirements and recommendations focus on functionality, accessibility and user engagement. An example of a proposed requirement that was dismissed was the option to record data of the contributing speakers. The project team dismissed this option in the app, because of increased complexity, when dealing with personal information.

Another requirement entailed consolidating all recording functionality onto a single page, saved with one single button click. This requirement was dismissed as data could get lost during recordings, in case of intermittent Internet connectivity. So, partitioning the actions on different screens remained the preferred design option.

The team debated multilingual functionality, but decided to include English translations of basic terms such as "yes" and "no" and numerical values in brackets next to recorded words, or vice versa. This makes it easier for proficient readers to recognize words based on their phonetic properties.

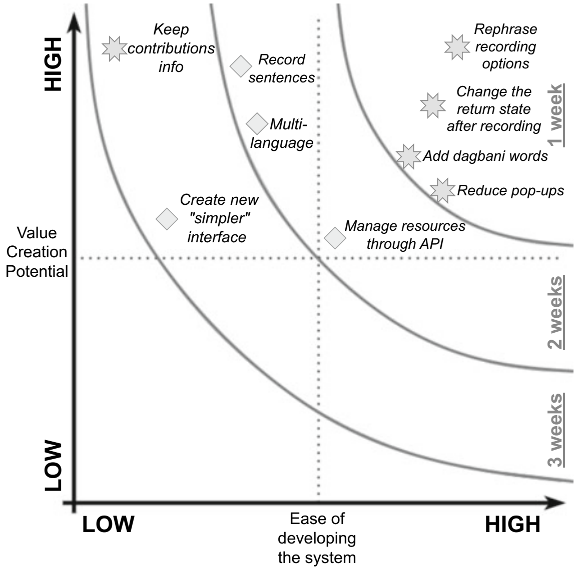


Fig. 4 Prioritization of mobile app feedback using the MoSCoW method [Stars: selected implemented feedback; Diamonds: feedback not selected]

Ultimately, the team implemented minor modifications to the app, including reduction of pop-up frequency during the uploading procedure, substitution of the "category" feature with "list", and alteration of the "word" option's return state following each recording, with the aim of enhancing user-friendliness. These modifications are considered significant to enhance user convenience and expediency, while avoiding potential ambiguity.

6.4 *Prioritizing Requirements Using MoSCoW*

When designing a system, it is necessary to balance the relevance of user requirements against the ease and cost of incorporating them in system. We used the MoSCoW method to map two dimensions, value for the user against cost of implementation. After deliberation the project team balanced the estimated relative value of each requirement, as show in the diagram on Fig. 4.

6.5 *User Feedback—3rd and Final Iteration*

During the 3rd iteration the mobile app was completely recreated and submitted for evaluation, to the satisfaction of the users in Ghana. A last comment entailed the incorporation of a text field for input of name and gender of the data collecting person. This feature was deemed indispensable for the purpose of audio file analysis in a subsequent phase. A salient facet was whether they would furnish their complete name or just a nickname. Following discussion, the decision was taken for a nickname, as many people share the same name, particularly in this region of Ghana, where the app is being utilized. This resolution will facilitate expansion of the app to broaden its user base.

7 Results and Conclusion

This study resulted in (i) a data collection app (ii) a general method to resource indigenous languages. Despite the contextual limitations and specific requirements of this project, we can generalize certain findings:

- Resourcing an indigenous language must be done in close collaboration with native speakers of the targeted language.
- Context analysis and stakeholder analysis are essential and need to be done at the start of the project.
- To bridge the communication gap between developer and user(s), field visits, test sessions and frequent meetings are required.
- Real-life demos followed by rapid prototyping and frequent user-tests are the appropriate for elicitation of requirements.
- User requirements are collected and documented, using a method of prioritization of value against costs/ease of implementation (MoSCoW).
- Context-awareness of the conditions and limitations are key success factors in designing for low-resource environments.

Acknowledgements We thank André Baart and Gideon Amakama Ali for their contributions to this study. We are thankful to the Internet Society, for the research grant that was provided to carry out the Tiballi project in northern Ghana.

References

1. Ndebele H (2020) Promoting indigenous African languages through information and communication technology localisation: a language management approach. Alternation Spec Issue 13:102–127
2. Toshniwal S, Sainath TN, Weiss RJ, Li B, Moreno P, Weinstein E, Rao K (2018) Multilingual speech recognition with a single end-to-end model. arXiv preprint arXiv:1808.02610
3. De Vries NJ, Badenhorst J, Davel MH, Barnard E, De Waal A (2011) Woefzela—an open-source platform for ASR data collection in the developing world. Conference paper
4. Stan GV (2021) "Small" language limited-vocabulary automatic speech recognition using ML. Web alliance for regreening in Africa
5. Bon A (2016) ICT4D 3.0, An adaptive, user-centered approach to innovation for development. https://doi.org/10.13140/RG.2.1.3260.5687

6. Cao L, Mohan K, Xu P, Ramesh B (2009) A framework for adapting agile development methodologies. Eur J Inf Syst 18(4):332–343. https://doi.org/10.1057/ejis.2009.28
7. Salah D, Paige RF, Cairns P (2014) A systematic literature review for agile development processes and user centred design integration. In: Proceedings of the 18th international conference on evaluation and assessment in software engineering, pp 1–10
8. Nielsen J, Molich R (1990) Heuristic evaluation of user interfaces. Association for Computing Machinery, pp 249–256. https://doi.org/10.1145/97243.97281
9. Boyer J (2014) Human integration design processes (HIDP)
10. Ahmad KS, Ahmad N, Tahir H, Khan S (2017) Fuzzy MoSCoW: a fuzzy based MoSCoW method for the prioritization of software requirements. In: 2017 international conference on intelligent computing, instrumentation and control technologies (ICICICT). IEEE, pp 433–437

A Cross-sectional Analytical Model for Cloud-Based ERP

Pranay Dadheech, Ajay Kumar Sharma, Mayank Patel, and Narendra Singh Rathore

Abstract This is well-known fact that to improve business processes, upsurge efficiency, and improve the decision-making capabilities, there is necessity to design and implement a cross-analytical model. Similarly, in the present context the role of cloud computing is increased to utilizing ERP with cloud support. This implicates opinion from stakeholders and users, identifying business processes that require improvement, data sources for integration, and analytical tools for data analysis. After that, next step is to customize the ERP system to meet organizational needs, including workflows, user access, and permission configurations. To standardize the integrated approaches into the ERP system, there is need of development of a model. An analytical model is developed using appropriate analytical tools, designing the model, and developing algorithms and statistical models. The analytical model is then tested and validated by running test scenarios, validating results, and making necessary adjustments. The present study tries to give a cross-analytical model to manage the cloud-based ERP system.

Keywords Cross section · Cloud · ERP · Model · Analytical · Organization · Tools

P. Dadheech (✉) · A. K. Sharma · M. Patel
Department of CSE, Geetanjali Institute of Technical Studies, Udaipur, India
e-mail: pranaydadhich98@gmail.com

N. S. Rathore
Geetanjali Institute of Technical Studies, Udaipur, India
e-mail: director@gits.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_49

1 Introduction

To enhance business processes, increase efficiency, and improve decision-making capabilities, a cross-analytical model of cloud computing with ERP adoption must be designed and implemented. The process starts with requirements gathering from stakeholders and users, identifying business processes for improvement, data sources for integration, and analytical tools for data analysis. Next, the cloud infrastructure needs to be setup by selecting a provider, configuring the infrastructure, and implementing security measures. Customization of the ERP system to meet organizational needs, including workflows, user access, and permission configurations is then carried out. Data is integrated from various sources into the ERP system by extracting and transforming it into a standardized format. An analytical model is then developed by selecting appropriate analytical tools, designing the model, and developing algorithms and statistical models. Testing and validation of the analytical model are done through running test scenarios, validating results, and making necessary adjustments. Lastly, the cross-analytical model is deployed into production, and ongoing maintenance and support are provided to ensure continued functionality and meeting of organizational needs [1, 2].

2 Aim and Outline of Model

The purpose of conducting a cross-sectional analysis of cloud-based ERP varies depending on the research question and the specific context of the study. However, some common objectives of cross-sectional analysis are to compare the performance of organizations that have adopted cloud-based ERP systems with those that have not, identify the factors that contribute to successful adoption and implementation of cloud-based ERP systems, and compare different cloud-based ERP systems based on their features, functionality, and performance [3].

In the present study, the objectives of cross-sectional analysis in cloud-based ERP studies are as follows:

- To identify the current state of adoption and usage of cloud-based ERP systems.
- To determine the significance of cross-sectional analysis of different organizations that have adopted cloud-based ERP systems.
- To collect facts on the prevalence, extent, and nature of adoption, providing insights into the current state of the market and the factors that influence organizations' decisions to move to the cloud.
- To evaluate the impact of cloud-based ERP on organizational performance.
- To identify the key factors that influence the adoption and success of cloud-based ERP systems.

3 Model for Cloud-Based ERP Adoption

Rendering to the authors AlBar and Hoque [1] in association of Moh'd Anwer Al-Shboul, few considerable factors are recognized under the few models, in order to cultivate the investigation model which is convenient to the adoption of cloud-based ERP. From them the study considering there key combine approaches. The projected model for the research is displayed in a view which empirically inspects the impression of TOE, innovation. Therefore the key factors for the research model are:

1. Technology-organization-environment (TOE) framework
2. The model of innovation resistance (MIR)
3. The diffusion of innovation (DOI) theory.

Above mentioned each of the crucial factors may be divided into sub-factors to sumptuous the physiognomies of every key factor. These sub-factors are displayed with the help of below mentioned table. Thus, it is depicted that there are 11 sub-factors which are well-thought-out for the study under three key models for the adoption of cloud-based ERPs particularly for the industries of developing countries [4, 5].

The technology-organization-environment (TOE) framework is a theoretical model that describes how various factors interact to affect the adoption and implementation of new technologies in organizations. Proposed by Tornatzky and Fleischer in 1990, the TOE framework has become widely used in research and practice (Table 1).

Developed by Everett Rogers, the diffusion of innovation (DOI) theory is a framework that explains how new ideas, products, and technologies spread through society over time. The theory posits that the adoption and diffusion of innovations follow a predictable pattern characterized with five stages (Fig. 1).

The model of innovation resistance is a theoretical framework that explains why individuals or organizations may resist adopting new innovations, even when they

Table 1 Key factor and sub-factors of the model

S. No.	Key factor	Sub-factor
1	TOE	• ICT-based cloud ERP skill • Technological and organizational development approach • Smart governance approach
2	MIR	• Comparative benefit aspect • Sophistication reduction • Endorsement aspect • Perceives aspects
3	DOI	• Security trail • Vendor dependency • Customization dependency • Availability

Fig. 1 Key factors for cross-sectional model

could be beneficial. The model suggests that there are five types of barriers to innovation adoption, which can be categorized into three groups: cognitive, emotional, and behavioral.

4 Research Process for Cloud ERP Adoption

The trend of adopting cloud-based enterprise resource planning (ERP) systems is on the rise among organizations seeking to enhance their efficiency, streamline their operations, and reduce costs. To investigate the cloud ERP adoption model, the following steps should be taken [6]:

- Clearly define the research problem and formulate research questions that can be answered. For instance, "What are the factors that affect the adoption of cloud-based ERP systems in organizations?"
- Conduct a thorough literature review to gain insights into the current state of research on cloud ERP adoption, identify research gaps, and build on existing knowledge.
- Develop a conceptual framework that identifies the organizational, technological, and environmental factors that influence cloud ERP adoption.
- Create testable and measurable research hypotheses based on the conceptual framework.
- Collect both qualitative and quantitative data using surveys, interviews, and other methods to test the research hypotheses.
- Analyze the data collected using statistical analysis techniques to determine the factors that influence cloud ERP adoption.
- Present the results in a comprehensive report that includes an executive summary, introduction, literature review, methodology, findings, discussion, and conclusion.
- Provide recommendations for organizations based on the findings to help them make informed decisions about cloud-based ERP adoption.

The cloud ERP adoption research model involves a detailed examination of the decision-making process and provides recommendations to help organizations make informed decisions about cloud ERP adoption [7].

5 Cross-sectional Analytical Model

The cross-sectional analytical model is fully implementation of research process for cloud adoption. At the initial stage, the research problem and questions should be defined clearly on the basis of factors TOE, MIR, and DOI. This step is accountable as the conceptual framework is developed to identify the organizational, technological, and environmental factors that impact cloud ERP adoption [8, 9] (Fig. 2).

The research hypotheses should be created based on the conceptual framework and should be testable and measurable. The qualitative and quantitative data should be collected using surveys, interviews, and other methods to test the research hypotheses.

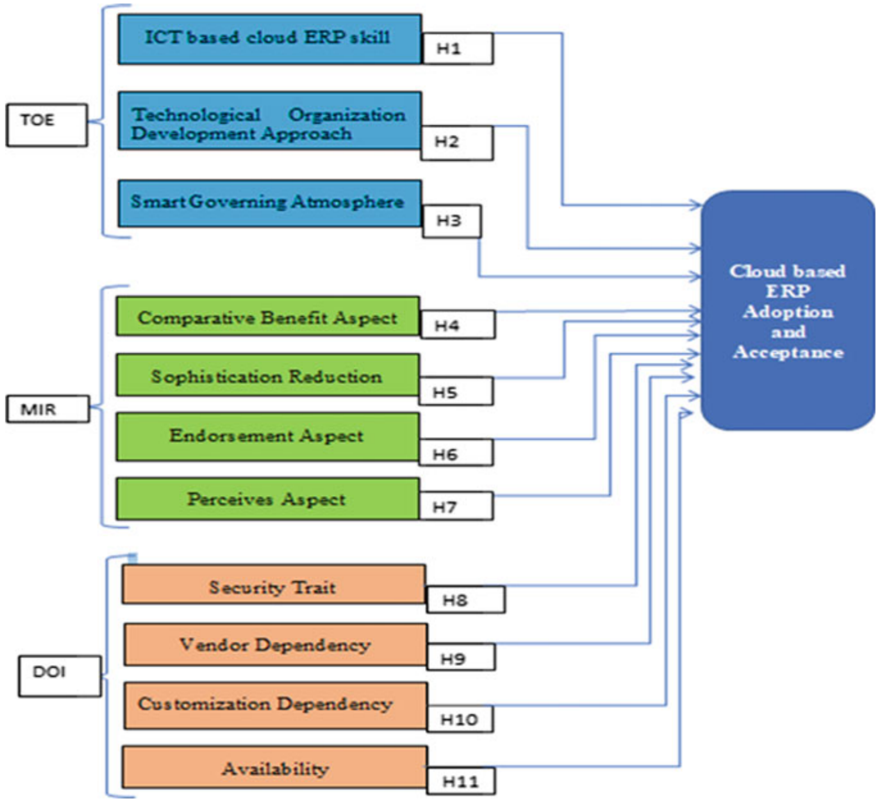


Fig. 2 Cross-sectional analytical model for cloud ERP adoption

The collected data should be analyzed using statistical analysis techniques to determine the factors that influence cloud ERP adoption [3].

The results should be presented in a comprehensive report that includes an executive summary, introduction, literature review, methodology, findings, discussion, and conclusion. Recommendations for organizations should be provided based on the findings to assist them in making informed decisions about cloud-based ERP adoption.

6 Conclusion

The projected cross-sectional analytical model gives a clear picture in order to adopt the cloud-based ERP system for any organization. There is need to develop hypothesis in the association of factors and assumptions which satisfy both factors and requirements of the organizations. Satisfaction of both the domains is essential part of the model. The quantitative test with hypothesis gives a significance picture to suggest that degree of success of cloud-based ERP model on the projected organization. The projected model is nicely designed to solve the problem of cloud-based ERP adoption. The study and model give a proper path to accept the cloud in the organization to manage the ERP system.

References

1. Al-Shboul MA (2019) Towards better understanding of determinants logistical factors in SMEs for cloud ERP adoption in developing economies. Bus Process Manag J 25:887–907
2. Sharma A, Garg S (2015) Comparative Study of cloud computing solutions. IJCST 6(4). ISSN: 0976-8491
3. Al Mudawi N, Beloff N, White M (2022) Developing a framework of critical factors affecting the adoption of cloud computing in government systems (ACCE-GOV). In: Arai K (ed) Intelligent computing. Lecture notes in networks and systems, vol 283. Springer, Cham
4. Alouffi B, Hasnain M, Alharbi A, Alosaimi W, Alyami H, Ayaz M (2021) A systematic literature review on cloud computing security: threats and mitigation strategies. IEEE Access 9:57792–57807
5. Ahn B, Ahn H (2020) Factors affecting intention to adopt cloud-based ERP from a comprehensive approach. Sustainability 12:6426
6. Anderegg T (2006) MRP/MRPII/ERP/ERM—confusing terms and definitions for a Murkey alphabet soup. Available at http://www.wlug.org.nz/EnterpriseSpeak. Accessed on Dec 2010
7. Hassan A, Bhatti SH, Shujaat S, Hwang Y (2022) To adopt or not to adopt? The determinants of cloud computing adoption in information technology sector. Decis Analytics J 5:100138
8. Song C, Sohn Y (2022) The influence of dependability in cloud computing adoption. J Supercomput 78:12159–12201
9. Alsmadi D, Halawani M, Prybutok V, Al-Smadi R (2022) Intention, trust and risks as core determinants of cloud computing usage behavior. J Syst Inf Technol 24(3):178–201

A Study of Linkage Between Strategic Decision and Performance of Organization with Reference of Hospitals of Rajasthan

Prerna Tiwari, Sandeep Chourasia, and Sanjay Gour

Abstract In the current situation, performance of an organization is the foremost question as customers are too much aware about the services and the environment provided to them. The standing of the organization in the market totally depends on the productivity and a facility provided by the organization, the thing is too much crucial when organization is dealing as the service industry. It is well-known fact that the performance of the organization is moreover dependent on the decision of the top/strategic management along with management of operational level staff. This is correspondingly true that motivation of the employee to do well for the organization affects the performance of the organization. From time to time, it is witnessed that a good leader on the leading position in the organization plays a significant role to make the organization identifiable despite other mentioned factors. This will happen with appropriate strategic decisions which matter a lot in the service industries. The projected paper is an attempt to assess the linkage between leading working position, strategic decision and performance of organization with special reference of hospitals of Rajasthan (Ajmer Zone).

Keywords Performance · Strategic decision · Organization · Motivation · Linkage

1 Introduction

This is mutual thought that an organization is driven by the top management or by a being that has been holding leading position for executing the organization's management and taking strategic decisions. Also there may be a team of certain persons/board of management and some regulatory body, the entire will be termed as

P. Tiwari · S. Chourasia
School of Management Studies, Sangam University, Bhilwara, Rajasthan, India

S. Gour (✉)
Computer Science and Engineering, Jaipur Engineering College and Research Centre, Jaipur, India
e-mail: sanjay.since@gmail.com

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_50

top management. This is a notable fact that performance service class organization is accomplished by the regular employees and working professionals who are working for providing services [1]. As this study is based on the hospitals, therefore, whole thing is services of the organizations/hospitals. Likewise there are numerous activities which are managed by the motivated employee. The employee motivation derives by the healthier performance appraisal scheme implementation and execution in the hospitals.

The performance has arisen by means of positive organizational performance moment. These days the hospital/health industries services are also notable in the professional world which bring strong-willed prospective to come across the challenge of continuous quality improvement to satisfy the patient demands. Thus hospitals/health industries are compulsory to confirm the reputation of their services to meet the satisfaction level of the patient and medical practitioners [2].

The current study is fundamentally a valuation study of the view of the employee to recognize the protagonist of foremost management and its straight association from the performance of the hospital. This study is carried out by the opinion of employee and staff of hospital of Ajmer Zone about the role of top management and their strategic decision in order to welfare/appreciation providing to employee and its impact on the performance of the hospitals.

2 Factor of Organization Performance

In general, performance of any organization is the issue of worth, working procedure followed in the organization. Likewise it is witnessed that, the organization performance is the blend of the performance of the employees and management strategy [3, 4], which is further noted as sustainability of organization and maintaining of this is known as strategic practices (Fig. 1).

The key factors which are too much critical to cover the performance of the organization comprise mixture of management and employee performance, as well as strategies, appraisal system and policies to implement the strategies [5, 6]. If the appraisal system is faultless and helpful to employee/staff motivation, then it will give encouraging sites for the organization. Consequently there is necessity of linkage between subsequent factors of the organization.

1. Management performance
2. Employee performance
3. Quality management
4. Appraisal system.

The noteworthy objective of human resources management is appropriate operation of the resources in a supreme way so that objectives may be accomplished positively [7].

Fig. 1 Factors for organization performance

3 Aim of the Study

The performance of any hospital is attained by the consistent employees and staff in the professional manner regarding services who are employed for patients. Likewise there are other activities doing in the parallel which are also much managed by the supportive staff. The employee motivation is derived by the healthier appraisal system. Motivating the forces in the accurate way is also significant feature which is managed by top management by strategic decisions [8, 9].

The objective of present study is to find out the reality of the proclamation that performance of the hospitals/organization is linked with strategies of leading working positions. Consequently, the present study is a struggle to measure, i.e. performance of the organization and its linkage with strategic management of the organization.

4 Hypothesis

The present hypothesis tries to examine about the relationship between the performance of the organization and its linkage with strategic management of the organization with reference to hospitals of Ajmer Zone of Rajasthan. In such context the hypothesis is framed as:

Null Hypothesis (H_0): Performance of the hospitals is directly linked with the appraisal policy drive by strategic management.
Alternate Hypothesis (H_1): Performance of the hospitals is directly not linked with the appraisal policy drive by strategic management.

5 Assessment of Hypothesis

To achieve the evaluation of hypothesis, we are captivating opinion of 250 respondents. According to the nature of the hypothesis, chi-square test is found suitable to obtain the result from various hospitals.

The view of the employees were gathered on the five point Likert scale as the response, viz. strongly agree, agree, moderate, not agree and strongly not agree from the assumption mentioned above (Table 1).

The employee must put their single opinion on the assumption, same will depict the view of respondents in allusion of their acceptance level about the statement. Here, for the present study we are taking assumption about the statement that performance of the hospitals is directly linked with the appraisal policy of strategic management (Table 2).

Table 1 Five-point scale

Nomenclature	Strongly agree	Agree	Moderate	Not agree	Strongly not agree
Point of Likert scale	5	4	3	2	1

Table 2 Observed frequency

Opinion for supposition	Management persons	Employee	Total
Strongly agree	47	54	101
Agree	30	50	80
Moderate	8	25	33
Not agree	7	12	19
Strongly not agree	8	9	17
Total	100	150	250

To treasure the explanation about the hypothesis, we are captivating the response of employee and management persons. The joint figure for opinion strongly agree is noted 101, for agree 80, for the moderate it is 33, for not agree it is 19 and for strongly not agree option it is recorded 17. These are denoted as the observed frequency for the hypothesis. Overall, we can say that 181 opinions are positive with statement and 36 are negative whereas 33 are neutral (Fig. 2; Tables 3 and 4).

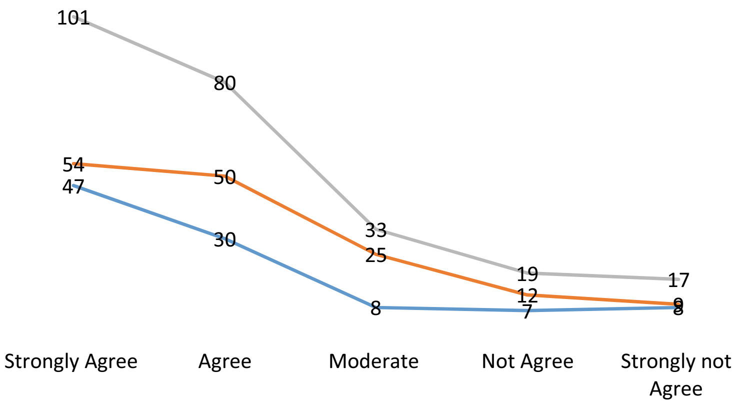


Fig. 2 Observed frequency

Table 3 Expected frequency

Total of type of opinions	Total of management person/employee	Expected frequency
101	100	40.40
80	100	32.00
33	100	13.20
19	100	7.60
17	100	6.80
101	150	60.60
80	150	48.00
33	150	19.80
19	150	11.40
17	150	10.20

Table 4 Calculation of X^2

Observed frequency (OF)	Expected frequency (EF)	$(OF - EF)^2$	$(OF - EF)^2/EF$
47	40.4	43.56	1.08
30	32	4	0.13
8	13.2	27.04	2.05
7	7.6	0.36	0.05
8	6.8	1.44	0.21
54	60.6	43.56	0.72
50	48	4	0.08
25	19.8	27.04	1.37
12	11.4	0.36	0.03
9	10.2	1.44	0.14
Chi-square value			**5.85**

6 Analysis of Result

Subsequent are the calculation as per the hypothesis test chi square and the explanation of the result.

$$\begin{aligned}\text{Degree of Freedom} &= (r-1)(c-1)\\ &= (5-1)(2-1)\\ &= 4\end{aligned}$$

where the tabulated value @5% level of significance is = 9.488.

Consequently, the intended value of chi-Square = 5.85.

Tabulated value of chi square = 9.488.

At this time, the intended value of chi square (5.85) is lesser than tabulated value 9.488 at 0.5% level of significance; henceforth the null hypothesis is accepted. It means that performance of the hospitals is directly linked with the appraisal policy drive by strategic management.

7 Conclusion

As per the study and projected hypothesis, opinion of the management persons and employee was collected from the different hospitals of the Ajmer Zone. After analysis of the opinion of the employee and management persons against hypothesis with chi-square test, it is found that hypothesis gets accepted. Thus, we can say that performance of the hospitals is directly linked with the appraisal policy drive by strategic management. The study springs a believed direction that performance of the organization depends on the strategic decision and implements it as management practices for the sustainability of the organization. Here it is notable that administration along with the operational management persons and employees as work force impacts on the performance of the organization by the decision of top management. The organization is itself quality working place and is dependent on the motivation of employee with good appraisal system provided by strategic management.

References

1. Altpeter M, Schneider EC, Whitelaw N (2014) Examining strategies to build and sustain healthy aging programming collaboratives. Health Educ Behav 41(1):27–33
2. Ament SMC, Gillissen F, Moser A, Maessen JMC, Dirksen CD, Meyenfeldt MF, Weijden T (2014) Identification of promising strategies to sustain improvements in hospital practice: a qualitative case study. BMC Health Serv Res 14:641
3. Burli S, Kotturshettar B, Kalghatgi P (2011) Impact of quality management practices on the organisational performance of small and medium scale manufacturing industries. Int J Manag Res Rev 1(4):63–77
4. Lamberta G, Ouedraogob N (2008) Empirical investigation of ISO 9001 quality management systems impact on organisational learning and process performances. Total Qual Manag 19(10):1071–1085
5. Sun H, Zhao Y, Yau HK (2009) The relationship between quality management and the speed of new product development. TQM J 21(6):576–588
6. Tanninen K, Puumalainen K, Sandstrom J (2010) The power of TQM: analysis of its effects on profitability, productivity and customer satisfaction. Total Qual Manag 21(2):171–184
7. Piskar F (2007) The impact of the quality management system ISO9000 on customer satisfaction of Slovenian companies. Manag Glob Transitions 5(2):45–61
8. Marimuthu M, Paulose H (2016) Emergence of sustainability based approaches in healthcare: expanding research and practice. Procedia Soc Behav Sci 224:554–561
9. Ramirez B, West DJ, Costell MM (2013) Development of a culture of sustainability in health care organizations. J Health Organ Manag 27:665–672

Secure and Explainable Artificial Intelligence (XAI) in Cloud Ecosystems: Challenges and Opportunities

Darshi Khalasi, Devang Bathwar, Jitendra Bhatia, Malaram Kumhar, and Vinodray Thumar

Abstract The widespread adoption of artificial intelligence (AI) in cloud ecosystems has revolutionized various domains, offering significant advantages such as improved efficiency, scalability, and cost-effectiveness. However, the use of AI also poses significant security risks, and the need for more transparency in AI models makes it difficult to understand how decisions are made. This paper reviews recent research efforts toward developing secure and explainable AI (XAI) for cloud ecosystems, highlights the critical need for secure XAI in cloud ecosystems, and explores the challenges and potential solutions. The paper also discusses a case study involving the security of cloud-based health care with intrusion detection systems (IDS). Finally, the paper concludes with open research directions in this domain and their potential impact on cloud ecosystems.

Keywords Secure AI · Explainable AI · Cloud ecosystem · Machine learning · Privacy-preserving techniques · Federated learning · Transparency · Accountability

D. Khalasi (✉) · D. Bathwar
Department of Electronics and Communication, Institute of Technology, Nirma University, Ahmedabad, India
e-mail: 20bec056@nirmauni.ac.in

D. Bathwar
e-mail: 20bec016@nirmauni.ac.in

J. Bhatia · M. Kumhar
Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, India
e-mail: jitendra.bhatia@nirmauni.ac.in

M. Kumhar
e-mail: malaram.kumhar@nirmauni.ac.in

V. Thumar
Department of Computer Engineering,Vishwakarma Government Engineering College, GTU, Ahmedabad, India
e-mail: vdthumar@vgecg.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_51

1 Introduction

A cloud ecosystem is a network of services, applications, and platforms collaborating to provide users with cloud computing services. The ecosystem includes cloud service providers, users, third-party software vendors, developers, and consultants. Cloud computing has transformed how data is stored, processed, and managed, leading to the widespread adoption of AI-based applications that process sensitive data. Artificial intelligence (AI) technologies' rapid proliferation has transformed how businesses operate and deliver services. Cloud computing, with its inherent scalability and accessibility, has become an ideal platform for deploying AI models and applications. AI adoption in cloud computing has several advantages, including increased productivity, accuracy, and decision-making. However, these AI models may face various security risks because AI models are susceptible to hostile attacks, data breaches, etc. Furthermore, it can be challenging to comprehend how decisions are made when using AI models, which can produce biased results and reduce confidence in AI-based systems. Figure 1 shows the various cybersecurity attacks in 2022.

The major contribution of this paper is as follows:

- We explored the critical requirements of developing secure and explainable AI for cloud ecosystems, essential for building trustworthy and robust AI-based systems.
- We review the recent research efforts toward achieving secure and explainable AI and discuss the challenges and opportunities in this field.
- We identify the open research directions for promoting the security and transparency of AI-based systems in cloud ecosystems.

The rest of the paper is organized as follows: Sect. 2 discusses the background of secure and explainable AI. Section 3 presents the related work. Section 4 presents the implementation steps involved in XAI in the cloud ecosystem. Section 5 discusses

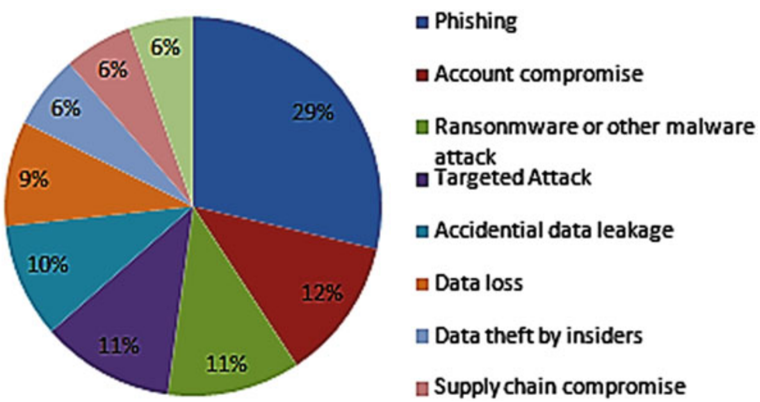


Fig. 1 Cloud security attacks [1]

the need for secure and explainable AI in the cloud ecosystem. Section 5 presents a case study on the healthcare system that adopts AI in the cloud ecosystem. Section 6 offers open research directions. Finally, Sect. 7 concludes the paper.

2 Background

Secure and explainable AI is critical for building trustworthy and robust AI-based systems in cloud ecosystems. Secure AI aims to protect AI models from malicious attacks and ensure data confidentiality, integrity, and availability. Furthermore, it seeks to provide transparency and accountability in AI models by enabling the users' decisions.

2.1 Secure Artificial Intelligence

Security involves ensuring the confidentiality, integrity, and availability of data. Organizations implement robust security control for secure AI for the entire AI life cycle. This comprises encryption, access controls, monitoring, and auditing. Some critical considerations for building a secure AI system in a cloud ecosystem are

- *Data Privacy*: Protecting privacy is an essential task in building a secure AI system; it ensures that sensitive data is not exposed by implementing appropriate security controls like encryption, access controls, and data randomization,
- *Secure Infrastructure*: A secure cloud infrastructure ensures that AI systems are protected from external hazards. Cloud security includes network security, firewalls, and intrusion detection and prevention systems.
- *User Access Controls*: To prevent unauthorized access to AI systems, which involves implementing access controls such as multi-factor authentication and role-based access control.
- *Secure AI Models*: A secure AI system is essential to ensure that AI models are secure. By implementing measures such as model validation and testing, including security considerations into the development process.

  Constructing a secure AI system in a cloud ecosystem requires a systematic approach to defending data privacy, executing a secure infrastructure, controlling user access, protecting AI models, and conducting regular security audits. By implementing these measures, organizations can help to ensure that their AI systems are protected from cyber threats and can operate securely and effectively in the cloud.

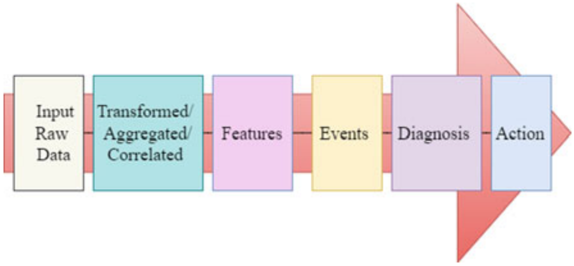


Fig. 2 Data workflow in XAI [2]

2.2 *Explainable Artificial Intelligence*

An emerging area of artificial intelligence called explainable AI (XAI) aims to develop machine learning models and systems that can explain their decision-making procedures in a simple and straightforward. XAI is particularly crucial in cloud computing, where machine learning models are frequently trained on enormous amounts of data and used to make decisions that can significantly affect businesses and individuals. One aspect of XAI in the cloud is the requirement for accountability and transparency. Both cloud service providers and AI developers need to take responsibility for the judgments made by their machine learning models and be open about how they gather and use data. Figure 2 shows the data workflow in XAI.

- *Interpretable Machine Learning (IML)*: One approach to XAI is using interpretable machine learning algorithms. In the cloud ecosystem, various interpretability techniques, such as decision trees, linear models, and rule-based systems, are offered by tools such as scikit-learn, TensorFlow, and Keras used for implementing IML
- *Model Explainability Services*: To offer model explainability services is another approach to XAI that by providing users with explanations for the AI models' decisions. These services can be integrated into the cloud ecosystem, allowing users to request reasons for specific decisions or to receive proofs automatically with every prediction made by the model. Visualization tools can be integrated into the cloud ecosystem, giving users a better understanding of the model's decision-making process.
- *Explainable AI Frameworks*: Several open-source frameworks provide XAI capabilities, such as ELI5, SHAP, and LIME. These frameworks can be integrated into the cloud ecosystem, allowing users to perform XAI tasks easily.

3 Related Work

AI security and comprehensibility are key components of creating reliable AI systems. Securing and explaining AI becomes even more difficult in the context of the cloud ecosystem, where data and computation are frequently distributed across numerous servers and devices. Table 1 summarizes recent studies on secure and explainable AI in the cloud ecosystem.

Bayuk et al. [3] discussed the significance of introducing security in cloud computing environments. The author presents that new metrics must be developed to consider the special characteristics of cloud computing because traditional security metrics are frequently insufficient for cloud environments. The author offers several cloud security metrics to assess the security of cloud environments. One such metric is the monthly incident rate, which counts the security incidents in a cloud environment during a specific time frame. The percentage of data breaches discovered and remedied within a predetermined time frame is another metric businesses can use to gauge the efficiency of their incident response procedures.

Arora et al. [4] discussed the challenges of ensuring data security and privacy in cloud computing environments. The authors argue that new approaches are needed to address the unique challenges of cloud computing, as traditional security measures are insufficient for cloud environments. The importance of each layer is explained by providing examples of specific security measures that can be applied to ensure data security and privacy in the cloud are hereby discussed. The paper offers important insights into developing effective security measures for cloud computing environments and taking a multi-layered approach to cloud security.

Divya et al. [5] observed the use of artificial intelligence (AI) to enhance the security of cloud environments publically used. The authors argue that using AI can help address the challenges of securing public cloud environments, as they are highly complex and dynamic. The paper discusses some cases for AI in cloud security, including threat detection and response, security automation, and risk assessment, while validating the potential benefits of each use case by providing examples of specific AI techniques that can be used to increase the efficiency of cloud security. With this, the paper highlights the potential benefits of using AI to enhance cloud security and provides valuable insights about the associated challenges and opportunities.

El-ghazal et al. [9] state a secure and explainable deep learning approach for cloud-based IoT applications by implementing federated learning and model interpretability. A case study of the approach applied to a smart home IoT dataset is also provided. Wang et al. [10] discuss the secure and efficient algorithm for outsourced matrix computation in cloud computing, which ensures the confidentiality and integrity of the data while minimizing the time needed by a node to process a packet.

McMahan et al. [7] proposed a framework for privacy-preserving deep learning, which uses centralized learning and different privacy techniques to enable multiple groups to jointly train a model on their data without giving away sensitive information. Haiqin et al. [11] proposed a secure federated learning framework that allows multiple units to collaborate on a machine learning task while protecting data privacy by implementing homomorphic encryption.

Pantelis et al. [12] provide a comprehensive survey of the state-of-the-art explainable AI techniques, including model-agnostic and model-specific methods, as well as their applications in various domains. Ibrahim et al. [13] proposed a survey of explainable AI techniques specifically tailored for the cloud computing environment,

Table 1 Existing works and their contributions

Author	Year	Core contribution	Scope of improvement
Bayuk et al. [3]	2011	Metrics assess the security of cloud computing environments and quantitative measures	SoSE approach to cloud computing security
Wentao Liu et al. [6]	2012	Give insight into the security issues with cloud computing and its novel suggestions for solutions to those issues	Use of block chain, machine learning improve cloud security. Integrating new technologies into cloud security frameworks
Arora et al. [4]	2017	Demonstration of a hybrid cryptographic system (HCS). Uses encryption techniques at various levels to ensure data security and privacy	Backup and recovery tools could be included to prevent data loss in the event of an attack
McMahan et al. [7]	2017	Provides a framework for privacy-preserving deep learning, to enable multiple parties to train a model on their data without revealing sensitive information	Algorithms to achieve numerical stability Sacred nature of difficulty for continued success; Formal security framework
Mohandas et al. [8]	2020	AI applied to prevent cyberattacks. Before harming the data kept by cloud service providers (CSPs), cyber threats will be "halted in their tracks"	Benchmarking various AI algorithms or models and evaluating; contrasting their performance with conventional cloud-based methods
Divya et al. [5]	2020	Primary commitment is its knowledge into the security issues with distributed computing. Work on the security and constancy of cloud-based frameworks	Developing new threat models that reflect the evolving nature of cloud security threats

including explainable deep learning and explainable reinforcement learning. Sajid et al. [14] surveys the state-of-the-art explainable machine learning techniques applicable in the cloud computing environment, including techniques for model interpretation, visualization, and explanation. Wang et al. [15] explored the trade-off between accuracy and explainability in AI models. The authors presented a hybrid approach that combines interpretable rule-based models with deep neural networks and fostered a reasonably profound learning structure that helps clinicians understand the forecasts made by the artificial intelligence framework.

4 Implementation of XAI in Cloud Ecosystem

The XAI has been incorporated into the machine learning life cycle. It is taking on the responsibility and sealing the expectation to explain and translate black-box algorithms, which are used for stakeholder critical business decision-making processes,

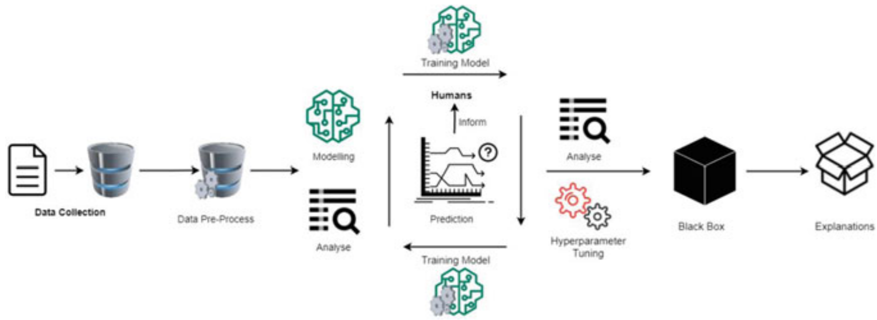


Fig. 3 Process of XAI in cloud ecosystem [16]

ensure that their AI systems are visible, explicable, and responsible. Figure 3 shows the process of XAI in the cloud ecosystem.

- *Data Collection*: This first step deals with collecting the data representing the problem domain. The data must be properly structured and labeled to train the AI system.
- *Model Training*: The next step is to train the AI model using the collected data. Ensuring that the AI model is transparent and explainable is important during this step and is achieved by selecting appropriate algorithms and documenting the model's architecture and parameters.
- *Model Testing and Evaluation*: Once the AI model has been trained, it must be tested and evaluated to ensure its functioning as expected. It tests the model on a separate dataset which evaluates its performance metrics. Accuracy and precision are the two parameters that are being evaluated.
- *Implementation*: XAI techniques are implemented when tested and evaluated to make the model more transparent and explainable. This involves techniques such as visualization of the decision-making process and model debugging tools.

4.1 Key Features of Explainable AI in Cloud Ecosystem

Explainable artificial intelligence is significant for building trust, guaranteeing straightforwardness, and empowering human oversight in computer-based intelligence frameworks. With regard to the cloud environment, there are a few key highlights that are fundamental for executing logical artificial intelligence.

Cloud-based XAI platforms provide methods to interpret AI models effectively. This involves techniques such as importance feature analysis and decision boundary visualization, which help users to understand how the model arrives at its predictions or classifications. XAI platform provides transparency into the architecture and design of the AI models. Users have visibility into the model's structure, including the layers, connections, and parameters, allowing them to understand the underlying logic and processes.

XAI stages support extracting rules or choice trees from black-box models. This assists clients in understanding the dynamic cycle by changing complex models into comprehensible principles, which can be analyzed and approved. XAI stages in the cloud ought to offer ongoing clarifications, empowering clients to comprehend and approve computer-based intelligence forecasts or choices sooner rather than later. This element is especially significant while sending simulated intelligence models in time-sensitive applications like money, medical services, or independent frameworks.

An effective XAI environment should work with an input circle between the artificial intelligence framework and the client. Clients ought to have the option to give input, seek clarification on some critical issues, or solicit extra clarifications, which can further develop the artificial intelligence model's straightforwardness and address expected issues. A vital element of computer-based intelligence in the cloud system is to engage clients with the capacity to comprehend, approve, and trust the results of artificial intelligence models through interpretable clarifications and straightforward dynamic cycles.

5 XAI in Security Model of Cloud Ecosystem

Explainable artificial intelligence is a concept that focuses on developing AI models and systems that can provide clear and understandable explanations for their decisions and behaviors. When it comes to the security model of a cloud ecosystem, implementing XAI can bring several benefits, which are listed below:

- *Intrusion Detection and Threat Analysis*: XAI can help understand and explain the reasons behind a security breach or incident within a cloud ecosystem. By describing how a particular attack or threat was detected, security analysts can better understand the nature of the threat and take appropriate actions to prevent similar incidents in the future.
- *User Behavior Analysis*: XAI can assist in analyzing user behavior within a cloud environment to identify any abnormal or potentially malicious activities. XAI can help administrators and security teams make informed decisions about user access and privileges by explaining why specific actions are flagged as suspicious or risky.
- *Vulnerability Assessment*: XAI techniques can be used to evaluate the vulnerabilities and weaknesses in a cloud infrastructure. By explaining the factors that contribute to the identification of vulnerabilities, security experts can prioritize their efforts to patch or mitigate these weaknesses, thereby enhancing the overall security of the cloud ecosystem.
- *Model Auditing and Validation*: XAI enables security experts to audit and validate the effectiveness and reliability of AI models used within the cloud ecosystem.

Fig. 4 Key features of cloud IDS

5.1 Cloud Intrusion Detection System

Cloud Intrusion Detection System (IDS) is a security mechanism designed to detect and prevent unauthorized access, malicious activities, and potential threats within a cloud computing environment. The key features of the cloud IDS are shown in Fig. 4. It monitors network traffic, system logs, user behaviors, and other relevant data sources to identify suspicious or abnormal activities that may indicate a security breach.

The workflow of a cloud IDS typically involves several steps to monitor, analyze, and respond to potential security threats within a cloud environment. Figure 5 shows the workflow of cloud IDS.

- *Data Collection*: The cloud IDS collects the relevant data from various sources within the cloud ecosystem. Data can be network traffic logs, system logs, application logs, user activities, and different data points that provide activities within the cloud environment.
- *Preprocessing and Data Normalization*: The preprocessing and normalization are done to ensure consistency and compatibility. This is done by cleaning the information and changing extra information into a normalized design.
- *Signature-based and Anomaly Detection*: Signature-based detection techniques are applied to compare the collected data against known patterns or signatures

associated with known attacks. If a match is found, an alert is generated. Anomaly detection involves establishing baselines of normal behavior and comparing current observations against these baselines.

- *Machine Learning Analysis*: Machine learning algorithms are used to analyze data and identify potential threats. These algorithms learn from historical data to detect patterns, trends, and anomalies that indicate malicious activities. Machine learning models can continuously adapt and improve their detection capabilities over time.
- *Alert Correlation and Analysis*: The alert found is sent to security personnel or administrators for investigation. The cloud IDS correlates alerts and analyzes them collectively to gain a holistic view of the security landscape within the cloud environment. Security Information and Event Management (SIEM) platforms aggregate and analyze alerts from various IDS sensors and other security tools.
- *Continuous Monitoring and Feedback*: The cloud IDS continuously monitors the cloud environment for new threats and anomalies. In addition, it collects feedback from security analysts and administrators regarding the accuracy of alerts, false positives, and false negatives. This feedback is used to refine and improve the IDS's detection capabilities.

5.2 Research Challenges

Implementations of cloud-based IDS may face number of difficulties, mostly because of the peculiar qualities and complexity of cloud environments. Here are some typical challenges that cloud IDS encounters:

- *Scalability*: With many virtual machines, containers, and services, cloud environments are dynamic and can scale quickly. A significant challenge can be ensuring that the cloud IDS can manage the high volume of network traffic and monitor all the resources in a scalable manner.
- *Network Complexity*: Distributed networks, virtual networks, and microservices architectures are frequent components of cloud environments, which can make it difficult to monitor and analyze network traffic. It can be challenging to capture and analyze network traffic across various components and services due to the distributed nature of cloud infrastructure. IDS solutions must be able to decrypt and examine encrypted traffic without jeopardizing security or infringing on privacy laws.
- *False Positives and False Negatives*: It is difficult to maintain the accuracy of intrusion detection while minimizing false positives (which mistakenly classify the benign activity as a threat) and false negatives (which fail to detect real threats). It can be challenging to balance IDS rules and machine learning models.

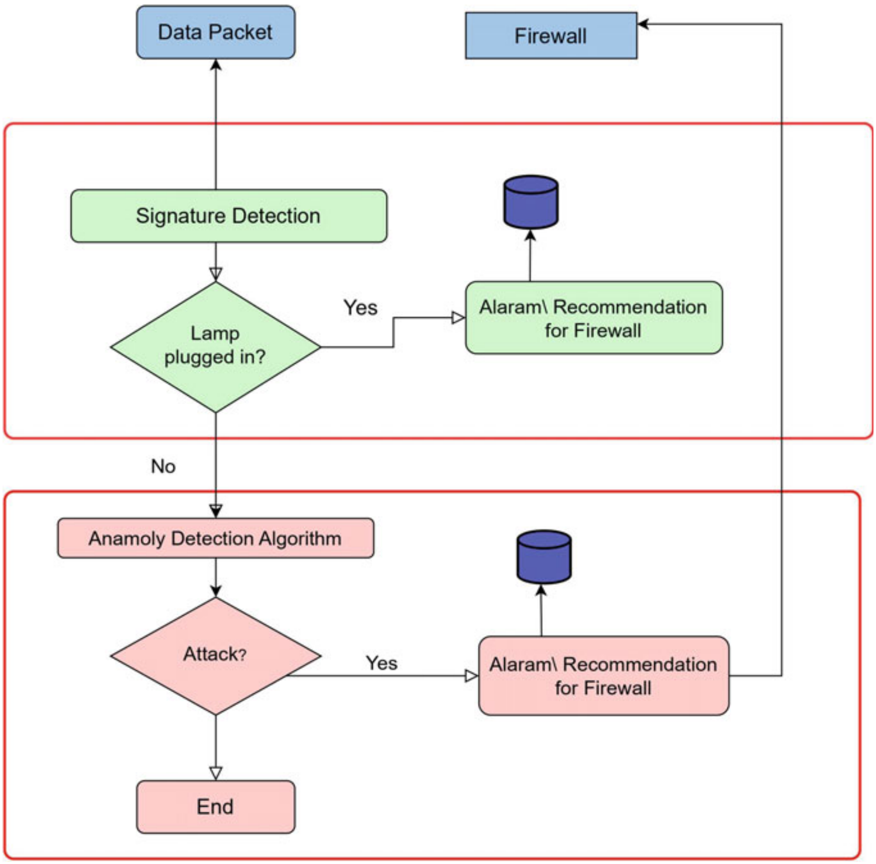


Fig. 5 Workflow of cloud IDS [17]

6 Case Study

In this section, the author discusses a case study MEDITECH [18] that adds advanced security to its cloud-based healthcare with cloud IDS. There were many challenges that MEDITECH faced regarding paper-based documentation, limited data accessibility of the patient, and inefficient workflow. So the solution for these challenges is electronic health record (EHR) systems. MEDITECH creates EHR to improve the interactions between clinicians and patients. The business enables large and small healthcare organizations to provide safe, affordable patient care. Software for health information management, patient care and patient safety, emergency department management, oncology, genomics, population health, laboratories, blood banks, revenue cycle management, home health, virtual care, and many other areas of healthcare are all included in MEDITECH's user-friendly and portable offerings.

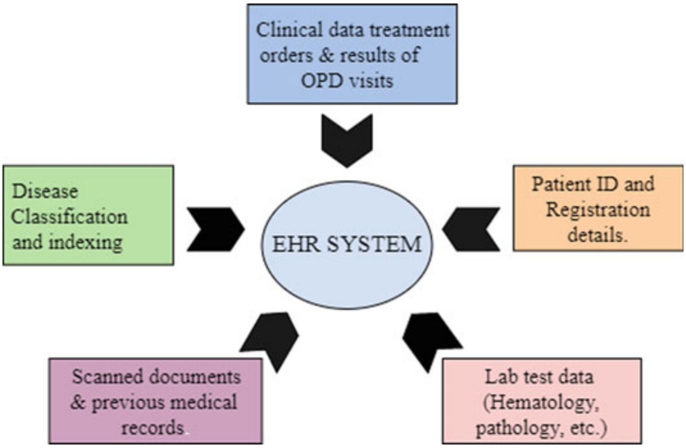


Fig. 6 EHR system by MEDITECH [19]

The implementation of EHR systems offered by MEDITECH aimed to address these challenges by digitizing and centralizing patient information, improving data accessibility, streamlining workflows, enhancing data security, and empowering patients to engage in their care. The key features of the EHR System are shown in Fig. 6. EHR systems have significantly transformed healthcare by addressing these problems and enabling more efficient, effective, patient-centered care delivery.

Health care, including the MEDITECH platform, requires an EHR system for many reasons, which are listed below:

Centralized and Accessible Patient Information: EHR systems provide a centralized repository for storing and accessing patient health information. Healthcare providers can quickly and easily retrieve patient data such as medical history, diagnoses, medications, allergies, laboratory results, and imaging reports. This accessibility improves care coordination and enables informed decision-making.

Improved Patient Safety: EHR systems help enhance patient safety by reducing errors and promoting accurate and timely documentation. Features such as computerized physician order entry (CPOE) and decision support systems can help prevent medication errors, alert healthcare providers to potential drug interactions or allergies, and support evidence-based care practices.

Efficient Workflow and Documentation: EHR systems streamline healthcare workflows by automating various tasks, including patient registration, scheduling, documentation, and billing. Electronic documentation replaces paper-based records, reducing the need for manual data entry and paperwork. This automation saves time, improves efficiency, and reduces the risk of errors associated with manual processes.

Data Analytics and Population Health Management: EHR systems support data analytics and population health management initiatives. By aggregating and analyzing patient data at a larger scale, healthcare organizations can identify trends, patterns, and insights for population health management, disease surveillance, quality improvement, and research purposes.

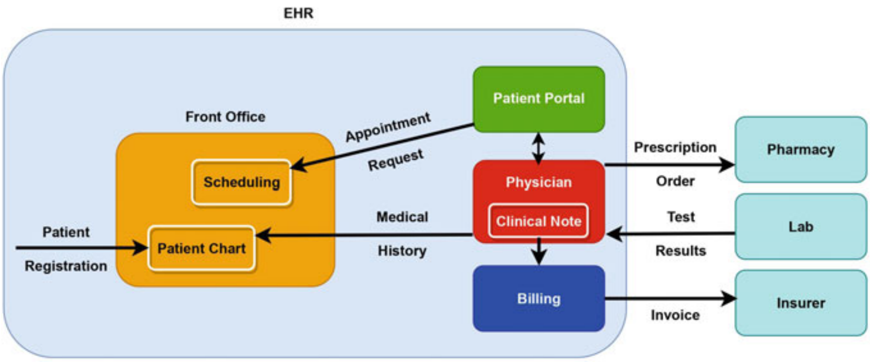


Fig. 7 Workflow of EHR system [20]

MEDITECH works to meet or exceed the requirements of industry standards and security frameworks. The whole workflow of the EHR system is described in Fig. 7. Additionally, it supports the organization's requirements for compliance. As a result, MEDITECH will be able to keep providing its healthcare clients and their patients with cutting-edge, affordable, and secure technology.

7 Open Research Directions

Despite significant progress, there are still open research issues in developing secure and explainable AI for cloud ecosystems. Creating more robust and efficient secure AI techniques that can withstand sophisticated attacks such as model poisoning and evasion attacks is necessary. Also, there is a need to develop more advanced explainable AI techniques that can handle complex AI models such as deep neural networks. In addition, it encourages techniques that support computer-based intelligence models' planning and prediction while preserving the security of sensitive data stored in the cloud. This might include secure multi-party calculation, homomorphic encryption, or differential security to safeguard hidden information. Finally, there is a need to develop standards and guidelines for secure and explainable AI in cloud ecosystems to promote interoperability and adoption.

8 Conclusion

Secure and explainable AI is critical for building trustworthy and robust AI-based systems in cloud ecosystems. Recent research efforts have significantly progressed toward developing secure and explainable AI techniques and frameworks. However, there are still open research directions in this field, and addressing these challenges can significantly impact the security and transparency of AI-based systems in cloud ecosystems. The case study, however, also emphasizes the complex challenges and factors involved in creating and implementing AI-based solutions for regulatory compliance. The complexity of regulations, the quality and availability of data, interoperability with legacy systems, and the explainability of AI decisions are all significant challenges that must be addressed to ensure that the solution is efficient and meets these needs.

References

1. S. R. Department (2022) Most common cloud security attacks worldwide (Online). https://www.statista.com/statistics/1320178/common-cloud-security-attacks-worldwide. Accessed 19 May 2023
2. M. V. N. A. M. AI and Cloud. The flow of data in an AI-driven user interface (Online). https://www.juniper.net/us/en/research-topics/what-is-explainable-ai-xai.html. Accessed 19 May 2023
3. Bayuk J (2011) Cloud security metrics,. In 2011 6th international conference on system of systems engineering, IEEE, pp 341–345
4. Arora A, Khanna A, Rastogi A, Agarwal A (2017) Cloud security ecosystem for data security and privacy. In: 2017 7th international conference on cloud computing, data science & engineering—confluence, pp 288–292
5. Gochhait S, Chinta DSR (2020) Leveraging AI for secure public cloud. Eur J Molecular Clin Med 7(6):2314–2323
6. Liu W (2012) Research on cloud computing security problem and strategy. In: 2012 2nd international conference on consumer electronics, communications and networks (CECNet), IEEE, pp 1216–1219
7. McMahan B, Moore E, Ramage D, Hampson S, y Arcas BA (2017) Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics, PMLR, pp 1273–1282
8. Mohandas S (2021) The role of AI in cloud computing
9. Minh D, Wang HX, Li YF, Nguyen TN (2022) Explainable artificial intelligence: a comprehensive review. Artif Intell Rev 1–66
10. Javed AR, Ahmed W, Pandya S, Maddikunta PKR, Alazab M, Gadekallu TR (2023) A survey of explainable artificial intelligence for smart cities. Electronics 12(4):1020
11. Wu Z, Li Q, He B (2022) Practical vertical federated learning with unsupervised representation learning. IEEE Trans Big Data
12. Papastefanopoulos V, Kotsiantis S et al (2021) Explainable AI: a review of machine learning interpretability methods. Entropy 23(18):1–45
13. Kok I, Okay FY, Muyanli O, Ozdemir S (2022) Explainable artificial intelligence (xai) for internet of things: a survey. arXiv preprint arXiv:2206.04800
14. Nazir S, Dickson DM, Akram MU (2023) Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks. Comput Biol Med 106668
15. Atakishiyev S, Salameh M, Yao H, Goebel R (2021) Explainable artificial intelligence for autonomous driving: a comprehensive overview and field guide for future research directions, arXiv preprint arXiv:2112.11561
16. Pandian S (2023) Explainable Artificial Intelligence (XAI) for AI & ML Engineers (Online). https://www.analyticsvidhya.com/blog/2022/10/explainable-artificial-intelligence-xai-for-ai-ml-engineers. Accessed 19 May 2023
17. Rebahi Y (2015) Virtual security appliances: the next generation security . https://www.researchgate.net/figure/IDS-process-flow-diagram-The-different-components-of-the-architecture-interwork-in-the_fig3_283256631

18. Gregory TM, Lebovitz M (2021) How MEDITECH adds advanced security to its cloud-based healthcare solutions with Cloud IDS (Online). https://cloud.google.com/blog/products/identity-security/how-meditech-uses-google-cloud-id. Accessed 19 May 2023
19. iTech (2020) EHR vs EMR. Why the difference is important in healthcare (Online). https://itechindia.co/blog/ehr-vs-emr-why-the-difference-is-important-in-healthcare. Accessed 19 May 2023
20. Altexsoft (2019) Electronic health record systems: features, EHR vendors, and adoption advice (Online). https://www.altexsoft.com/blog/electronic-health-record-systems. Accessed 19 May 2023

Skin Cancer Image Augmentation Techniques Using AI: A Survey of the State-of-the-Art

Aniket Y. Patil, Yashwant S. Ingle, Nuzhat Faiz Shaikh, Parikshit Mahalle, and Janki Barot

Abstract Even if more and more high-quality public datasets are available, one of the biggest problems with deep learning for skin lesion diagnosis is still the paucity of training samples. On several computer vision tasks, deep convolutional neural networks have displayed impressive performance. While the models developed by these algorithms often outperform more conventional machine learning techniques, they require more extensive datasets to be accessible for training. Data augmentation has emerged as a common technique for addressing this problem, especially in domains where massive datasets are not frequently accessed, which is frequently the case when working with medical imaging. The goal of data augmentation is to create more data used to train the model, and it has been found to enhance performance when verified on a different, unexplored dataset. This survey devotes much time to using augmentation techniques based on Basic Data Augmentation algorithms, GANs, and VAE. Data Augmentation, a data-space solution to the issue of limited data, is the main topic of this survey. This study seeks to provide insight into these methodologies and confidence in the validity of the models generated as artificial intelligence models trained with augmented data find their way into the clinic.

Keywords Data augmentation · Skin cancer · Deep learning

A. Y. Patil · Y. S. Ingle (✉) · P. Mahalle
Vishwakarma Institute of Information Technology, Pune, India
e-mail: yashwant.ingle218@gmail.com

P. Mahalle
e-mail: parikshit.mahalle@viit.ac.in

Y. S. Ingle · N. F. Shaikh
SKNCOE, Pune, India

J. Barot
Silver Oak University, Ahmedabad, India

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_52

1 Introduction

Today, the most significant cause of mortality in the world is cancer. More than half of cancer diagnoses globally are for skin cancer [1]. Every year, globally, approximately 2–3 million cases of non-melanoma skin cancer are identified, while around 130,000 cases of melanoma skin cancer are reported [2]. Late disease identification and therapy are the primary causes of skin cancer fatalities. The two primary categories it frequently separates are melanoma and non-melanoma skin cancer [3]. Early detection and treatment are the most efficient ways to prevent skin cancer-related mortality. An increasing number of computer-aided diagnosis methods are being created to classify skin cancer due to the fast growth and development of computer hardware and computer vision algorithms. The deep learning-based convolutional neural network (CNN) technique is beneficial in this area.

The CNN models' successful outcomes urge scientists to develop models with better accuracy. Building more complicated designs typically produces these consequences [4]. Remember that the quantity of trainable parameters is frequently used to quantify model complexity. A model's complexity increases with the number of trainable parameters it contains. One of the most evident issues when utilizing complex CNN models is the overfitting problem [5], defined as the performance difference between the training and validation/test stages when the model loses its ability to generalize. Overfitting typically happens when a model is too sophisticated for the data or the data are insufficient [6]. At the model architecture level, regularization approaches such as dropout [7], ridge regression (l_2 regularization) [8], and Lasso regression (l_1 regularization) [9] are used. These methods' primary goal is to make neural network models less complicated during training, which is thought to be the leading cause of overfitting, especially when the model is trained on short datasets. These methods might be seen as results of the ongoing struggle to achieve better performance through the invention of increasingly sophisticated deep neural networks, such as VGG-16 [10], ResNet [11], Inception-V3 [12], and DenseNet [13].

CNN works well for picture identification, but many practical implementations require large amounts of data for accuracy [14]. ISIC, HAM10000 is studied (see Fig. 1). Academics are using data augmentation and transfer learning to handle the data challenge.

2 Motivation

Image segmentation aims to find a group of pixels or voxels that are comparable in an area of interest (ROI). As a result, clinical characteristics relating to volume and form may be quantitatively analyzed using medical image segmentation. Deep learning segmentation algorithms on medical pictures are frequently assessed by contrasting the findings with images that clinical experts manually segmented. Since

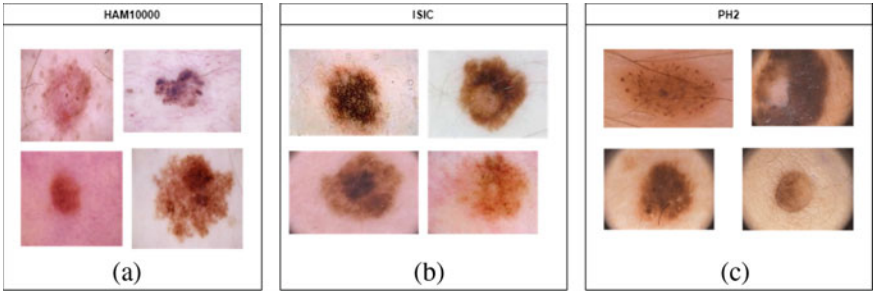


Fig. 1 **a** HAM10000 dataset collects 10,015 photos of pigmented skin lesions gathered from various sources. **b** Over 26,000 labeled pictures of skin lesions, including melanoma, seborrheic keratosis, and nevi, are included in the International Skin Imaging Collaboration (ISIC) dataset. **c** The PH^2 dataset consists of 200 dermoscopic pictures of benign and cancerous pigmented skin lesions, including melanomas, that were gathered from the dermatology department of Hospital Pedro Hispano in Portugal

labeling tagged data for model training and assessment are a labor-intensive and time-consuming activity due to the pixel-level nature of segmentation, data augmentation approaches (see Fig. 2) are crucial for developing segmentation algorithms in this industry. Data augmentation has been used in several studies to classify melanoma, including Matsunaga et al. [15], González-Daz [16], Menegola et al. [17], and Esteva et al. [18].

In comparison to sliding window CNN segmentation techniques [19], U-net, a well-liked convolutional neural network, performs better. Elastic deformation is used in U-net design, which ensures superior outcomes by producing more realistic variations for data augmentation. In picture segmentation, elastic deformation is also employed to prevent overfitting [20]. The general process of data augmentation using convolution neural architecture is shown in Fig. 2.

These approaches, known as data augmentation, are frequently employed in medical applications to build more accurate deep learning models from tiny, labeled datasets. This approach creates new data items by fundamentally modifying existing training samples [21]. To address the overfitting problem, the best way to reduce the gap between training and testing datasets is to create enhanced data through data augmentation. According to Shorten's work [4] on data augmentation, CNN may be better by overcoming data-space limitations, including limited dataset sizes and

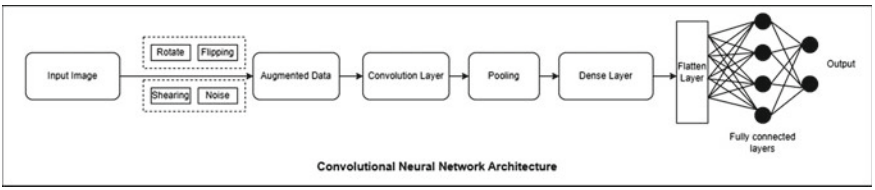


Fig. 2 General process of data augmentation using convolutional network architecture

unequal data distribution. Here are the examples: Random erasing was suggested to solve the occlusion issue [22], rotation and flipping to partially solve the viewpoint issue [23–25], brightness to solve the lighting issue [26], and cropping and zooming to solve the scaling and background problems.

3 Literature Survey

This section covers data augmentation strategies, which researchers utilize. Data augmentation improves training datasets by changing samples. Researchers have examined many augmentation methods to increase machine learning model functionality and generalization. Generative models, text, audio, and time series data augmentations, color and lighting changes, image distortions, and geometric changes are used. The wide range of data augmentation methods researchers utilize can teach us about cutting-edge strategies for supplementing and enriching training data in numerous fields. Basic Data Augmentation and Generative Models approaches studied.

3.1 Basic Augmentation

To improve images, this section includes cropping, flipping, affine transformations, and color disturbance. One image from the dataset is edited and added back to the data pool to increase its size.

Geometric transformation. Most fundamental augmentations employ geometric transformation. Scaling, translating, rotating, reflecting, shearing, and temporal perspective transforms, including skew, are the most prevalent. The image's orientation and direction are altered. This method preserves picture feature dimensions during transformation, eliminating data loss.

Image cropping. In order to crop an image, it is usual practice to select a random window size from the original image and scale it to a predetermined size. It is typically used when there is a class imbalance. Random cropping can be used to provide an effect very similar to translation. Depending on the extent of cropping reduction used, this adjustment could not preserve the label. After scaling, each pixel's coordinates change, while x and y stand for the pixel's original location in the picture. Cropping pictures to a specific size before training is a common practice training data may contain examples of varying sizes [4, 27].

Noise injection. A popular definition of noise is a random fluctuation in color or brightness information [28]. Noise injection aims to simulate noisy images. It could seem difficult for neural networks when there is data noise. Data from the real world are rarely flawless [29]. Noise can reduce neural networks' accuracy and generalization performance when tested on real-world data. At the absolute least, the data used to train deep learning models cannot be as clean as the data used to test them. The three well-known noise types that may enhance image data are speckle, salt-and-pepper, and Gaussian [30]. Deep learning models may be resistant to some forms of noise when they are used to enhance data.

Color transformations. By manipulating the color channels of the images, color transformation is performed in RGB color spaces. Images can have varied color casts by altering the ratio between each color channel. Create a color histogram that depicts the image for more complex color alterations. Similar to how photo editing software handles it, the intensity values of these histograms may be altered to adjust the lighting.

Random erasing. A data augmentation approach called random erasing [31] avoids broad changes to specific picture pixels' values. Instead, it substitutes a random value for the pixel values inside a randomly sized area in the image. Unlike the other techniques of data augmentation discussed above, it increases the data variety holistically without growing the amount of the data.

Combination. The combination is a data augmentation technique that combines two or more original photos to create a new image. Ionue [32] gave an example of how the pairing of samples could be turned into a successful enhancement technique. Summers and Dinneen [33] looked further into the idea of blending images in an illogical fashion. They considered combining photos using nonlinear techniques to create fresh training examples.

Kernel filters. Image processing sharpens and blurs. Kernel filters are popular. Applying a high-contrast vertical or horizontal edge filter or a Gaussian blur filter to a matrix over a picture sharpens or blurs the edges. Data augmentation blurring may improve motion blur resistance during testing. Data augmentation employing kernels of varying values may also create visuals with specific features [4]. Enhancing photographs for data augmentation may also capture extra information about fascinating objects. The Sobel [34] or Canny [35] filters may detect edges, whereas high-contrast vertical or horizontal edge filters can sharpen or blur them. This method's downside is that it mimics CNNs' internals. Thus, network layers are better for kernel filter construction than dataset augmentation.

3.2 Generative Models

Generative modeling is a fascinating method of data augmentation. Generative modeling is the process of constructing artificial instances from a dataset while maintaining the original set's features. The goal of generative modeling using neural networks is to learn about data without supervision, which might be advantageous for conventional classification problems. Collecting training data for unsupervised learning is inherently easier and less expensive.

GAN. The extremely intriguing and enormously well-liked generative modeling framework known as GANs was developed using the adversarial training ideas previously outlined. Since its inception, GANs have emerged as one of the most important advancements in generative deep learning approaches [36]. Researchers

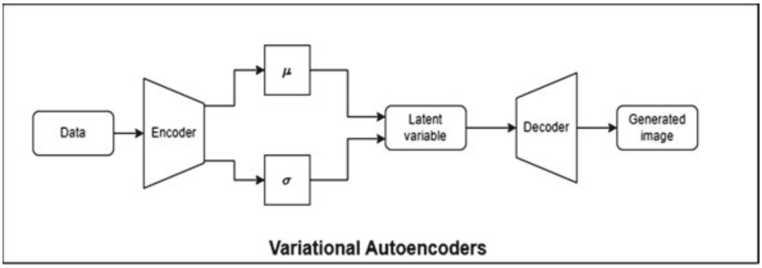


Fig. 3 Process of data augmentation using variational autoencoder (VAE)

studying medical image processing have been interested in Generative Adversarial Networks (GANs) [37], one of several generative models. Due to their superior performance compared to other generative models, GANs have recently been applied in the medical image (MI) generation [38, 39], medical image editing in latent space [40], MI segmentation [41], and MI classification. The generator must be able to produce data in a way that prevents the discriminator from telling the difference between fake and real.

VAE. Generative modeling with variational autoencoders is worthwhile. The GAN framework can improve variational autoencoder samples. VAEs, or variational autoencoders (Fig. 3), may represent data points in a low-dimensional fashion.

KL divergence is used to compare the estimated posterior distribution and a prior distribution across latent space. The ELBO helps train VAEs since the reconstruction loss and KL divergence components must be adjusted for effective data representation.

In order to enhance the self-supervised learning performance on skin cancer classification tasks for teledermatology, Whitney et al. [42] looked at how to use GAN-generated skin cancer pictures. BCN20000 and HAM10000 will be used to test the suggested methodology. They initially use unlabeled data to train StyleGAN to produce high-quality skin cancer pictures. Then, using an enhanced training dataset that comprises StyleGAN-generated pictures and annotated training images, trained a feature encoder using self-supervised learning.

4 Challenges and Opportunities

Dermoscopy, which uses polarized light to magnify things, is now used by dermatologists to evaluate patients. The patient's medical history, ethnicity, social habits, and sun exposure all affect the diagnosis. Suspicious lesions are surgically removed and sent to the lab for permanent paraffin section processing and pathologist assessment on representative glass slides after an office biopsy.

AI-enabled CAD technologies are poised to revolutionize medicine and health care, particularly in the area of medical imaging. According to studies [18, 43], AI algorithms for disease identification in medical imaging are on par with or even superior than physician performance. The International Skin Imaging Collaboration (ISIC), which provides digital datasets of skin lesion images with professional annotations, is a driving factor behind automated CAD solutions for the identification of melanoma and other skin cancers.

4.1 Challenges

In a short amount of time, deep learning algorithms have outperformed the benchmarks of well-known computer vision datasets; the same pattern may be anticipated for the skin lesion diagnostic problem. To increase diagnosis accuracy, deep learning algorithms often need a sizable quantity of training data that represents each class of skin lesions and is varied, balanced, and of good quality. There are many more problems with diagnosing skin cancer using AI solutions for datasets of skin lesions of different modalities, as detailed below.

- Deep learning performance with Imbalanced Datasets

Despite tweaking strategies like a penalty for false negatives in tiny skin lesion classes during training using tailored loss functions, unbalanced datasets often hurt deep learning system performance. Most skin lesions are benign. Most deep learning architectures use balanced datasets like ImageNet, which includes 1000 classes and 1000 images per class [40].

- Clinical meta-data and medical history of patients

The patient's and family's history of skin cancer, age, ethnicity, sex, general anatomic location, size, and form of the suspicious skin lesion, and dermoscopy must be utilized together. Thus, only image-based deep learning algorithms used to identify skin cancer are unsuccessful for patient and clinical data.

- Differences within classes

Skin lesions within a class vary in color, attribute, texture, size, and location. Thus, these skin lesions are subdivided by appearance. Melanin's pigment makes most melanomas black. Some melanomas are reddish or pink and have normal skin tone. BCC has numerous subtypes, including nodular BCC, superficial BCC, morphea form BCC, and basosquamous carcinoma, which can range from white to red.

4.2 Opportunities

These performance scores do not match physicians' skin cancer diagnosis due to the concerns discussed above. Deep learning techniques are opaque because they only learn from imaging dataset pixel values and lack domain knowledge or logical inference abilities to determine the association between distinct skin states [44]. However, deep learning may soon identify skin cancer well.

5 Discussions

Data augmentation efforts support industries and machine learning models. List the survey's top data augmentation methods. Add noise, occlusion, scaling, flipping, rotation, and translation. Consider implementation simplicity, flexibility to different tasks, and model generalization endurance when assessing their appeal. Explore machine learning data augmentation methods. Explain why some methods are better for time series analysis, natural language processing, picture categorization, and object recognition. Survey task-specific augmentation methods.

Survey responses showed domain-specific traits. Discussed survey data augmentation. Combining methods may increase model robustness, variety, or realism. Show how combining approaches improves model performance.

Most studies supplement data using many ways. This makes assessing each strategy tough. Most GPU tests accelerate training and deployment. Preprocessing improves photo quality. GANs and variational autoencoders learn complicated representations. Pioneer deep CNN visualization [45, 46]. Humanizing convolutional networks may aid augmentation. Smartphone applications can instantly measure skin cancer risk and identify and treat patients' early, boosting survival.

6 Conclusion

Data augmentation improves datasets. In this study, we examined data augmentation approaches that help deep learning models overcome overfitting and perform better, particularly in picture segmentation and classification. Picture segmentation precision was measured by the Jaccard index, dice coefficient, sensitivity, specificity, and accuracy. Combining all tactics can improve data-hungry deep learning systems.

Comparing several of the procedures to others that did not meet the necessary standards, such as due to small availability of sample sizes, situations where the population standard deviation is unknown, or both, make the t-distribution most helpful, the type of training data, models, and augmentation rules used affect TTA performance. GAN-based data augmentation indicates toward aiming to balance the benefits of translation-based and noise-based techniques obtaining the high quality of the former and the unlimited sample availability of the latter even if such conciliation is a challenging unsolved task. Improved sampling techniques that can choose the highest-quality, or better yet, most pertinent samples from the infinite sample of a noise-based methodology, may be more feasible. But with data augmentation, several biases, including those caused by lighting, occlusion, scale, backdrop, and many more, may be avoided or at the very least significantly reduced.

References

1. Codella NC et al (2017) Deep learning ensembles for melanoma recognition in dermoscopy images. IBM J Res Dev 61(4/5):5:1–5:15
2. Almaraz-Damian J-A, Ponomaryov V, Sadovnychiy S, Castillejos-Fernandez H (2020) Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures. Entropy 22(4):484
3. Elgamal M (2013) Automatic skin cancer images classification. Int J Adv Comput Sci Appl 4(3)
4. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1):1–48
5. Brownlee J (2018) Better deep learning: train faster, reduce overfitting, and make better predictions. In: Machine learning mastery
6. Guo Y, Cai J, Jiang B, Zheng J (2018) Cnn-based real-time dense face reconstruction with inverse-rendered photo-realistic face images. IEEE Trans Pattern Anal Mach Intell 41(6):1294–1307
7. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15(1):1929–1958
8. Farebrother RW (1976) Further results on the mean square error of ridge regression. J R Stat Soc: Ser B (Methodol) 38(3):248–250
9. Ranstam J, Cook J (2018) LASSO regression. J Br Surg 105(10):1348–1348
10. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
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 770–778
12. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826
13. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708
14. Mustafa S. (2017) Feature selection using sequential backward method in melanoma recognition. In: 2017 13th international conference on electronics, computer and computation (ICECCO), IEEE, pp 1–4
15. Matsunaga K, Hamada A, Minagawa A, Koga H (2017) Image classification of melanoma, nevus and seborrheic keratosis by deep neural network ensemble. arXiv preprint arXiv:1703.03108
16. Díaz IG (2017) Incorporating the knowledge of dermatologists to convolutional neural networks for the diagnosis of skin lesions. arXiv preprint arXiv:1703.01976
17. Menegola A, Fornaciali M, Pires R, Bittencourt FV, Avila S, Valle E (2017) Knowledge transfer for melanoma screening with deep learning. In: 2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017), IEEE, pp 297–300

18. Esteva A et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118
19. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI2015: 18th international conference, Munich, Germany, 5–9 Oct 2015, Proceedings, Part III 18. Springer, Heidelberg, pp 234–241
20. Quan TM, Hildebrand DGC, Jeong W-K (2021) Fusionnet: a deep fully residual convolutional neural network for image segmentation in connectomics. Front Comput Sci 3:613981
21. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
22. Zhong Z, Zheng L, Kang G, Li S, Yang Y (2020) Random erasing data augmentation. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, no 07, pp 13001–13008
23. Divon G, Tal A (2018) Viewpoint estimation—Insights & model. In: Proceedings of the European conference on computer vision (ECCV), pp 252–268
24. Ning X, Nan F, Xu S, Yu L, Zhang L (2020) Multi-view frontal face image generation: A survey. Concurr Comput: Pract Exp, e6147
25. Massa F, Marlet R, Aubry M (2016) Crafting a multi-task CNN for viewpoint estimation. arXiv preprint arXiv:1609.03894
26. Liu H, Wang C, Peng Y (2021) Data augmentation with illumination correction in sematic segmentation. J Phys: Conf Ser 2025(1):012009
27. Lu W, Xing X, Cai B, Xu X (2019) Listwise view ranking for image cropping. IEEE Access 7:91904–91911
28. Ravishankar A, Anusha S, Akshatha H, Raj A, Jahnavi S, Madhura J (2017) A survey on noise reduction techniques in medical images. In: 2017 international conference of electronics, communication and aerospace technology (ICECA), vol 1: IEEE, pp 385–389
29. Nazaré TS, da Costa GBP, Contato WA, Ponti M (2018) Deep convolutional neural networks and noisy images. Progress in pattern recognition, image analysis, computer vision, and applications: 22nd Iberoamerican Congress, CIARP 2017, Valparaíso, Chile, 7–10 Nov 2017, Proceedings 22. Springer, Heidelberg, pp 416–424
30. Boonprong S, Cao C, Chen W, Ni X, Xu M, Acharya BK (2018) The classification of noise-afflicted remotely sensed data using three machine-learning techniques: Effect of different levels and types of noise on accuracy. ISPRS Int J Geo Inf 7(7):274
31. Boyat AK, Joshi BK (2015) A review paper: noise models in digital image processing. arXiv preprint arXiv:1505.03489
32. Inoue H (2018) Data augmentation by pairing samples for images classification. arXiv preprint arXiv:1801.02929
33. Summers C, Dinneen MJ (2019) Improved mixed-example data augmentation. In: 2019 IEEE winter conference on applications of computer vision (WACV), IEEE, pp 1262–1270
34. Kanopoulos N, Vasanthavada N, Baker RL (1988) Design of an image edge detection filter using the Sobel operator. IEEE J Solid-State Circuits 23(2):358–367
35. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 6:679–698
36. Mutepfe F, Kalejahi BK, Meshgini S, Danishvar S (2021) Generative adversarial network image synthesis method for skin lesion generation and classification. J Med Signals Sens 11(4):237
37. Goodfellow I et al (2014) Adv Neural Inf Process Syst
38. Shahsavari A, Ranjbari S, Khatibi T (2021) Proposing a novel Cascade Ensemble Super Resolution Generative Adversarial Network (CESR-GAN) method for the reconstruction of super-resolution skin lesion images. Inform Med Unlocked 24:100628
39. Shen T, Hao K, Gou C, Wang F-Y (2021) Mass image synthesis in mammogram with contextual information based on GANs. Comput Methods Programs Biomed 202:106019
40. Fernández R, Rosado P, Vegas Lozano E, Reverter Comes F (2021) Medical image editing in the latent space of Generative Adversarial Networks. Intell-Based Med 5:100040.
41. Zhang J et al (2021) Dense GAN and multi-layer attention based lesion segmentation method for COVID-19 CT images. Biomed Signal Process Control 69:102901

42. Ren Z, Guo Y, Stella XY, Whitney D (2021) Improve image-based skin cancer diagnosis with generative self-supervised learning. In: 2021 IEEE/ACM conference on connected health: applications, systems and engineering technologies (CHASE), IEEE, pp 23–34
43. Liu X et al (2019) A comparison of deep learning performance against healthcare professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health 1(6):e271–e297
44. Marcus G, Davis E (2019) Rebooting AI: building artificial intelligence we can trust. Vintage
45. Wei JW, Tafe LJ, Linnik YA, Vaickus LJ, Tomita N, Hassanpour S (2019) Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. Sci Rep 9(1):3358
46. Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H (2015) Understanding neural networks through deep visualization. arXiv preprint arXiv:1506.06579

Detection of Quality and Shelf Life of Fruits and Vegetables Using Neural Networks

Priyanka D. More, Sachin R. Sakhare, Aditya U. Joshi, Lokesh P. Gaikwad, Vinayak P. Jakite, and Anom S. Warbhuvan

Abstract The quality and shelf life of fruits and vegetables affects the consumer's purchasing decision. In recent years, the use of neural networks for detecting the quality and predicting the life expectancy of foods has gained significant attention. In this research paper, we propose using neural networks for the same. We present a detailed analysis of the different neural network architectures and training techniques that have been used for this purpose and provide a comprehensive comparison of their performance. Our experimental results demonstrate the efficacy of this approach. In this research paper, we propose a solution for the detection using neural networks. We use the VGG-16 model, a deep learning architecture that has been widely used in image classification tasks, to classify the images of fruits and vegetables. The model training process involved the collection and annotation of a large dataset of images of tomatoes and potatoes at different stages of ripeness. We used this dataset to train the VGG-16 model, optimizing the hyperparameters through a grid search process. The results of our experiments showed that the VGG-16 model was able to achieve an accuracy of around 80% in the classification of the test images provided. This demonstrates the potential of neural networks for the automated detection of the ripeness and life of these perishable products, which can be a great practical use case to the food industry.

P. D. More (✉) · S. R. Sakhare · A. U. Joshi · L. P. Gaikwad · V. P. Jakite · A. S. Warbhuvan
Vishwakarma Institute of Information Technology, SPPU, Pune, India
e-mail: priyanka.more@viit.ac.in

S. R. Sakhare
e-mail: sachin.sakhare@viit.ac.in

A. U. Joshi
e-mail: aditya.21910991@viit.ac.in

L. P. Gaikwad
e-mail: lokesh.21911020@viit.ac.in

V. P. Jakite
e-mail: vinayak.21911071@viit.ac.in

A. S. Warbhuvan
e-mail: anom.21910758@viit.ac.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_53

Keywords Neural networks · Quality detection · Shelf-life detection · Machine learning

1 Introduction

India is a global agricultural powerhouse. Around 54.6% of the population does agriculture or related activities (according to the census 2011) and it makes around 17.4% to the country's Gross Value Added. When it comes to tomatoes, there are different types that are grown for different purposes.

In terms of quality, the tomatoes that are used for daily use, canning, or ketchup need to meet certain standards. They need to be ripened properly and should not have any blemishes. The tomatoes also need to be of the right size. Too small and they will not have enough flesh and too large and they will have too much water. The classification process is important in order to ensure that only the best tomatoes are used. There are a few different steps involved in the classification process. The first step is to sort the tomatoes by their stages (unripe, partially ripe, and ripe). The next step is to inspect the tomatoes for any blemishes. The final step is to grade the tomatoes based on their association with either of the categories defined [1]. Color is one of the most important factors in the classification process. The color of the tomato is an indication of its ripeness. The goal is to use only ripe tomatoes. There are three different colors that are typically used in the classification process. These colors are red, pink/green, and brown. Red tomatoes are typically the ripest and are the best choice. Pink/green tomatoes are not as ripe as red tomatoes, but they are still a good choice. Brown tomatoes are the least ripe and are typically only used if there is a shortage of the other two. Figure 1 represents images of good and damaged tomatoes and potatoes.

In this research paper, we will investigate the methods used to check the quality and shelf life of potatoes and classify them for use. There are a few different ways to check the quality of potatoes. One method is to look at the potato and inspect it for any blemishes or irregularities [2]. Another method is to use machine learning algorithms to classify the potatoes based on color, the presence of buds, and damaged spots or irregularities. The flesh should be white or pale yellow. If the flesh is brown

Fig. 1 **a** Good tomato, **b** Damaged tomato, **c** Good quality potato, and **d** Bad quality potato

or discolored, it is likely that the potato is of poor quality [3]. It is important to check the quality of potatoes before using them in any dish.

We also introduce a method for detecting the shelf life of fruits and vegetables using neural networks. Our approach utilizes high-resolution images of the produce and applies a CNN to classify the freshness of the products.

Machine learning can help with this by automatically classifying fruits and vegetables [4]. This can be done by taking pictures of fruits and vegetables and using a machine learning algorithm to identify them [5]. This research paper will discuss how machine learning (neural networks) can be used to classify, detect ripeness, and predict the expiry of fruits and vegetables (here, tomato and potato are used as an example) and how this can help people eat a healthier diet. Machine learning is basically teaching computers to learn from data. The benefits of using machine learning in this experiment include being able to help people eat a healthier diet, helping farmers to control inventory by predicting the shelf life of their products, and automatically checking the quality of fruits and vegetables. This can help people know if they are eating enough fruits and vegetables, and it can also help them choose healthier options.

2 Existing Related Solutions

This proposed system in the paper detects defective tomatoes by converting the image to an HSV scale and then negating it to mask the colors and identify the presence of damaged spots on the tomato. The use of the HSV scale makes it easier to identify the defect on the upper body of the tomatoes [6].

Supriya [7] in the paper uses the same principle as [6] and identifies the defect in the fruits. The damage appears darker than the shade of the given fruit and hence is represented as "white" on the HSV scale. This makes it easier to identify the damage on the outer skin of the fruit [7].

In this paper, various models are developed using the RGB scale to classify tomatoes. These modules detect the stage of the tomatoes based on colors. Due to the use of multiple modules, the accuracy of the system is high with an error margin of 0.000536995 [8].

The paper "Tomatoes Classifier using Color Histograms" also suggests the use of the HSV scale to create a Histogram and then map and classify the tomatoes. Naive Bayes Theorem is used to create the machine learning model [9].

Wasule [9] in the paper uses the same approach of the HSV scale. The authors were able to achieve an accuracy of 95% for this system [10].

The proposed system uses the CNN model to detect the freshness of fruits. A CNN model is trained on various images and it is used to predict the freshness of fruits [11].

The proposed system uses a flutter app and uses mango fruit as an example. It captures two images, one through a rear camera of smartphone and the other through a Seek thermal camera. The model predicts the shelf life of the fruits with an accuracy of about 90% [12].

3 Proposed System

Here are the steps in the system flowchart (Fig. 2):

1. Importing libraries: This step involves importing all the necessary libraries and modules required, such as VGG-16, NumPy, and Matplotlib.
2. Creating features with VGG-16: In this step, the VGG-16 model is used to extract features from the images of fruits and vegetables in the dataset.
3. Loading training and validation data: The training and validation data are then loaded and prepared for use in the model.
4. Training model: The CNN model is trained by extracted features based on the training data.
5. Graphing model: The model is then graphed to visualize its performance.
   a. Training and validation accuracies: The respective accuracies of the model are plotted on the graph.
   b. Training and validation losses: The respective losses of the model are also plotted on the graph.
6. Model evaluation on testing data: The model is then evaluated on the test data to determine its overall performance.
7. Testing images on model: The model is evaluated on a set of images for its performance on real-world data.

The dataflow for this project will involve the following steps (Fig. 3):

- A user will take a photo of a fruit or vegetable using a smartphone or other device.
- The user will submit the photo to the system, using the user interface of the system, such as a smartphone app or web application.

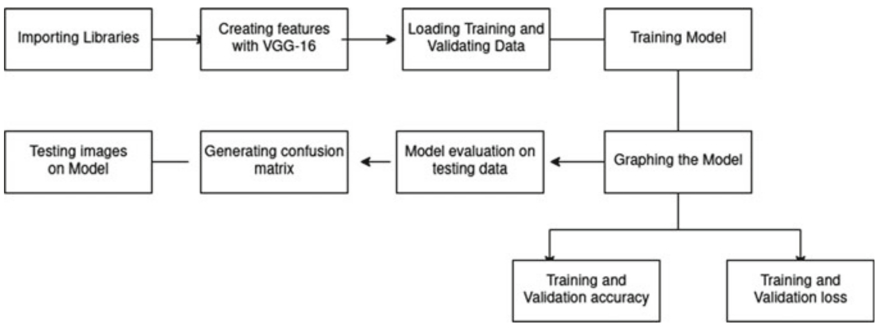


Fig. 2 System flowchart

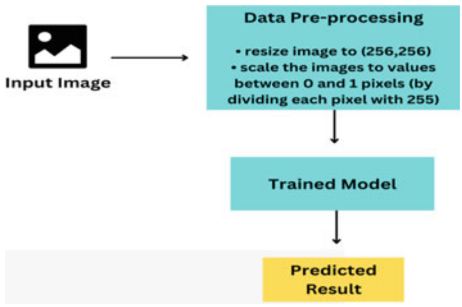


Fig. 3 Process flowchart

- The system will receive the photo and process it using the machine learning algorithms and models, to make a prediction about the quality and shelf life.
- The system will access the database or other data storage solution to retrieve any necessary information.
- The system will make the prediction about the quality and shelf life of the produce, and provide this prediction to the user via the user interface.

The proposed system consists of three main steps:

3.1 *Creating a Dataset*

We captured images of the tomato and potato using the camera and the internet. We have collected around 1500 images of both tomatoes and potatoes. Then, we classified them into different categories based on their quality and shelf life. The dataset is split into two sub-sets, namely training set and testing dataset. The training dataset will be used to train the machine learning models and the testing dataset will be used for validation.

Our dataset includes images of tomatoes and potatoes at various stages of freshness, from freshly harvested to overripe (Tables 1 and 2).

Table 1 Dataset—quality detection

Sr. no.	Image class	No. of images
01	Tomato good	228
02	Tomato	197
03	Tomato wilt	145
04	Tomato fruit cracking	165
05	Tomato anthracnose	136
06	Potato good	201
07	Potato bad	197

Table 2 Dataset—shelf-life detection

Sr. no.	Image class	No. of images
01	Tomato—0 days	149
02	Tomato—1–5 days	357
03	Tomato—5–10 days	800
04	Tomato—10–15 days	215
05	Potato—store it in dry place	76
06	Potato—throw it away	153

3.2 Creating a Machine Learning Model

The steps to train the model are as follows:

1. Collect all the images in the dataset.
2. Resize images to make each one of the same sizes.
3. Use data augmentation to scale images.
4. Create a CNN model (using the TensorFlow library in Python).
5. Train the model.

We now evaluate the model by finding out the accuracy of the model and then predicting the results on a set of test images.

We took multiple approaches to decide the best possible model. We tried changing the parameters to get a handful of models and track their results to decide the one with the most accurate prediction. We then compared the accuracy/effectiveness of each model and chose the one with the largest accuracy—73.33%.

The specifications were—20 Epochs, 80/20 split ratio, six convolutional layers, and optimizer—Adam. We then used VGG-16. Further, we fine-tuned VGG-16 using our created dataset to get the desired results. We were able to achieve 80% accuracy by using the same.

3.3 Prediction of Quality and Shelf Life

We captured new images to test the model. The model will now predict the quality and shelf life of the provided images with ~80% accuracy. These images are then fed into the CNN model, which uses its trained weights and biases to classify the freshness of the produce.

4 Result Analysis

The first thing we need to consider is which model will be used for the prediction. We took the one with the highest accuracy.

Based on Table 3, the model with specifications—CNN (VGG-16), 20 epochs, 16 layers, 80/20 split ratio, and optimizer—Adam—is the preferred model as it has the highest accuracy, i.e., 80%. Now we plot the confusion matrix and ROC curve for the model (Figs. 4 and 5).

The results would include GOOD or BAD for the quality while the range of days for their shelf life. The categories include "0 days", "1–5 days", "5–10 days", and "10–15 days" (Fig. 6 and Table 4).

Table 3 Evaluation of implemented models

Model used	No. of epochs	No. of layers	Split ratio	Optimizer	Accuracy
CNN	20	6	80/20	Adam	73.33
CNN	30	3	80/20	Adam	66.74
CNN	23	2	80/20	Adam	60.63
CNN	25	6	80/20	Adam	59.87
CNN	27	6	70/30	Adam	60.67
CNN (VGG-16)	27	6	60/40	SGD	45.24
CNN (VGG-16)	20	16	80/20	Adam	80

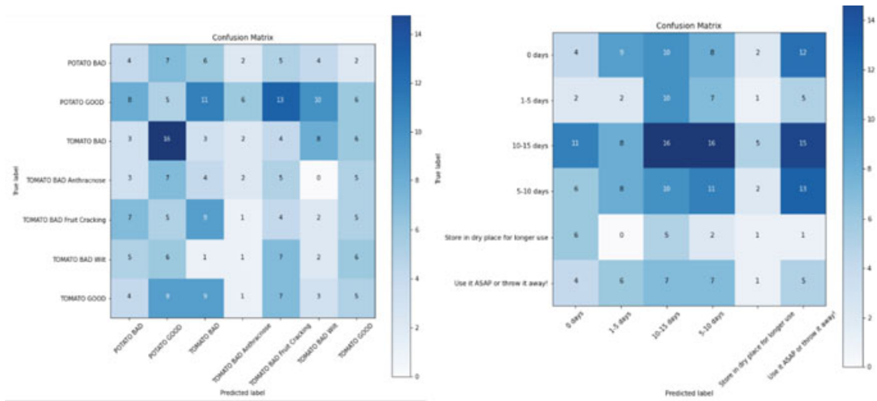


Fig. 4 Confusion matrix—**a** quality detection and **b** shelf-life detection

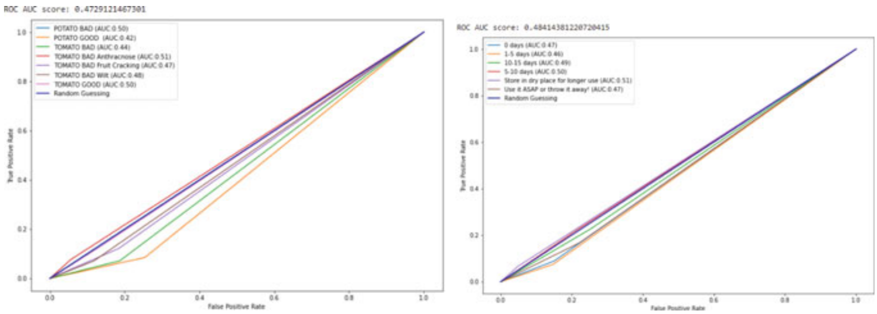


Fig. 5 ROC curve—**a** quality detection and **b** shelf-life detection

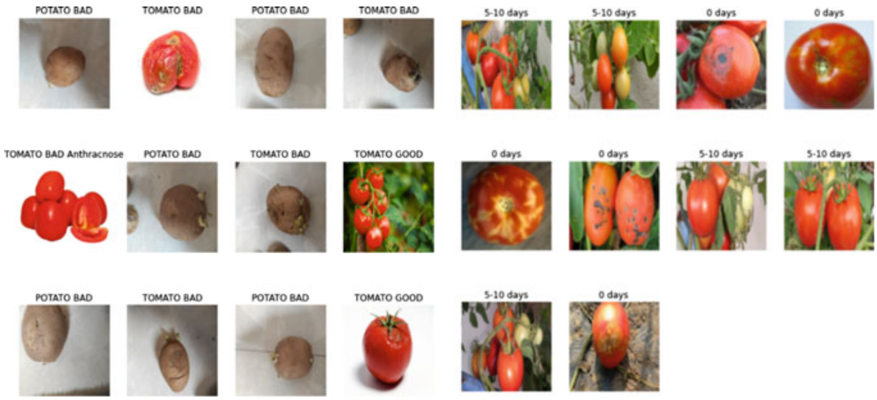


Fig. 6 Results—**a** quality detection and **b** shelf-life detection

Table 4 Comparison with existing solutions

Features	Zuo [13]	Sonwani et al. [14]	Kaur and Gupta [15]	Proposed solution
Concept used	SVM/Decision tree	CNN an actuators	Random Forests	CNN(VGG-16)
Preprocessing	Required	Required	Not required	Not required
Accuracy (%)	99	~95	~92	~80
Drawbacks	No shelf-life detection	Extensive setup required	No quality detection	None (accuracy can be increased)

5 Conclusion

In conclusion, our research suggests that the use of ML for the detection of the quality and shelf life of our daily foods is a promising approach. Future work in this area could focus on improving the performance of the proposed approach and extending it to other types of produce. Our results denote that the proposed system can accurately predict the freshness of tomatoes and potatoes and estimate their shelf life, providing valuable information for the consumer and the food industry.

References

1. Antony M, Kumar R (2021) A comparative study on predicting food quality using ML techniques. In: 2021 7th international conference on advanced computing & communication systems (ICACCS)
2. Bhargava A, Bansal A (2021) Fruits and vegetables quality evaluation using computer vision: a review. J King Saud Univ-Comput Inf Sci 33(3):243–257
3. Baygın M (2022) Image classification with SqueezeNet based deep feature generator. Turkish J Sci Tech 17(1):121–134
4. Varghese RR, Jacob PM, Ranjan SSDM, Varughese JC, Raju H (2021) Detection and grading using machine vision, ICSCC
5. Ren A et al (2020) ML approach towards the quality assessment of fruits using non-invasive sensing. IEEE Sens J 20(4):2075–2083
6. Sagar K, Gauri J, Mayuri P, Aditya K, Himali P (2018) Tomato disease detection using img processing. IJAREEIE 7(12)
7. Supriya P, Vaishnavi J, Komal D, Kulkarni BP (2022) Fruit quality detection using Python. IRJET 7(05)
8. Marcos JV, Botello-Álvarez E, Pérez-Pinal J, Lara MC, Fabiola L, Sánchez MGB, Alejandro B (2019) Fuzzy classification of the maturity. Hindawi
9. Ceh-Varela E, Hernandez-Chan G (2015) Tomatoes classifier using histograms. In: International conference on advances in engineering science and management, Agra, 8 Nov 2015
10. Swapnil W, Deshmukh SM (2018) Quality detection and grading of tomatoes using raspberry pi. IJRITCC 6(7)
11. Albert-Weiss D, Osman A (2022) Deep learning for shelf-life based on an active learning approach. MDPI
12. Femling F, Olsson A, Fernandez F (2018) Food identification for retail apps. SITIS
13. Zuo H (2022) Analysis and detection of tomatoes quality using machine learning algorithm and image processing. https://doi.org/10.21203/rs.3.rs-2016895/v1
14. Sonwani E, Bansal U, Alroobaea R, Baqasah AM, Hedabou M (2022) An artificial intelligence approach toward food spoilage detection and analysis. Front Public Health, 9
15. Kaur K, Gupta OP (2017) A ML approach to determine maturity stages of tomatoes. Orient J Comp Sci Technol 10(3):683–690

Liver Diseases' Stage Detection Using Hybrid Technique of Supervised Machine Learning Classifiers

Sagar Patel, Chintan Shah, and Premal Patel

Abstract The danger of liver disease, one of the deadliest illnesses in the world, has been rapidly increasing among humans over the past several decades. For researchers, one of the most difficult problems is predicting the disease using massive medical information. To solve this issue, machine learning techniques like classification and clustering have been created. The primary aim of this study is to use classification algorithms to figure out a patient's risk of getting liver disease. The stage of the liver disease is also indicated, including cirrhosis liver, liver fibrosis, fatty liver, and healthy liver. As a result, the classification accuracy and processing speed of the proposed Hybrid Classifier (RF, SVC, XGBoost) and the methods NB, SVM, LOR, RF, DT, KNN, and RBTC are assessed. Using After considering these performance factors, the Hybrid Classifier is determined to be the best classifier with 99% accuracy.

Keywords Logistic Regression · Neural network · Dataset · Accuracy · SVM · HYBRID model

S. Patel (✉) · C. Shah · P. Patel
Department of Computer Engineering, College of Technology, Silver Oak University, Ahmedabad, Gujarat, India
e-mail: Sagarpatel.alpha@gmail.com

P. Patel
e-mail: premalpatel.ce@socet.edu.in

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_54

1 Introduction

This research offers software that makes it easier to upload information and learn the likelihood of developing liver disease. Machine learning techniques are used in this study to categorize if the liver state is normal. For the prediction, we employ NB, SVM, LOR, RF, DT, KNN, RBTC [1–5], and hybrid models. The health sector will benefit from using this approach to forecast diseases. Using the patient's blood results, the model will be useful in determining if the patient's liver state is normal or abnormal. The medical businesses will find this patient information useful in the process. The current models use a variety of machine learning algorithms that produce output with lower accuracy and are unable to handle big data bundles. The performance of the liver datasets throughout training and testing is seen [6]. These previously created methods have been adequate, but more needs to be done to increase the accuracy of their prediction rate for the detection of liver disease. The solution that is being suggested here makes use of the machine learning idea, and the models are trained before being tested. The ultimate outcome will be predicted by the model with the highest accuracy. You must first enter your personal information into the system, which includes your age, gender, total and direct bilirubin levels, total proteins, albumin, A/G ratio, SGPT, SGOT, and Alkphos [7]. The results of the user's blood test can reveal these numbers. When the user provides these inputs, the system compares the supplied data to the training dataset of the most accurate model and then forecasts the outcome as risky or not. There are several algorithms that are employed, including Logistic Regression, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB) [1, 2, 4, 5, 8], and Hybrid Classifier (RF, SVC, XGBoost), among others. The Indian Liver Patient Dataset (ILPD) [6], which was chosen from the UCI Machine Learning repository, is the dataset that was employed. It consists of 30,691 patients. The system is relatively easy to implement and design. The system works in practically all settings and calls for very little system resources.

2 Methodology

The various phases are as follows.

Dataset (Figs. 1 and 2):

Age of the	Gender of	Total Bilir	Direct Bili	Alkphos A	Sgpt Alan	Sgot Aspa	Total Prot	ALB Albu	A/G Ratio	Result
65	Female	0.7	0.1	187	16	18	6.8	3.3	0.9	0
62	Male	10.9	5.5	699	64	100	7.5	3.2	0.74	0
62	Male	7.3	4.1	490	60	68	7	3.3	0.89	0
58	Male	1	0.4	182	14	20	6.8	3.4	1	0
72	Male	3.9	2	195	27	59	7.3	2.4	0.4	0
46	Male	1.8	0.7	208	19	14	7.6	4.4	1.3	0
26	Female	0.9	0.2	154		12	7	3.5	1	0
29	Female	0.9	0.3	202	14	11	6.7	3.6	1.1	0
17	Male	0.9	0.3	202	22	19	7.4	4.1	1.2	1
55	Male	0.7	0.2	290	53	58	6.8	3.4	1	0
57	Male	0.6	0.1	210	51	59	5.9	2.7	0.8	0
72	Male	2.7	1.3	260	31	56	7.4	3	0.6	0
64	Male	0.9	0.3	310	61	58	7	3.4	0.9	1
74	Female	1.1	0.4	214	22	30	8.1	4.1	1	0
61	Male	0.7	0.2	145	53	41	5.8	2.7	0.87	0
25	Male	0.6	0.1	183	91	53	5.5	2.3	0.7	1
38	Male	1.8	0.8	342	168	441	7.6	4.4	1.3	0
33	Male	1.6	0.5	165	15	23	7.3	3.5	0.92	1
40	Female	0.9	0.3	293	232	245	6.8	3.1	0.8	0
40	Female	0.9	0.3	293	232	245	6.8	3.1	0.8	0

Fig. 1 Liver diseases dataset [7, 9]

ID	N_Days	Status	Drug	Age	Sex	Ascites	Hepatome	Spiders	Edema	Bilirubin	Cholester	Albumin	Copper	Alk_Phos	SGOT	Tryglicerid	Platelets	Prothromi	Stage
1	400	D	D-penicill	21464	F	Y	Y	Y	Y	14.5	261	2.6	156	1718	137.95	172	190	12.2	4
2	4500	C	D-penicill	20617	F	N	Y	Y	N	1.1	302	4.14	54	7394.8	113.52	88	221	10.6	3
3	1012	D	D-penicill	25594	M	N	N	N	S	1.4	176	3.48	210	516	96.1	55	151	12	4
4	1925	D	D-penicill	19994	F	N	Y	Y	S	1.8	244	2.54	64	6121.8	60.63	92	183	10.3	4
5	1504	CL	Placebo	13918	F	N	Y	Y	N	3.4	279	3.53	143	671	113.15	72	136	10.9	3
6	2503	D	Placebo	24201	F	N	Y	N	N	0.8	248	3.98	50	944	93	63	NA	11	3
7	1832	C	Placebo	20284	F	N	Y	N	N	1	322	4.09	52	824	60.45	213	204	9.7	3
8	2466	D	Placebo	19379	F	N	N	N	N	0.3	280	4	52	4651.2	28.38	189	373	11	3
9	2400	D	D-penicill	15526	F	N	N	Y	N	3.2	562	3.08	79	2276	144.15	88	251	11	2
10	51	D	Placebo	25772	F	Y	N	Y	Y	12.6	200	2.74	140	918	147.25	143	302	11.5	4
11	3762	D	Placebo	19619	F	N	Y	Y	N	1.4	259	4.16	46	1104	79.05	79	258	12	4
12	304	D	Placebo	21600	F	N	N	Y	N	3.6	236	3.52	94	591	82.15	95	71	13.6	4
13	3577	C	Placebo	16688	F	N	N	N	N	0.7	281	3.85	40	1181	88.35	130	244	10.6	3
14	1217	D	Placebo	20535	M	Y	Y	N	Y	0.8	NA	2.27	43	728	71	NA	156	11	4
15	3584	D	D-penicill	23612	F	N	N	N	N	0.8	231	3.87	173	9009.8	127.71	96	295	11	3

Fig. 2 Liver diseases stage dataset (Kaggle.com)

2.1 Exploratory Data Analysis

Data visualization: Data visualization allows us to see how the data are organized and what sort of relationships the qualities of the data hold. It is the quickest method for determining whether the input features match the output features [1, 7, 9].

Correlation Analysis: Three crucial aspects of correlations. They can provide information regarding the nature (shape), direction, and intensity (degree) of the link between two variables (Fig. 3).

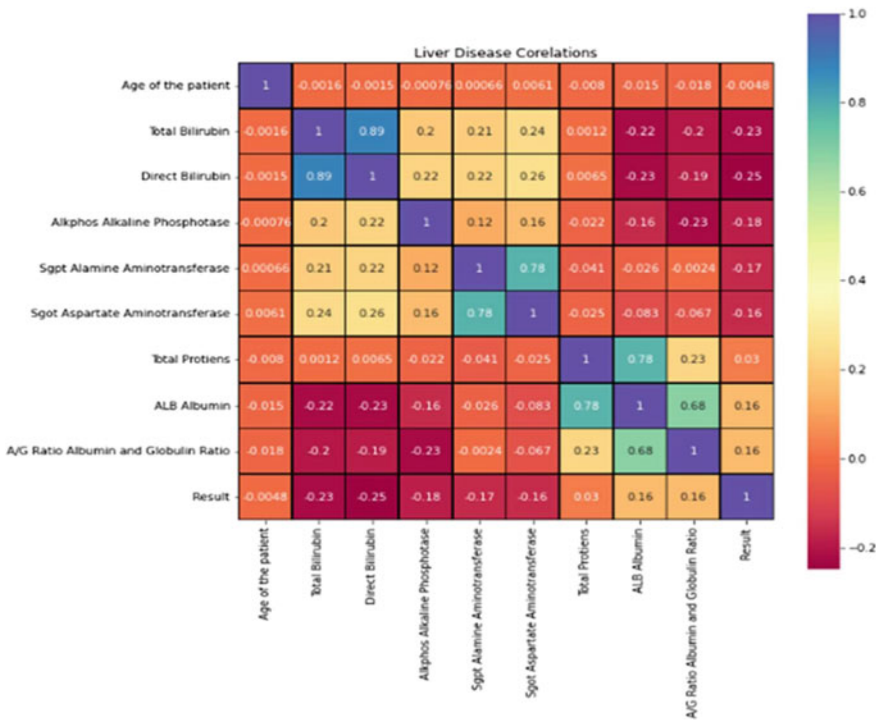


Fig. 3 Correlation matrix of the model

2.2 *Data Preprocessing*

This entails removing the text's null and most prevalent terms. The words in the dataset include both very long and small words, links, and numerous full stops. Before giving it to the algorithm, all of these need to be removed. Data Cleaning, Data Integration, Data Reduction, and Data Transformation are important phases in data preprocessing. Accuracy, completeness, consistency, timeliness, believes ability, and interpretability are all requirements that are met [1].

2.3 *Training Classification Model*

We split the data set across the testing and training phases in different ratios to get the best results. Now, in order to train the model and forecast the precise outcome, we use machine learning techniques such as NB [4, 8], Logistic Regression, RF, DT, KNN, SVM [1–3], and Hybrid Ensemble Classifier.

3 Modeling and Analysis

3.1 *SMOTE: Synthetic Minority Oversampling Technique*

Oversampling

Synthetic minority oversampling technique (SMOTE) is one of the most utilized oversampling techniques for addressing the issue of imbalance (Fig. 4).

It seeks to balance the distribution of classes by deliberately expanding minority class examples and then duplicating them.

New minority instances arise when minority instances that currently exist are combined. For the minority class, it creates virtual training records using linear interpolation. For each example in the minority class, one or more of the K-Nearest Neighbors are randomly chosen to serve as these synthetic training records. Following the oversampling procedure, the data are rebuilt and can be subjected to several categorization models.

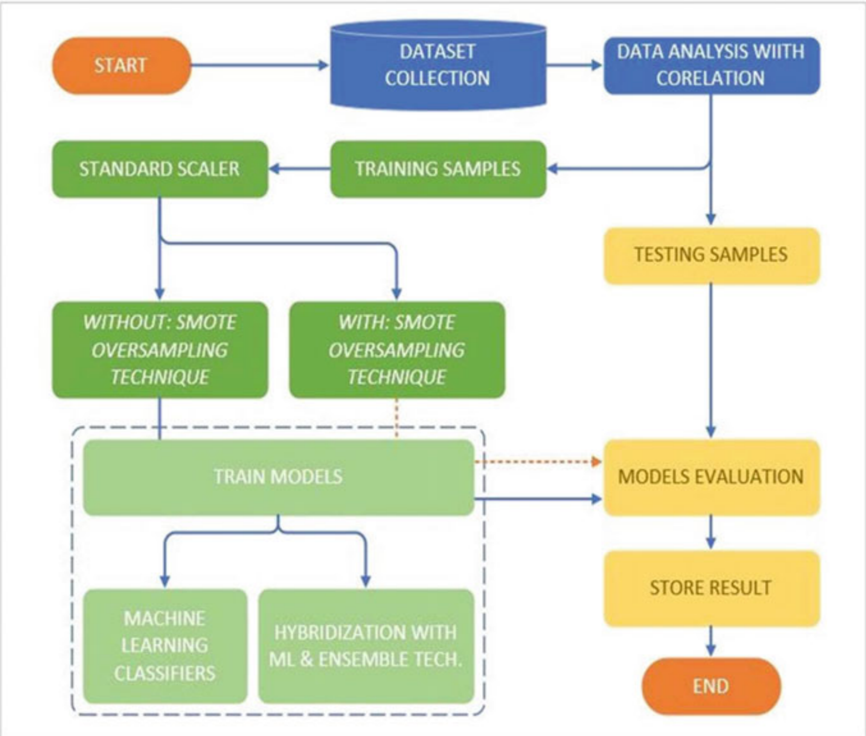


Fig. 4 Block diagram of the model [10]

Random oversampling involves choosing random examples from the minority class with replacement and adding multiple copies of this instance to the training data; therefore, it is likely that a single instance will be chosen more than once [2].

Undersampling

The reverse of random oversampling is random under sampling. In order to decrease the number of examples in the majority class in the processed data, this method aims to randomly pick and eliminate samples from the majority class [2].

The various Machine Learning Models used are:

Logistic Regression:

One of the more straightforward classification models is Logistic Regression. It is beneficial when experimenters wish to look at interactions between variables because of its parametric nature, which can be partially comprehended by looking at the parameters. Regrettably, the term "logistic regression" refers to the use of regression models to identify continuous response variables, as opposed to categorization, where the variable that responds is discrete. The conclusion of the Logistic Regression that calculated the chance that the variable that responded belonged to a specific class may have served as the phrase's inspiration. The beta variable, or factor, in this model is typically estimated via most likely hood estimation (MLE). Once the optimum coefficients (or coefficients, when there are numerous variables that are independent) have been determined, the conditional likelihoods for every observation can be computed, logged, and summed to give a forecast probability. A chance of less than.5 suggests 0 and a likelihood of above 0 predicts 1 if the classification is binary. Once the model has been calculated, it is advised to evaluate the model's quality of fit, or how well it forecasts the dependent variable [1–3].

K-Nearest Neighbor:

The specifics of the KNN technique's implementation are covered in this section. The algorithm for K-Nearest Neighbor is the entire training dataset. When an estimate is required for an instance of unseen data, the KNN method will search in the training database for the k-most similar situations. A class label is chosen for classification problems based on a majority vote, meaning that the label that is most commonly expressed around a particular data point is adopted. Although technically "plurality voting" applies, the term "majority vote" is more frequently employed in literature. The difference between both terms is that "majority voting" informally calls for a majority of more than 50%, which typically only applies when there are only two options. When there are many classes, such as four, you do not necessarily require 50% of the vote to come to a conclusion regarding a class; instead, you can assign a class label with a vote of more than 25% [1–3].

Support Vector Machine:

SVM seeks to identify the best possible hyperplane for classifying the data. SVM is implemented in Python using the scikit-learn library. The preprocessed data are divided into a training set and test set, each comprising 25% and 75% of the whole dataset. In a high- or infinite-dimensional space, a Support Vector Machine creates one hyperplane or a collection of hyperplanes. The hyperplane with the greatest distance from the closest training data point for any class (referred to as the functional margin) achieves a decent separation since, generally speaking, the margin has a positive correlation with the classifier's generalization error. Decision boundaries known as hyperplanes assist in categorizing the data points. Different classes can be given to the data points that fall on each side of the hyperplane. Additionally, the amount of features affects how big the hyperplane is. The hyperplane is essentially a line if there are just two input features. The hyperplane turns into a two-dimensional plane if there are three input features. When the quantity of features exceeds a certain point, it becomes impossible to imagine [1–3].

Hybridization:

In the hybridization process, either regression or classification models are put together using two-layer estimators. The first layer is composed of the baseline models, which were all used to predict the outcomes on the test datasets. A Meta-Classifier or Regressor, which makes up the second layer, uses all the previous forecasts from the baseline models as input to generate new predictions. Here, I've created a hybrid model for predicting liver disease and liver stage using three machine learning classifiers: RF, SVC, and XGBOOST [1–3].

Specifically, we will evaluate the following three algorithms:

- Random Forest.
- Support Vector Classifier.
- eXtreme Gradient Boosting Classifier.

Mlxtend:

A Python module called machine learning extensions (Mlxtend) contains helpful tools for routine data science activities. It includes many technologies that are beneficial for jobs involving data science and machine learning, such as:

1. Feature Selection.
2. Feature Extraction.
3. Visualization.
4. Ensembling and many more.

How to use the Stacking Classifier on the classification dataset is covered in this article [11].

```
HYBRID CLASSIFIER Accuracy is :99.96%
```

```
from sklearn.metrics import classification_report
STK_Pred=STK.predict(X_test)
STKreport = classification_report(Y_test, STK_Pred)
print(STKreport)
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      3561
           1       1.00      1.00      1.00      3489

    accuracy                           1.00      7050
   macro avg       1.00      1.00      1.00      7050
weighted avg       1.00      1.00      1.00      7050
```

Fig. 5 Classification report of liver disease prediction

Why Hybridization?

The bulk of data science and machine learning competitions are won using stacked models. They can increase the amount of accuracy that some models are now showing. We can get a great deal of stacked models by using an array of techniques for the first layer of the design because different algorithms capture different trends in training data. We can get better and more precise results by integrating the two models [1–3, 9].

4 Results and Discussion

This study's main goal was to forecast liver disease using a variety of machine learning techniques. We made predictions using a Hybrid Ensemble Classifier, which has a 99.96% accuracy rate and produces superior outcomes. My proposed Hybrid Classifier is compared against the NB [4, 8], SVM, LOR, RF, DT, KNN [1–3], and RBTC algorithms. We have noted the following Accuracy, Precision, Sensitivity, and Specificity for each algorithm (Figs. 5, 6, 7, and 8):

5 Comparison Chart

See Figs. 9 and 10.

```
HYBRID CLASSIFIER Accuracy is :98.39%
```

```
from sklearn.metrics import classification_report
STK_Pred=STK.predict(X_test)
STKreport = classification_report(y_test, STK_Pred)
print(STKreport)
```

```
              precision    recall  f1-score   support

         1.0       1.00      1.00      1.00         8
         2.0       0.97      0.97      0.97        62
         3.0       0.98      0.98      0.98       101
         4.0       1.00      1.00      1.00        77

    accuracy                           0.98       248
   macro avg       0.99      0.99      0.99       248
weighted avg       0.98      0.98      0.98       248
```

Fig. 6 Classification report of liver disease stages prediction

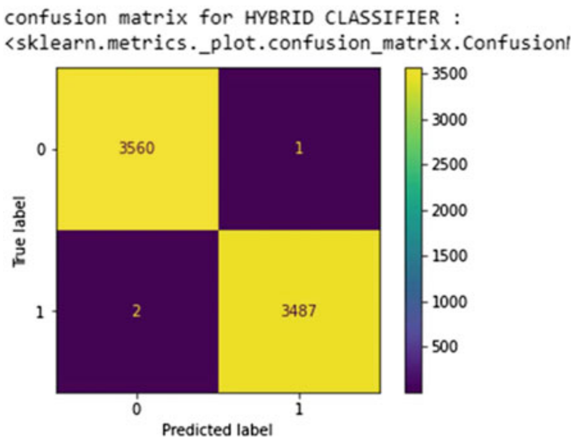


Fig. 7 Confusion matrix of liver disease prediction

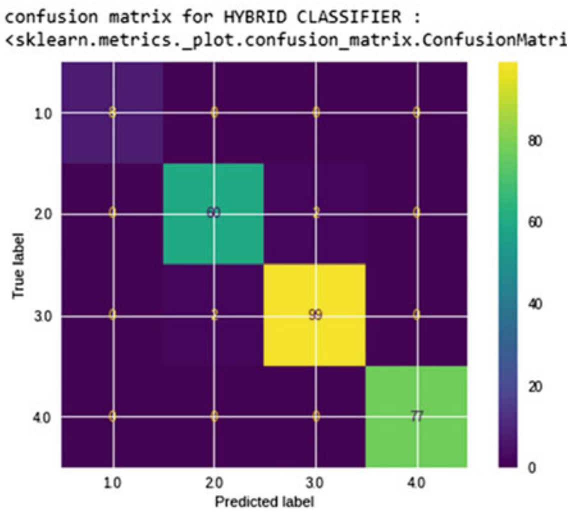


Fig. 8 Confusion matrix of liver disease stages prediction

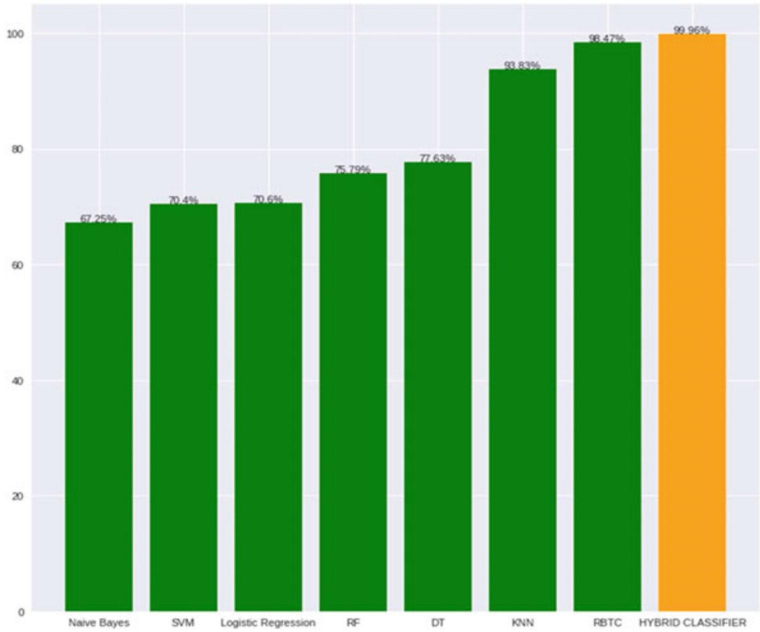


Fig. 9 Comparison chart of liver disease prediction using Hybrid Classifier

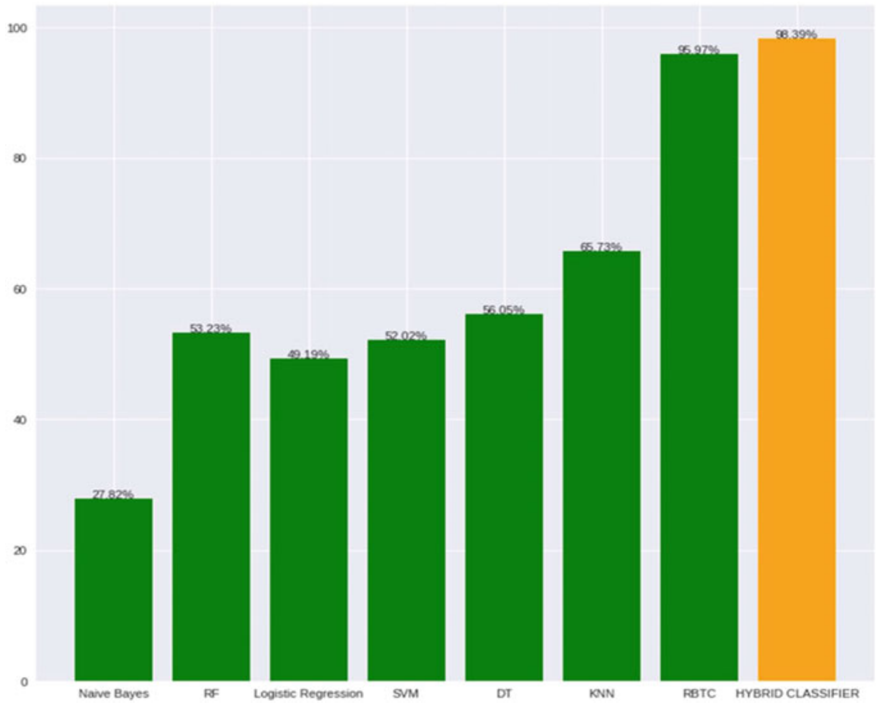


Fig. 10 Comparison chart of liver disease stage prediction using Hybrid Classifier

6 Conclusion

In this study, we have proposed methods for using machine learning to diagnose liver disease and predict the phase of liver disease in patients. SVM, RF, DT, NB, Logistic Regression, KNN, RFBTC [1–4, 8], and Mixed Classifier were a few of the many machine learning techniques that were examined. All the models were used in the system's implementation, and their performance was assessed. Based on certain performance metrics, the performance review was conducted. Our suggested model, which combines RF, SVC, and XGBOOST, has the highest accuracy, predicting liver disease stages with an accuracy of 99% and identifying them with an accuracy of 98%.

References

1. Kuzhippallil MA, Joseph C, Kannan A (2020) Comparative analysis of machine learning techniques for Indian liver disease patients, © IEEE
2. Kumar P, Thakur RS (2021) Liver disorder detection using variable—Neighbor weighted fuzzy K nearest neighbor approach. In: 18th international conference, Springer, Heidelberg

3. Harper PR (2005) Are view and comparison of classification algorithms for medical decision making. Health Policy 71(3):315–331
4. Christopher N, New automatic diagnose is of liver status using Bayesian classification
5. Ramana ER, Sorrell M, Maddrey WC (2007) Schiff's diseases of the liver, 10th edn Copyright©2007
6. Singha J, Baggab S, Kaurc R (2020) Software based prediction of liver disease with feature selection and classification techniques, © Elsevier
7. Gupta S, Karanth G, Pentapati N, Badri Prasad VR (2020) A web based framework for liver disease diagnosis using combined machine learning models, © IEEE
8. Sug P (1997) On the optimality of the simple Bayesian classifier under zero-one loss. Mach Learn 29(2–3):103–130
9. BUPA liver disorder dataset. UCI repository machine learning databases
10. Sorich MJ (2009) An intelligent model for liver disease diagnosis. Artif Intell Med 47:53–62
11. 16th edition Harrison's Principles of internal medicine

Revolutionizing Transportation: Advancements in Robot-Assisted Mobility Systems

Usman Ahmad Usmani, Ari Happonen, and Junzo Watada

Abstract We explored advancements in robot-assisted mobility systems and the impact of transportation revolutions. Current developments highlight technological breakthroughs in robotics and related transportation domains, including autonomous vehicles, last-mile delivery, and public transportation. We addressed ethical considerations, regulatory frameworks, and societal acceptance challenges while examining future developments like human–robot collaboration and the role of artificial intelligence in traffic flow optimization. By providing valuable insights into the transformative potential of robot-assisted mobility systems, the study serves as a resource for researchers, policymakers, and industry professionals involved in shaping the future of transportation.

Keywords Robot-assisted mobility · Digitalization · Artificial intelligence · Society 2.0 · Machine learning · Sensor fusion · Last-mile delivery · Robotics · Autonomous vehicles

1 Introduction

Transportation connects people, goods, and services across multiple destinations. As societies strive for improved efficiency, safety, and sustainability in transportation, emerging technologies open new possibilities for novel mobility innovations, including immense promises in robot-assisted mobility systems. Integrating robots into transportation infrastructure, next-generation communication networks, and

U. A. Usmani (✉)
Universiti Teknologi Petronas, Seri Iskandar, Perak, Malaysia
e-mail: usman_19001067@utp.edu.my

A. Happonen
LUT University, Yliopistonkatu 34, 53850 Lappeenranta, Finland
e-mail: ari.happonen@lut.fi

J. Watada
Waseda University, Tokyo 169-8050, Japan

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8_55

other vehicles can offer solutions to address both the sustainability challenges and enhance the overall transportation experience [1, 2]. Advancements in robotics are allowing the integration of intelligent machines into transportation systems, enabling them to perform traditional human tasks. This change could transform domains like autonomous vehicles, last-mile delivery, and public transportation by leveraging technologies like artificial intelligence, metaverse [3], augmented reality [4, 5], machine learning, sensor fusion, robot-assisted mobility systems which can enhance efficiency, sustainability, safety, and accessibility in ways previously unimaginable. In public transportation, robotic assist buses, trains, and stations could improve overall passenger experience, provide more real-time information, enhance accessibility for individuals with disabilities or ongoing rehabilitation periods [6, 7], and contribute to passenger safety, convenience, and operational efficiency.

Our study explored recent advancements in robot-assisted mobility systems through a review of relevant literature and case studies involving key technological breakthroughs and their application in transportation domains. By analyzing the current state of research, trends, and challenges, we shed light on the transformative potential of robot-assisted mobility systems. We cover areas like robots in autonomous vehicles [8] and self-driving cars, with robots assuming the responsibilities of driving, navigation, and decision-making. We will also touch on the technological advancements enabling autonomous vehicles to navigate complex environments, detect obstacles, and make real-time decisions for economical, safe, and efficient driving on public and other roads.

Furthermore, the utilization of robots in last-mile delivery was explorer as the most challenging area in logistics and supply chains for good (green) performance [9, 10] efficiency [11], and sustainability [12], although decarbonization of warehousing is not that easy too [13]. By employing robots for package delivery, logistics companies can optimize routes, reduce delivery times, and minimize environmental impact, e.g., on package handling, customers vs. process interaction, and have tools for special regulatory-related situations, for example, pandemics and social distancing [14] times.

With the technological advancements comes robot-assisted life style related integration challenges and society related ethical considerations, needs for specific and related regulatory frameworks, and potential societal acceptance/cultural-based resistance issues. All these include the show stoppers from the existing regulatory landscape and needs for future policy development. In addition to fully autonomous vehicles, there is human–robot collaboration in transportation, including cobots working side by side with humans in a synergistic setting, representing an exciting frontier of future lifestyle. On the algorithms side, robotization and automatization offer potential to use, e.g., artificial intelligence in optimizing traffic flows, reducing congestion, and rerouting traffic to safer and more economically sustainable routes.

By providing a comprehensive overview of the advancements in robot-assisted mobility systems, applications, challenges, and future prospects for emission reduction [15] and sustainability matters, our study makes three main contributions to the field of robot-assisted mobility systems in transportation: **(1) Comprehensive**

analysis of technological advancements: Studies on recent technological advancements tend to explore the integration of robots in autonomous vehicles, last-mile delivery activities, and public transportation. These studies highlight key breakthroughs in robotics, artificial intelligence, and sensor technologies. Through the synthesis of the current state of research and trends, our work offers valuable insights into the technological landscape and its potential impact on revolutionizing transportation. **(2) Examination of challenges and opportunities**: We address the challenges and opportunities associated with the integration of robot-assisted mobility systems, considering ethical considerations, regulatory frameworks, and societal acceptance. All of these aspects emphasize the importance of responsible deployment of robot-assisted solutions and are connected to ensuring long-lasting public trust. The aim is to provide a holistic understanding of the barriers and enablers in implementing robot-assisted mobility systems in transportation, thus guiding future development in the associated challenges. **(3) Exploration of future developments and impact**: The concept of human–robot collaboration in transportation, envisioning a future where humans and robots work together synergistically, is addressed for enhance mobility. Also, the role of artificial intelligence in optimizing traffic flow and reducing congestion is touched, related to efficient, less polluting, and sustainable transportation.

2 Advancements in Robot-Assisted Mobility

The integration of robots into transportation systems has ushered in a new era of innovation, transforming the way we perceive and experience mobility. This section of the paper focuses on the technological advancements in robot-assisted mobility systems across various domains of transportation. From autonomous vehicles to last-mile delivery and public transportation, significant breakthroughs in robotics, artificial intelligence, management systems, and sensor technologies have paved the way for better healthcare solutions [16–22], changes in education [23], industrial applications [24], improved constructions [25], less wasteful industrial operations [26], asset and material manipulations [27, 28], enhanced efficiency, safety, and accessibility [29]. By exploring the advancements, the following section especially focuses on providing a comprehensive overview of the cutting-edge technologies driving the revolution in transportation and their impact on reshaping the way we move. Through an in-depth analysis of the integration of robots in autonomous vehicles, last-mile delivery operations, and public transportation systems, this section sheds light on the transformative potential of robot-assisted mobility systems which is revolutionizing transportation.

2.1 Integration of Robots in Autonomous Vehicles

Robots in autonomous vehicles represent a major technological advancement in transportation, with transformative potential to reshape the landscape of autonomous transportation. Autonomous vehicles/self-driving cars rely on a combination of robotics, artificial intelligence, and sensor technology advancements to navigate and operate without human intervention. The integration of robots within these vehicles enables them to perceive their surroundings, make informed decisions, and execute actions in real-time. Where one crucial success factors is the development of advanced perception systems, including LiDAR, radar and camera systems, which enable vehicle to gather real-time data about its environment, including the detection and tracking of other vehicles, pedestrians, and potential obstacles. The fusion of sensor data through sophisticated algorithms allows for an accurate understanding of the vehicle's surroundings, essential for safe navigation, and also the vehicle's current condition, improving possibilities for pre-emptive maintenance [30, 31], fleet-level [32, 33] vehicle and a group of vehicle decisions, and balancing of, e.g., EV chargers' resource utilization in the routes of vehicles currently actively moving between different points and in potential need to charge up their batteries.

Furthermore, advancements in decision-making algorithms and control systems enable autonomous vehicles to interpret sensor data, plan routes, and make appropriate driving decisions. Machine learning and artificial intelligence techniques are utilized to analyze complex data patterns, do segmentation [34, 35] and make the previously mentioned optimized decisions. The algorithms can consider factors such as traffic conditions, road regulations, location data reliability [36, 37], vehicle status like estimated distance to fuel/charge up and maintenance needs [38], predicted road utilization rates, and potential hazards to ensure safe, sustainable and efficient navigation, reflecting successful, robust and reliable control systems in use. These systems execute the driving actions based on the decisions made by the vehicle's algorithms, making precise control over steering, acceleration, and braking to maintain vehicle stability with a smooth and safe ride. At the same time, extended communication and connectivity vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication systems allow autonomous vehicles to enhance overall transportation network safety by exchanging information with other vehicles and the surrounding infrastructure. V2V and V2I allows cooperative behavior, such as coordinated merging, lane changing, and intersection navigation, enhancing overall smooth, reliable, and safe traffic flows.

Multiple times mentioned, additional safety comes from human errors and associated risks related to minimization efforts, which is related to wider range of information these driving algorithms have available for them, than what humans tend to use for the on-the-fly driving decisions. Moreover, autonomous vehicles have the potential to increase traffic efficiency, reduce congestion, and improve fuel efficiency through optimized vehicle driving behaviors and cooperative traffic activities. Despite the advancements, there are also challenges like ethical considerations, decision-making in unavoidable accident situations and liability issues, regulatory-related limitations, and public acceptance/resistance to change and so on.

2.2 Robot-Assisted Last-Mile Delivery

Last-mile delivery, the final leg of the supply chain, is often characterized to be the most challenging, least efficient miles, touted with congestion issues and high in delivery costs. Robot-assisted last-mile delivery has emerged as a promising solution to overcome these challenges. In robotized last-mile delivery operations, solutions include ground-based robots and drones [39], for a seamless, optimized, and streamlined last-mile delivery process, including efficient and intelligent package-handling mechanisms. The robots are equipped with advanced navigation systems, sensing capabilities, and efficient package-handling mechanisms to ensure reliable and efficient deliveries [40]. In packet handling, robots are equipped with robotized arms or grippers capable of safely picking up and manipulating packages of various sizes and shapes. Advancements in computer vision and machine learning algorithms enable robots to accurately perceive and manipulate packages, ensuring secure and damage-free delivery.

In the navigation side, ground-based robots tend to be equipped with sophisticated sensors, including LiDAR, cameras, and inertial measurement units, to navigate in dynamic and complex urban environments. The sensors provide real-time data for the navigation system, enabling it to detect and avoid obstacles, navigate through pedestrian-heavy areas, and adapt to changing traffic conditions. Drones emerged as a solution for last-mile delivery, particularly in areas with limited road infrastructure or when time-sensitive deliveries are required. Drone technology has witnessed significant advancements in terms of range, payload capacity, and autonomous flight capabilities. Drones equipped with computer vision systems and GPS technology can accurately locate delivery destinations and perform autonomous take-off, landing, and navigation. On the other hand, regulatory considerations, airspace management for drones, and public acceptance are crucial factors that need to be addressed for wide range use of drones in last-mile delivery.

Artificial intelligence and machine learning algorithms help optimize last-mile delivery operations, by allowing robot to learn from historical delivery data, optimize routes, and make real-time decisions to improve delivery efficiency, with capability to assist in predicting and adapting to changing customer demands, weather conditions, and delivery patterns. This sort of robot-assisted last-mile delivery offers eco-friendly, sustainable, cost-reducing advantages, including increased delivery speed, reduced emissions, and total costs [41]. Robots can also operate 24/7, providing faster and more efficient deliveries, using slow traffic, non-congested, and traffic-free times of the day, compared to human-operated vehicles, mostly focusing on “normal” weekly working hours. Moreover, robots can make complex optimization decisions for routes, reducing the distance, time, emissions, and even brake dust generated on the route, adjusting as easily to ICE or EV vehicles specialties what comes to CO_2 emissions. This leads to a more sustainable and environmentally friendly delivery process.

2.3 Robots Integrated into Public Transportation Systems

Robots are playing an increasingly significant role in enhancing public transportation systems, offering improved passenger experience, increased accessibility, operational efficiency, and the potential to optimize resource allocation and traffic management, resulting in reduced congestion and improved system performance. Robots in public transportation systems serve and support workflows from passenger assistance and information provision tasks, all the way to maintenance and security operations, which can create a seamless and enjoyable experience for commuters while also optimizing operational processes [42, 43]. In passenger assistance, robots enhance tasks such as ticketing, answering frequently asked questions, supporting way- and route-finding, helping locate amenities within transportation hubs and provide real-time travel information. Equipped with natural language processing capabilities and interactive interfaces, these robots can communicate with passengers in a user-friendly and intuitive manner. This enhances the passenger experience by providing personalized assistance, reducing waiting times, and delivering accurate and timely information. Moreover, robots can contribute to enhancing accessibility and assist individuals with disabilities or limited mobility by providing personalized support, such as helping with boarding and alighting, offering priority seating, or providing audio-visual announcements. By integrating robots as accessible companions, public transportation becomes more inclusive and societally sustainable and ensures that all individuals can travel with ease and independence. In addition to passenger assistance, robots can work in maintenance and monitoring of public transportation infrastructure. Inspection robots equipped with sensors and cameras can autonomously assess the condition of tracks, tunnels, roads, sidewalks, and other critical public transportation infrastructure/components. By identifying potential maintenance needs, robots support and enable timely repairs, which reduce downtime, leading to improved operational efficiency and enhancing safety.

Robotic systems can be employed for security [44] and safety enforcement of public transportation systems. Surveillance robots with cameras and video analytics software robots do transportation hubs monitoring, detecting suspicious activities or unauthorized access. These robots can also assist in emergency situations by providing real-time video feeds to operators, helping them make informed decisions and coordinate effective responses. Especially in monitoring and safety-related matters, robots contribute to improved operational efficiency, as they perform boring, tedious, and repetitive tasks consistently and without fatigue or loss of interest.

Concerns regarding privacy, data security, and public acceptance need to be addressed tough. Ensuring a seamless collaboration between humans and robots is crucial, as robots should complement and support human staff, improving their work quality output instead of replacing them. Additionally, as robots will be handling, storing, tagging, and generating new user-level data, data security, personal data privacy, and system intruder protection will be matters to be taken seriously in the future to come.

2.4 Challenges and Opportunities in Robot-Assisted Mobility Systems

The challenges and opportunities that arise in the context of robot-assisted mobility systems relate to the increasingly prevalent integration of robots in domains of mobility, including personal mobility devices like exoskeletons [45]. As robots and automation in movement become increasingly prevalent in various domains of mobility, it is crucial to address the ethical, regulatory, and societal implications associated with their deployment. The integration of robot-assisted mobility systems in transportation brings forth a host of ethical and regulatory challenges to be addressed. With more sophisticated and autonomous systems, it is essential to establish frameworks that ensure their ethical operation and regulatory compliance [46], including the ethical consideration of allocating the responsibility and liability properly. These ethical frameworks will need to help to establish the task of guiding the decision-making processes of autonomous systems. With autonomous robots and automated transportation, questions arise regarding who should be held accountable in the event of accidents or failures. Determining the allocation of responsibility between manufacturers, operators, and users of robot-assisted mobility systems becomes a pressing issue, with ethical dilemmas, such as choosing between protecting occupants or pedestrians during critical situations. In this context, related challenges are:

A. **Data Privacy and Security**: The collection and processing of vast amounts of data in robot-assisted mobility systems raise concerns about data privacy and security. Robust data protection measures, ensuring secure data handling practices, and obtaining informed consent are essential safeguards for individuals' personal information.
B. **Ethical Decision-Making**: As robots become increasingly autonomous, ethical considerations arise regarding decision-making in complex scenarios. Establishing guidelines and frameworks for ethical behavior and ensuring transparency in the decision-making process are crucial to address moral dilemmas and ensure responsible robot behavior.
C. **Liability and Accountability**: Determining liability and accountability in the event of accidents or system failures involving robot-assisted mobility systems is a complex challenge. Clear legal frameworks and regulations are needed to establish responsibility among manufacturers, operators, and users, ensuring that appropriate parties are held accountable.
D. **Human–Computer Interaction**: Maintaining a balance between human control and autonomous capabilities is needed to ensure safe and effective human–robot

interaction. Understanding the impact of automation on human behavior and addressing potential user trust issues with appropriate levels of user control are important ethical considerations.

E. **Fairness and Bias**: Robot-assisted mobility systems must be designed and implemented in a manner that ensures fairness and prevents biases. Avoiding discriminatory practices, addressing algorithmic biases, and promoting diversity and inclusivity in the development process are vital to avoid exacerbating existing societal inequalities.
F. **Social Acceptance and Trust**: Public acceptance and trust of robot-assisted mobility systems play a crucial role in their successful integration. Open and transparent communication, public engagement, and addressing concerns about job displacement and loss of human interaction are the factors to gain societal acceptance.
G. **International Standards and Regulations**: Establishing international standards and regulations for the development, deployment, and operation of robot-assisted mobility systems between stakeholders, including governments, industry, and academia, can ensure consistency, safety, and ethical practices across different regions and jurisdictions [47, 48].
H. **Ethical Considerations in Decision-Making Algorithms**: Ensuring that decision-making algorithms employed in robot-assisted mobility systems are designed to prioritize safety, fairness, and human well-being is essential. Transparent and accountable algorithmic decision-making processes can help mitigate potential ethical challenges and promote responsible technology use.
I. **Continuous Monitoring and Evaluation**: Ethical considerations in robot-assisted mobility systems should be an ongoing process, no matter which area it is applied to [38]. Continuous monitoring, evaluation, and adaptation of ethical guidelines and regulations are necessary to address emerging issues and ensure that these systems align with evolving societal values and expectations [49].

Considering these ethical and regulatory considerations is crucial for the responsible development and deployment of robot-assisted mobility systems. Striking a balance between technological advancements and societal values ensures that these systems contribute positively to mobility while upholding ethical principles and addressing potential risks.

2.5 Societal Impact, Sustainable and Robotization Acceptance

The integration of robot-assisted mobility systems in transportation not only poses technological and operational challenges but also has a profound impact on society, including the potential effect on employment and the workforce. As robot-assisted mobility systems become more prevalent, there is a concern about the displacement of human workers in the transportation sector. By anticipating these changes, we can develop strategies to mitigate negative consequences, involving opportunities for identification for reskilling and retraining affected workers to take on new roles that align with the evolving transportation landscape. Ensuring a just transition and supporting affected individuals will be essential in addressing the societal impact of automation in transportation.

A. **Job Displacement and Employment**: The integration of robot-assisted mobility systems may lead to concerns about job displacement in transportation and delivery services. Managing the potential impact on employment and implementing measures for reskilling and job transition mitigates societal disruption and helps with just transition for affected workers.
B. **Changing User Behavior and Adoption**: The acceptance and adoption of robot-assisted mobility systems depend on user behavior and preferences. Encouraging positive user experiences, addressing user concerns, and promoting user-centric design can foster acceptance and facilitate the integration of these systems into everyday life [50–52].
C. **Social and Cultural Norms**: Robot-assisted mobility systems can challenge existing social and cultural norms regarding transportation and human–robot interactions. Understanding and addressing cultural perceptions, social acceptance, and expectations are important for ensuring that these systems align with societal values and norms.
D. **Trust and Reliability**: Building trust in robot-assisted mobility systems is crucial for their widespread acceptance. Demonstrating reliability, safety, and transparency in system performance, data handling [53], and decision-making processes can enhance public trust and confidence in these technologies.
E. **Inclusivity and Accessibility**: Inclusivity considerations ensure that robot-assisted mobility systems are accessible to individuals from diverse backgrounds, people with disabilities, anxiety, etc., challenges [54], and differences in shapes and sizes. Designing systems that cater to different mobility needs, providing accessible interfaces, and addressing potential barriers can enhance societal acceptance and ensure equitable access.
F. **Impact on Social Interactions**: The integration of robot-assisted mobility systems can impact social interactions and human-to-human connections in transportation settings. Balancing the benefits of increased efficiency and convenience with the importance of human interaction and social bonds is necessary for maintaining healthy social dynamics and community engagement.
G. **Perceived Threats and Ethical Concerns**: Perceptions of technology as a threat to personal privacy, autonomy, or employment can affect societal acceptance. Addressing these concerns through transparent communication, ethical guidelines, and regulatory frameworks helps alleviate fears and promote positive societal attitudes toward robot-assisted mobility systems.
H. **Environmental Awareness and Sustainability**: The societal impact of robot-assisted mobility systems extends to environmental awareness and sustainability. Promoting the benefits of reduced emissions, energy efficiency, and sustainable

transportation options can garner societal support and encourage the adoption of these systems as a means to contribute to a greener future.

I. **Ethical Considerations and Values Alignment**: Aligning robot-assisted mobility systems with societal, ethical considerations is crucial for acceptance. Ensuring that these systems operate in a manner that respects privacy, avoids discrimination, and upholds societal values fosters trust and positive reception among the public.

By addressing the societal impact and fostering acceptance of robot-assisted mobility systems, we can create an environment where these technologies are embraced as transformative solutions that enhance transportation, improve accessibility, and contribute to a more sustainable and inclusive society.

3 Future Developments and Impact of Robot-Assisted Mobility Systems

Robot-assisted mobility systems have rapidly evolved in recent years, promising transformative changes in transportation. As technological advancements continue to push the boundaries of robotics, artificial intelligence, and automation, the future holds exciting possibilities for the integration of robots in various aspects of mobility. This section explores the potential future developments and the profound impact that robot-assisted mobility systems can have on transportation, society, and the overall urban landscape. By examining emerging trends and envisioning potential scenarios, this section aims to provide insights into the transformative potential of these systems and the implications they may have in shaping the future of mobility.

3.1 Advancements in Robot-Assisted Mobility Technologies

Advancements in robot-assisted mobility technologies are bringing us closer to a future of intelligent and efficient mobility solutions. The integration of advanced autonomous navigation systems, sensor fusion, and machine learning paves the way for robots to navigate complex environments with precision and adaptability. Sensor fusion techniques enable robots to gather comprehensive data from multiple sensors, providing them with a holistic understanding of their surroundings and enhancing their decision-making abilities.

A. **Advanced Autonomous Navigation Systems**: We expect to see significant advancements in autonomous navigation systems, employing new algorithms, sensor fusion techniques, and machine learning capabilities to enable robots to navigate complex environments efficiently.
B. **Sensor Fusion**: The integration of multiple sensors, such as LiDAR, cameras, and radar, will enable robots to gather rich and comprehensive data about their

surroundings, obtaining accurate perception of the environment and allowing them to make informed decisions for safe and dynamic moving in unpredictable scenarios and do manipulations to their environment [55].

C. **Machine Learning and Artificial Intelligence**: AI and ML algorithms enable robots to learn from their interactions with the environment, notice new patterns, and use past experiences to avoid problematic situations, in addition to change actions to a more sustainable direction and in the long term help more sustainable product design too [56].
D. **Human–Robot Interaction**: Future advancements will focus on enhancing human–robot interaction in the context of mobility systems. Natural language processing, gesture recognition, and facial expression analysis enable robots to understand and respond to human commands and inquiries effectively. The result is more intuitive and user-friendly interactions with robot assistants.
E. **Integration of 5G Connectivity**: The integration of 5G connectivity in robot-assisted mobility systems shall enable faster and more reliable communication between robots, infrastructure, and other vehicles. This high-speed and low-latency connectivity facilitates real-time data exchange for quick responses and actions in changing conditions.
F. **Edge Computing and Cloud Robotics**: The combination of edge and cloud computing gives robots access to powerful computational resources and leverage big data for decision-making and anomaly detection [57]. Edge computing allows faster processing and reduced reliance on cloud connectivity, while cloud robotics enables shared knowledge access and leverages advanced robotics algorithms for complex tasks.
G. **Multi-Robot Collaboration**: Advancements in robot-assisted mobility technologies will foster the collaboration between multiple robots in a coordinated manner. This collaboration can enhance the efficiency of transportation systems, allowing robots to work together to optimize traffic flow, share information, and adapt to dynamic conditions.
H. **Enhanced Safety Features**: Future developments will focus on improving safety features in robot-assisted mobility systems, including the development of advanced collision avoidance systems [58, 59], for self-moving robots, real-time monitoring of the environment, and predictive analytics for accident prevention and hazards anticipation.
I. **Integration of Augmented Reality**: Augmented reality technology can be integrated into robot-assisted mobility systems to enhance the user experience. Passengers will receive real-time information and visual cues through augmented reality interfaces, making navigation and interaction with the transportation system more intuitive and engaging.
J. **Energy Efficiency and Sustainability**: Future advancements in robot-assisted mobility technologies should focus on improved energy efficiency, comparing different mobility options in emissions and investments like it is done in the energy production side, reduced CO_2 emissions, and promotion of total sustainability. This includes the development of energy-efficient propulsion

systems, integration of renewable energy sources, and optimization algorithms to minimize energy consumption and environmental impact.

3.2 Implications and Impact on Mobility and Society

The integration of robot-assisted mobility systems brings forth far-reaching implications and impacts on both mobility and society. From transforming how we move to influence social dynamics and urban planning, the implications of these systems extend beyond mere convenience.

A. **Improved Transportation Efficiency**: Improved transportation efficiency is achieved by optimizing traffic flow, reducing congestion, and minimizing travel times. Intelligent routing algorithms and real-time data analysis help to identify the most efficient routes, reducing delays and enhancing overall mobility.
B. **Enhanced Safety**: Robot-assisted mobility systems have the potential to enhance safety on the roads. Advanced collision avoidance systems, real-time monitoring, and predictive analytics mitigate risks and accidents. By reducing human error and improving response times, these systems contribute to safer transportation experiences with autonomous navigation.
C. **Increased Accessibility**: Robot-assisted mobility systems increase accessibility to transportation for individuals with disabilities, elderly populations, and those in underserved areas. These systems offer personalized assistance, provide accessible interfaces, and accommodate diverse mobility needs, ensuring equitable access to transportation options.
D. **Reduced Environmental Footprint**: Electric-powered robots and autonomous vehicles help reduce greenhouse gas emissions and air pollution compared to traditional transportation modes. Optimizing traffic flows reduce congestion and lead to more fuel-efficient travel.
E. **Urban Planning and Infrastructure**: The deployment of robot-assisted mobility systems can influence urban planning and infrastructure development. Efficient transportation systems shape urban design, leading to smarter, more sustainable, and livable cities. Optimized road networks, designated lanes for autonomous vehicles, and improved public transportation integration are examples of how these systems can influence urban infrastructure.
F. **Employment and Workforce Changes**: The adoption of robot-assisted mobility systems may have implications for employment in the transportation sector. These technologies may lead to job displacement in certain areas, and they also create new employment opportunities in new areas such as robot maintenance, system monitoring, and data analysis. Ensuring a just transition and supporting affected workers through reskilling and reemployment initiatives is crucial.
G. **Social Dynamics and User Acceptance**: The integration of robot-assisted mobility systems can impact social dynamics and user acceptance. Public attitudes, acceptance, and trust in these technologies play a significant role in their successful implementation. Building public awareness, addressing concerns, and

fostering inclusivity in the design and deployment of these systems are essential for widespread acceptance.

H. **Data Privacy and Security**: The implementation of robot-assisted mobility systems involves the collection and processing of vast amounts of data. Ensuring robust data privacy and security measures is critical to protect individuals' personal information and maintain public trust. Implementing secure data protocols, encryption techniques, and transparent data handling practices are a good starting point to enhance data privacy and security matters.

I. **Economic Implications**: The adoption of robot-assisted mobility systems can have economic implications at various levels. The development and deployment of these technologies can stimulate innovation, create new job opportunities, and drive economic growth. Additionally, cost savings resulting from increased transportation efficiency and reduced congestion can have positive economic impacts on individuals, businesses, and governments.

J. **Social Equity and Inclusion**: Robot-assisted mobility systems will have impact on social equity, equality, and inclusion. Ensuring that these systems are accessible to all, including disadvantaged communities, with reasonable price range with any potential biases addressed and discrimination removed, promoting good public/mass transport system.

The implications and impact of robot-assisted mobility systems on mobility and society are multifaceted, encompassing aspects such as efficiency, safety, accessibility, environmental sustainability, urban planning, employment, social dynamics, and equity. Understanding and addressing these implications are vital for harnessing the full potential of the technology development while minimizing any potential negative consequences.

4 Discussions and Conclusion

The study highlighted the implications of robot-assisted mobility systems in public and robotized transportation. Through an exploration of technological advancements, such as autonomous navigation, sensor fusion, machine learning, and human–robot interaction, it is evident that these technologies hold great potential to revolutionize the way we envision transportation when integrating robot-assisted mobility systems to current traditional transportation systems, numerous benefits can be seen. These include improved transportation efficiency, enhanced safety, increased accessibility, improved sustainability, and reduced environmental carbon footprint. The systems can optimize traffic flow, move private and public transportation to a wider range of road options between certain destinations to reduce general congestion, employ advanced safety and protection features, provide equitable access to transportation for all individuals, and contribute to sustainability efforts.

Furthermore, the impact of robot-assisted mobility systems extends beyond transportation itself. It influences urban planning and infrastructure development, creates new employment opportunities, and shapes social dynamics. Ensuring public acceptance and addressing concerns regarding data privacy, security, and social equity are crucial for successful implementation. In conclusion, the summarized findings underscore the transformative potential of robot-assisted mobility systems and their profound impact on mobility and society. Embracing these advancements while considering the associated challenges is essential for shaping a future of transportation that is efficient, safe, accessible, environmentally sustainable, and socially inclusive. By harnessing the power of these technologies and robotization, our societies can pave the way for a transportation ecosystem that meets the evolving needs of individuals and communities, enhancing the quality of life for all, although the industry might need new ways to innovate [60] platform, networks, and deeper trust-based [61] co-operation models, with proper consideration for cultural effects [62], in modern research and development units' collaboration forms [30] for sustainable business model development [63], to jump in next society 2.0 level.

References

1. Mahalakshmi R, Kavitha M, Gopi B, Kumar SM (2023) Women safety night patrolling IoT robot. In: 2023 5th ICSSIT, IEEE, pp 544–549
2. Janjua MB, Arslan H (2023) A survey of symbiotic radio: methodologies, applications, and future directions. Sensors 23(5):2511
3. Jafar RMS, Ahmad W, Sun Y (2023) Unfolding the impacts of metaverse aspects on telepresence, product knowledge, and purchase intentions in the metaverse stores. Technol Soc 74:102265
4. Ferraguti F, Farsoni S, Bonfè M (2022) Augmented reality and robotic systems for assistance in percutaneous nephrolithotomy procedures: recent advances and future perspectives. Electronics 11(19):2984
5. Fu Y, Lin W, Yu X, Rodríguez-Andina JJ, Gao H (2023) Robot-assisted teleoperation ultrasound system based on fusion of augmented reality and predictive force. IEEE Tran Ind Electron 70:7449–7456
6. Fareh R, Siddique T, Rashid H et al (2022) Development of a robot-assisted rehabilitation program for upper-extremity disabilities. ICRAE, pp 137–141
7. Selvaganapathy SG, Hema Priya N et al (2023) Machine learning based robotic-assisted upper limb rehabilitation therapies: a review. Comput Vis Robot: Proc CVR 2022:59–73
8. Abdelsalam A, Happonen A, Kärhä K, Kapitonov A, Porras J (2022) Toward autonomous vehicles and machinery in mill yards of the forest industry: technologies and proposals for autonomous vehicle operations. IEEE Access 10:88234–88250. https://doi.org/10.1109/ACCESS.2022.3199691
9. García Alcaraz JL, Díaz Reza JR, Arredondo Soto KC et al (2022) Effect of green supply chain management practices on environmental performance: case of Mexican manufacturing companies. Mathematics 10(11):1–19. https://doi.org/10.3390/math10111877
10. Oubrahim I, Sefiani N, Happonen A (2022) Supply chain performance evaluation models: a literature review. Acta Logist 9(2):207–221. https://doi.org/10.22306/al.v9i2.298
11. Hemilä J, Salmela E, Happonen A (2007) The role of the logistics service provider in VMI operations. In ISL 2007:449–454. https://doi.org/10.5281/zenodo.3376516
12. Oubrahim I, Sefiani N, Happonen A (2023) The influence of digital transformation and supply chain integration on overall sustainable supply chain performance: an empirical analysis from manufacturing companies in morocco. Energies 16:2–25. https://doi.org/10.3390/en16021004

13. Minashkina D, Happonen A (2020) Decarbonizing warehousing activities through digitalization and automatization with WMS integration for sustainability supporting operations. In: E3S Web of conferences, vol 158, Article: 03002, pp 1–7, https://doi.org/10.1051/e3sconf/202015803002
14. Happonen A, Tikka M, Usmani U (2021) A systematic review for organizing hackathons and code camps in Covid-19 like times: literature in demand to understand online hackathons and event result continuation. In: 2021 International conference on data and software engineering (ICoDSE), pp 7–12, https://doi.org/10.1109/ICoDSE53690.2021.9648459
15. Auvinen H, Santti U, Happonen A (2020) Technologies for reducing emissions and costs in combined heat and power production. In: E3S Web conforence, vol 158, pp 1–6, https://doi.org/10.1051/e3sconf/202015803006
16. Uma S (2023) Blockchain and AI: disruptive digital technologies in designing the potential growth of healthcare industries. In: AI and blockchain in healthcare, pp 137–150
17. Zahra A, Sood N, Lala SRF, Alam S (2023) Next-generation technologically empowered telehealth systems. In: Extended reality for healthcare systems, pp 51–75
18. Peralta-Ochoa AM, Chaca-Asmal PA, Guerrero-Vásquez LF, Ordoñez-Ordoñez JO, Coronel-González EJ (2023) Smart healthcare applications over 5G networks: a systematic review. Appl Sci 13(3):1469
19. Manjila S, Rosa B, Price K, Manjila R, Mancatelli M, Dupont PE (2023) Robotic instruments inside the MRI bore: key concepts and evolving paradigms in imaging-enhanced cranial neurosurgery. World Neurosurg 11(176):127–139
20. Unger M, Berger J, Melzer A (2021) Robot-assisted image-guided interventions. Front Robot AI 8:664622
21. Soliman C, Furrer MA, Lawrentschuk N (2023) New robots and how this has changed operative technique in renal cancer surgery. In: Robotic surgery for renal cancer, pp 99–110
22. Schonfeld E, Stienen MN, Veeravagu A (2022) Future perspective of robot-assisted minimally invasive spine surgery. In: Technical advances in minimally invasive spine surgery: navigation, robotics, endoscopy, augmented and virtual reality, pp 351–364
23. Rayner R, Kerwin K, Valentine N (2022) Robot-assisted teaching-the future of education? In: EcoMechatronics: challenges for evolution, development and sustainability, pp 329–357
24. Ke X, Yu Y, Li K et al (2023) Review on robot-assisted polishing: status and future trends. Robot Comput-Integr Manuf 80:102482
25. Xiao B, Chen C, Yin X (2022) Recent advancements of robotics in construction. Autom Constr 144:104591
26. Happonen A, Minashkina D (2020) State of the art preliminary literature review: sustainability and waste reporting capabilities in management systems. In: E3S Web of conference 211:03014–12. https://doi.org/10.1051/e3sconf/202021103014
27. Durrer J, Agrawal P, Ozgul A, Neuhauss SC, Nama N, Ahmed D (2022) A robot-assisted acoustofluidic end effector. Nat Commun 13(1):6370
28. Jafari N, Lim M, Hassani A, Cordeiro J, Kam C, Ho K (2022) Human-like tele-health robotics for older adults-a preliminary feasibility trial and vision. J Rehabil Assist Technol Eng 9:20556683221140345
29. Lorson F, Fügener A, Hübner A (2023) New team mates in the warehouse: human interactions with automated and robotized systems. IISE Trans 55(5):536–553
30. Happonen A, Minashkina D, Nolte A, Medina Angarita MA (2020) Hackathons as a company—University collaboration tool to boost circularity innovations and digitalization enhanced sustainability. AIP Conf Proc 2233(1):1–11. https://doi.org/10.1063/5.0001883
31. Metso L, Happonen A, Rissanen M (2022) Estimation of user base and revenue streams for novel open data based electric vehicle service and maintenance ecosystem driven platform solution, Lecture Notes in Mechanical Engineering, pp 393–404, https://doi.org/10.1007/978-3-030-93639-6_34
32. Metso L, Happonen A (2021) Data sharing concept for electric car services: fleet level optimization and emission reduction based on monitored data. In: MCMD 2020 conference, p 9, 16–18 Feb

33. Kortelainen H, Happonen A, Kinnunen S-K (2016) Fleet service generation-challenges in corporate asset management, Lecture Notes in Mechanical Engineering, Springer, Heidelberg, pp 373–380, https://doi.org/10.1007/978-3-319-27064-7_35
34. Usmani UA, Watada J, Jaafar J et al (2021) A reinforcement learning based adaptive ROI generation for video object segmentation. IEEE Access 9:161959–161977
35. Usmani UA, Happonen A, Watada J (2022) Enhanced deep learning framework for fine-grained segmentation of fashion and apparel. In: Intelligent computing: proceedings of the 2022 computing conference, vol 2, pp 29–44
36. Lehtinen M, Happonen A, Ikonen J (2008) Accuracy and time to first fix using consumer-grade GPS receivers. In: 16th international conference on software, telecommunications and computer networks, pp 334–340, https://doi.org/10.1109/SOFTCOM.2008.4669506
37. Jahkola O, Happonen A, Knutas A, Ikonen J (2017) What should application developers understand about mobile phone position data. In: CompSysTech'17, ACM, pp 171–178, https://doi.org/10.1145/3134302.3134346
38. Metso L, Happonen A, Rissanen M, Efvengren K, Ojanen V, Kärri T (2020) Data openness based data sharing concept for future electric car maintenance services, in advances in asset management and condition monitoring. Smart Innov Syst Technol 166(36):429–436. https://doi.org/10.1007/978-3-030-57745-2_36
39. Gala UB, Narwane VS, Raut RD, U.H., Govindarajan, et al (2021) A viability study using conceptual models for last mile drone logistics operations in populated urban cities of India. IET Collab Intell Manuf 3(3):262–272. https://doi.org/10.1049/cim2.12006
40. Chen Z, Zhang H, Zaferiou A, Zanotto D, Guo Y (2022) Mobile robot assisted gait monitoring and dynamic margin of stability estimation. IEEE Trans Med Robot Bionics 4(2):460–471
41. Usmani UA, Usmani MU (2014) Future market trends and opportunities for wearable sensor technology. Int J Eng Technol 6(4):326
42. Bogue R (2023) The first half century of industrial robot: 50 years of robotic developments. Ind Robot 50(1):1–10
43. Usmani UA, Roy A, Watada J et al (2022) Enhanced reinforcement learning model for extraction of objects in complex imaging. In: 2021 computing conforence, vol 1, pp 946–964
44. Tushar W, Yuen C, Saha TK, Nizami S, Alam MR, Smith DB, Poor HV (2023) A survey of cyber-physical systems from a game-theoretic perspective. IEEE Access 11:9799–9834
45. Hasan SK, Dhingra AK (2022) Biomechanical design and control of an eight DOF human lower extremity rehabilitation exoskeleton robot. Results Control Optim 7:100107
46. Usmani UA, Haron NS, Jaafar J (2021) A natural language processing approach to mine online reviews using topic modelling. In: COMS2 2021, Rev. Selected Papers, pp 82–98
47. Usmani UA, Watada J et al (2022) A systematic review of privacy-preserving blockchain in e-medicine. Biomedical and other applications of soft computing. Springer, Cham, pp 25–40
48. Moniruzzaman MD, Rassau A, Chai D, Islam SMS (2022) Teleoperation methods and enhancement techniques for mobile robots: a comprehensive survey. Robot Auton Syst 150:103973
49. Usmani UA, Jaafar J (2022) Machine learning in healthcare: current trends and the future. In: International conference on artificial intelligence for smart community: AISC 2020:659–675
50. Pierce K, Pepler DJ, Craig SG, Jenkin M (2023) Considerations for developing robot-assisted crisis de-escalation practices. Appl Sci 13(7):4337
51. Usmani UA, Happonen A, Watada J (2022) Enhancing artificial intelligence control mechanisms: current practices, real life applications and future views. In: Proceedings of the future technologies conference (FTC), vol 1, pp 287–306
52. Shi Y, Wang T, Yu J, Xiao S, Xiong L, Yang L (2023) Virtual potential field-based motion planning for human-robot collaboration via kinesthetically guided teleoperation. In: 2023 7th ICRCA, pp 37–44
53. Ilyas M, Khaw HY, Selvaraj NM, Jin Y et al (2021) Robot-assisted object detection for construction automation: data and information-driven approach. IEEE Trans Mechatron 26(6):2845–2856

54. Rasouli S, Gupta G, Nilsen E, Dautenhahn K (2022) Potential applications of social robots in robot-assisted interventions for social anxiety. Int J Soc Robot 14(5):1–32
55. Park D, Hoshi Y, Mahajan HP et al (2020) Active robot-assisted feeding with a general-purpose mobile manipulator: design, evaluation, and lessons learned. Robot Auton Syst 124:103344
56. Ghoreishi M, Happonen A (2020) Key enablers for deploying artificial intelligence for circular economy embracing sustainable product design: three case studies. AIP Conf Proc 2233(1):1–19. https://doi.org/10.1063/5.0001339
57. Usmani UA, Happonen A, Watada J (2022) A review of unsupervised machine learning frameworks for anomaly detection in industrial applications, Lecture Notes in Networks and Systems, vol 507, Chapter: 11, pp 158–189, https://doi.org/10.1007/978-3-031-10464-0_11
58. Vann W, Zhou T, Zhu Q, Du E (2023) Enabling automated facility maintenance from articulated robot Collision-Free designs. Adv Eng Inform 55:101820
59. Usmani UA, Watada J, Jaafar J, Aziz IA, Roy A (2021) Particle swarm optimization with deep learning for human action recognition. Int J Innov Comput Inform Control 17(6):1843–1870
60. Happonen A, Santti U, Auvinen H, Räsänen T, Eskelinen T (2020) Digital age business model innovation for sustainability in university industry collaboration model. In: E3S web of conforence 211(04005):1–11. https://doi.org/10.1051/e3sconf/202021104005
61. Happonen A, Siljander V (2020) Gainsharing in logistics outsourcing: trust leads to success in the digital era. IJCENT 6(2):150–175. https://doi.org/10.1504/IJCENT.2020.110221
62. Santti U, Eskelinen T, Rajahonka M et al (2017) Effects of business model development projects on organizational culture: a multiple case study of smes. Technol Innov Manag Rev 7(8):15–26. https://doi.org/10.22215/timreview/1096
63. Eskelinen T, Räsänen T, Santti U et al (2017) Designing a business model for environmental monitoring services using fast MCDS innovation support tools. Technol Innov Manag Rev 7(11):36–46. https://doi.org/10.22215/timreview/1119

Author Index

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 754, https://doi.org/10.1007/978-981-99-4932-8