Lecture Notes in Networks and Systems 782

Simon Fong
Nilanjan Dey
Amit Joshi *Editors*

ICT Analysis and Applications

Proceedings of ICT4SD 2023, Volume 2

Springer

Lecture Notes in Networks and Systems

Volume 782

Series Editor

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Advisory Editors

Fernando Gomide, Department of Computer Engineering and Automation—DCA, School of Electrical and Computer Engineering—FEEC, University of Campinas—UNICAMP, São Paulo, Brazil

Okyay Kaynak, Department of Electrical and Electronic Engineering, Bogazici University, Istanbul, Türkiye

Derong Liu, Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, USA

Institute of Automation, Chinese Academy of Sciences, Beijing, China

Witold Pedrycz, Department of Electrical and Computer Engineering, University of Alberta, Alberta, Canada

Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Marios M. Polycarpou, Department of Electrical and Computer Engineering, KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus

Imre J. Rudas, Óbuda University, Budapest, Hungary

Jun Wang, Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong

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).

Simon Fong · Nilanjan Dey · Amit Joshi
Editors

ICT Analysis and Applications

Proceedings of ICT4SD 2023, Volume 2

Springer

Editors
Simon Fong
University of Macau
Macau, Macao

Nilanjan Dey
Department of Computer Science
and Engineering
Techno International New Town
Kolkata, India

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-99-6567-0 ISBN 978-981-99-6568-7 (eBook)
https://doi.org/10.1007/978-981-99-6568-7

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Paper in this product is recyclable.

Preface

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

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

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

Research submissions in various advanced technology areas were received and after a rigorous peer-reviewed process with the help of program committee members and 185 external reviewers for 1100+ papers from 16 different countries including Saudi Arabia, USA, Singapore, Denmark, Norway, Denmark, Ghana, United Arab Emirates, Netherlands, Iraq, Bangladesh, Japan, Malaysia, 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 physically at Goa and virtually on Zoom including one inaugural and one keynote session.

Technology is the driving force of progress in this era of globalization. Information and communications technology (ICT) has become a functional requirement for the socioeconomic growth and sustained development of any country. The influence of information and 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 Startup 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.

Macau, Macao Simon Fong
Kolkata, India Nilanjan Dey
Ahmedabad, India Amit Joshi

Contents

Editors and Contributors

About the Editors

Simon Fong graduated from La Trobe University, Australia, with a 1st Class Honors B.Eng. Computer Systems degree and a Ph.D. Computer Science degree in 1993 and 1998, respectively. He is now working as Associate Professor in the Computer and Information Science Department of the University of Macau. Prior to his academic career, he took up various IT posts in Melbourne, Hong Kong, and Singapore. He has published over 396 international conference and journal papers, mostly in the areas of data mining and metaheuristics. He is also one of the founding members of the Data Analytics and Collaborative Computing Research Group in the Faculty of Science and Technology.

Nilanjan Dey is Associate Professor at Techno India College of Engineering, Kolkata, India. He is Visiting Fellow at the Department of Biomedical Engineering, University of Reading, UK, and Visiting Professor at Duy Tan University, Vietnam. He was Honorary Visiting Scientist at Global Biomedical Technologies Inc., CA, USA (2012–2015). He is Research Scientist of the Laboratory of Applied Mathematical Modeling in Human Physiology, Territorial Organization of Scientific and Engineering Unions, Bulgaria, and Associate Researcher of Laboratoire RIADI, University of Manouba, Tunisia. He is Scientific Member of Politécnica of Porto. He is Affiliated Member of Signal Processing and Computational Intelligence Applied to Power Systems (PSCOPE) research group, Federal University of Juiz de Fora, Brazil. He was awarded his Ph.D. from Jadavpur University in 2015.

Dr. Amit Joshi is currently Director of Global Knowledge Research Foundation, also Entrepreneur and Researcher who has completed his graduation (B.Tech.) in Information Technology and M.Tech. in Computer Science and Engineering and completed his research in the areas of cloud computing and cryptography in medical imaging with a focus on analysis of the current Government Strategies and World forums needs in different sectors on security purposes. He has an experience of around 12 years in academic and industry in prestigious organizations. He is Active Member of ACM, IEEE, CSI, AMIE, IACSIT-Singapore, IDES, ACEEE, NPA, and many other professional societies. Currently, he is Chairman of Computer Society of India (CSI) Udaipur Chapter and Secretary for the Association of Computing Machinery (ACM) Udaipur Professional Chapter. Further currently, he is also the International Chair of InterYIT at the International Federation of Information Processing (IFIP, Austria). He has presented and published more than 50 papers in national and international journals/conferences of IEEE and ACM. He has also edited more than 20 books which are published by Springer, ACM, and other reputed publishers. He has also organized more than 40 national and international conferences and workshops through ACM, Springer, and IEEE across five countries including India, UK, Thailand, and Europe. Apart from this academic involvement, he is also associated with the business and industry community across the globe. He is currently the Director of Emanant TechMedia (P) Limited which focuses on ICT and web-based services.

Contributors

Nasrin Aasofwala LJ University, Ahmedabad, India

Basant Agarwal Department of Computer Science and Engineering, Central University of Rajasthan, Ajmer, Rajasthan, India

Priya Aggarwal IT Department, Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi, India

Renuka Agrawal Symbiosis Institute of Technology, Pune, Maharashtra, India

G. Akhil Lovely Professional University, Delhi, Punjab, India

Esra Albhnasawi Department of Computer Science, King Faisal University, Al Hofuf, Kingdom of Saudi Arabia

Wadha Alfenais Department of Computer Science, King Faisal University, Al Hofuf, Kingdom of Saudi Arabia

Noura Almusaynid Department of Computer Science, King Faisal University, Al Hofuf, Kingdom of Saudi Arabia

Batool Alsowaiq Department of Computer Science, King Faisal University, Al Hofuf, Kingdom of Saudi Arabia

Gauri Amin Computer Department, Cummins College of Engineering for Women, Pune, Maharashtra, India

B. Aravind Reddy B V Raju Institute of Technology, Electronics and Communication Engineering, Narsapur, Medak, India

T. Arumuganathan ICAR-Sugarcane Breeding Institute, Coimbatore, India

M. Ashwin PES University, Bengaluru, India

Anjali Askhedkar Dr. Vishwanath Karad, MIT World Peace University, Pune, India

Rajas Aswale Pimpri Chinchwad College of Engineering, Pune, India

R. N. Awale VJTI, Mumbai University, Matunga, India

Yash Ayare Department of AI &DS, Vishwakarma Institute of Information Technology, Pune, India

Bibi Erum Ayesha Chaitanya Bharathi Institute of Technology(A), Hyderabad, India

Palamakula Ramesh Babu Chaitanya Bharathi Institute of Technology(A), Hyderabad, India

Shravani Bahulekar Vishwakarma Institute of Technology, Pune, Maharashtra, India

Parashuram Bannigidad Department of Computer Science, Rani Channamma University, Belagavi, India

Alpa R. Barad GLS University, Ahmedabad, India

Prafulla Kumar Bariha Department of Medicine, Veer Surendra Sai Institute of Medical Sciences and Research, Burla, India

Janki Barot Silver Oak University, Ahmedabad, India

Santi Kumari Behera Department of Computer Science Engineering, VSSUT, Burla, India

Pooja Bhati MAE Department, Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi, India

Ankit R. Bhavsar GLS University, Ahmedabad, India

Ganesh Bhutkar Vishwakarma Institute of Technology (VIT), Pune, India

P. T. Bhuvan PES University, Bengaluru, India

Awantika Biswal Department of Electronic Communication, SUIIT, Sambalpur University, Sambalpur, India

Monali R. Borade Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India

A. C. Chandana PES University, Bengaluru, India

Harshada Chavan MKSSS's Cummins College of Engineering for Women, Pune, India

Shakti Raj Chopra Lovely Professional University, Phagwara, India

Pranav Dani Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

Shruti Danve Dr. Vishwanath Karad, MIT World Peace University, Pune, India

Poorva Darade MKSSS's Cummins College of Engineering for Women, Pune, India

Ashis Das Department of Computer Science and Engineering, SUIIT, Sambalpur University, Sambalpur, India

Anuradha S. Deshpande JSPM's Imperial College of Engineering and Research, Pune, India

Manikrao Dhore Vishwakarma Institute of Technology, Pune, Maharashtra, India

Amol Dhumane Pimpri Chinchwad College of Engineering, Pune, India

Mandar P. Diwakar Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India

Gouravi Dorkhande Computer Department, Cummins College of Engineering for Women, Pune, Maharashtra, India

Smrity Dwivedi Department of Electronics Engineering, IIT-BHU, Varanasi, India

Laukik Gadam JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Monica Gahlawat L.J. Institute of Computer Applications, LJ University, Ahmedabad, India

Pimmy Gandotra Indian Institute of Technology, Delhi, New Delhi, India;
School of Electronics and Communication Engineering, SMVDU, Katra, J&K, India;
Department of Electronics and Communication Engineering, Jaypee Institute of Information Technology, Noida, India

Dweepna Garg Department of Computer Engineering, Devang Patel Institute of Advance Technology and Research (DEPSTAR), Faculty of Technology and Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

Nysa Gaur On My Own Technology, Mumbai, India

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

H. M. Gireesha KLE Technological University, Hubballi, Karnataka, India

Mohit Kumar Goel Lovely Professional University, Delhi, Punjab, India

Vinaya Gohokar Dr. Vishwanath Karad, MIT World Peace University, Pune, India

J. S. Gokul Department of ISE, CMR Institute of Technology, Bengaluru, India

Avinash Golande JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Aadya Golwalkar Medi-Caps University, Indore, India

Sudhanshu Gonge Symbiosis Institute of Technology, Pune, Maharashtra, India

Santosh Gore Sai Info Solution, Nashik, India

Chhaya Gosavi Computer Department, Cummins College of Engineering for Women, Pune, Maharashtra, India

V. Gowri Manohari Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Aditya Gujrathi JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, India

Punam Gulande VJTI, Mumbai University, Matunga, India

Shrawani Gulhane Department of AI &DS, Vishwakarma Institute of Information Technology, Pune, India

B. Hari Kiran Lovely Professional University, Delhi, Punjab, India

Priyanka Harjule Department of Mathematics, Malaviya National Institute of Technology, Jaipur, Rajasthan, India

Vinayak Hegde Department of Computer Science, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Mysuru, India

Darshan Ingle Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

Yashwant S. Ingle Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India

Nalini C. Iyer KLE Technological University, Hubballi, Karnataka, India

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

Reetu Jain On My Own Technology, Mumbai, India

Sanjay Kumar Jain Department of Computer Engineering, NIT Kurukshetra, Thanesar, India

Surbhi Jangid Department of Computer Engineering, NIT Kurukshetra, Thanesar, India

Vihaan Jasani Department of Information Technology, K.J. Somaiya Institute of Technology, Mumbai, India

Sangeeta Jaybhaye Vishwakarma Institute of Technology, Pune, Maharashtra, India

B. Jegan Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

Rakesh Kumar Jha Indian Institute of Information Technology, Design and Mnaufacturing, Jabalpur, Madhya Pradesh, India

A. Jitendra Lovely Professional University, Delhi, Punjab, India

J. Joan Niveda Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Rahul Joshi Symbiosis Institute of Technology, Pune, Maharashtra, India

Sampada Joshi MKSSS's Cummins College of Engineering for Women, Pune, India

Yash Kadam Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India

Sahil Kakurle Vishwakarma Institute of Technology, Bibwewadi, Pune, India

Abhijeet Kamble Vishwakarma Institute of Technology, Bibwewadi, Pune, India

Malhar Kapshe Vishwakarma Institute of Technology, Bibwewadi, Pune, India

K. Karunya Sathyabama Institute of Science and Technology, Chennai, India

T. P. Kausalya Nandan B V Raju Institute of Technology, Electronics and Communication Engineering, Narsapur, Medak, India

N. Kavya SJCIT, Chickballapur, India

Omkar Khade Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India

Chaitra L. Korlahalli KLE Technological University, Hubballi, Karnataka, India

Jinal Kotadia Department of Computer Science and Engineering, Devang Patel Institute of Advance Technology and Research (DEPSTAR), Faculty of Technology and Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

Ketan Kotecha Symbiosis Institute of Technology, Pune, Maharashtra, India

B. Koushika Varsha B V Raju Institute of Technology, Electronics and Communication Engineering, Narsapur, Medak, India

Pinna Murali Krishna Lovely Professional University, Phagwara, India

Ramu Kuchipudi Chaitanya Bharathi Institute of Technology(A), Hyderabad, India

Jyoti Kukade Medi-Caps University, Indore, India

Achanta Tharun Kumar Lovely Professional University, Phagwara, India

Hari Pranav Arun Kumar Department of ISE, CMR Institute of Technology, Bengaluru, India

P. Kumar Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

Rahil Ratan Kumar School of Computer Science and Engineering), Galgotias University, Greater Noida, Uttar Pradesh, India

Santosh Kumar Department of AI &DS, Vishwakarma Institute of Information Technology, Pune, India

Brejesh Lall Indian Institute of Technology, Delhi, New Delhi, India

Indira Linginani VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India

Mokshit Lodha Dr. Vishwanath Karad, MIT World Peace University, Pune, India

Atharva Dhoble Hemakshi Lohiya Department of AI &DS, Vishwakarma Institute of Information Technology, Pune, India

Dheeraj B. Lokhande Shree Siddheshwar Women's College of Engineering, Solapur, India

R. Madhumathi Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

Tushar Mahajan JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, India

Parikshit N. Mahalle Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India

Nitin Mahankale Symbiosis Centre for Management Studies, Pune, India

M R. Mahesh NCET, Bengaluru, India

Dhruv R. Makadia Department of Computer Science and Engineering, Devang Patel Institute of Advance Technology and Research (DEPSTAR), Faculty of Technology and Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

S. Manasa SJCIT, Chickballapur, India

Ankita Mane Computer Department, Cummins College of Engineering for Women, Pune, Maharashtra, India

T N. Manjunath BMSIT, Bengaluru, India

Aarnav Marathe Department of Information Technology, K.J. Somaiya Institute of Technology, Mumbai, India

Ganesh Masute JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, India

Manoj Mathur School of Planning and Architecture, New Delhi, India

Deepanjali Mishra School of Humanities, KIIT University, Bhubaneswar, India

Trapti Mishra Medi-Caps University, Indore, India

Soumabha Mitra Sathyabama Institute of Science and Technology, Chennai, India

Amar Kumar Mohapatra IT Department, Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi, India

Bassam Hameed Mohsen Department of English, Hodiedah University, Hodiedah, Yemen

Abhay Monasara Department of Computer Science and Engineering, Devang Patel Institute of Advance Technology and Research (DEPSTAR), Faculty of Technology and Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

Sarvesh Morey JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, India

Swanand Mulay JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Anilkumar V. Nandi KLE Technological University, Hubballi, Karnataka, India

P. Nandini Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Bhawna Narwal IT Department, Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi, India

M. Naveenganesh Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

Pratima Nayak ApMoSys Technologies Pvt. Ltd., Navi Mumbai, India

Sandesh Nayak Department of Mines and Geology, Government of Rajasthan, Jaipur, India

Isha Nikam MKSSS's Cummins College of Engineering for Women, Pune, India

Pandit B. Nirmal Department of English, Sant Tukaram College of Arts and Science, Parbhani, India

P. C. Nissimagoudar KLE Technological University, Hubballi, Karnataka, India

Mukta Ogale Computer Department, Cummins College of Engineering for Women, Pune, Maharashtra, India

Prabhat Padhy ApMoSys Technologies Pvt. Ltd., Navi Mumbai, India

Manas Kumar Panigrahi Department of Computer Science and Engineering, SUIIT, Sambalpur University, Sambalpur, India

Dhairya Panjwani Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

Prashant Panse Medi-Caps University, Indore, India

V. Parvathy School of Humanities, KIIT University, Bhubaneswar, India

Kalyani Patel Gujarat University, Ahmedabad, India

Nidhi Patel Department of Information Technology, K.J. Somaiya Institute of Technology, Mumbai, India

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

Shubham Patel Symbiosis Institute of Technology, Pune, Maharashtra, India

Nikhil Patil Vishwakarma Institute of Technology (VIT), Pune, India

Nishant Patil JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Priyanshi Patil Vishwakarma Institute of Technology (VIT), Pune, India

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

Tejas Patil Symbiosis Institute of Technology, Pune, Maharashtra, India

Veeraboena Pavan Lovely Professional University, Delhi, Punjab, India

Sudeep Pawar Pimpri Chinchwad College of Engineering, Pune, India

Varsha Pimprale MKSSS's Cummins College of Engineering for Women, Pune, India

Parth Pishte Vishwakarma Institute of Technology (VIT), Pune, India

S. R. Pooja VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India

Pooja Department of Mathematics, Malaviya National Institute of Technology, Jaipur, Rajasthan, India

Mustafa Poonawala Dr. Vishwanath Karad, MIT World Peace University, Pune, India

D. Poornima Sathyabama Institute of Science and Technology, Chennai, India

Apoorva Popalghat Vishwakarma Institute of Technology (VIT), Pune, India

D. V. Pranav PES University, Bengaluru, India

Battini Pranay Lovely Professional University, Phagwara, India

Amarnath Pulagam Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Sanjeev Kr. Punia School of Computer Science and Engineering), Galgotias University, Greater Noida, Uttar Pradesh, India

S. Raghavendar Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

Swasti Rai JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Darshan Rander Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

Mansing Rathod Department of Information Technology, K.J. Somaiya Institute of Technology, Mumbai, India

Kyatham Hanishka Reddy VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India

K. Risheedhar Varma Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

V. Roshini Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

Nilesh P. Sable Department of Information Technology, Vishwakarma Institute of Information Technology, SPPU Pune, Pune, India

Surojit Saha Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

Raman Sahu Department of Electronic Communication, SUIIT, Sambalpur University, Sambalpur, India

S. P. Sajjan Department of Computer Science, Rani Channamma University, Belagavi, India

Salla Sanjana VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India

Suresh Sankrayananarayanan Department of Computer Science, King Faisal University, Al Hofuf, Kingdom of Saudi Arabia

T. Satyanarayana Murthy Chaitanya Bharathi Institute of Technology(A), Hyderabad, India

Tejas Sawant Pimpri Chinchwad College of Engineering, Pune, India

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

Chintan Shah Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India;
Department of Computer Engineering, College of Technology, Silver Oak University, Ahmedabad, Gujarat, India

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

Pratham Shah Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

Siddhanth Shah Symbiosis Institute of Technology, Pune, Maharashtra, India

Farhan Shaikh Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

M. Shalini Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

D. Sharath Srinivas Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

Anushka Sharma Medi-Caps University, Indore, India

Arun Sharma Indira Gandhi Delhi Technical University for Women, New Delhi, Delhi, India

Kamiya Sharma L.J. Institute of Computer Applications, LJ University, Ahmedabad, India

Kaivalya Shende Vishwakarma Institute of Technology, Pune, Maharashtra, India

Bhavya Sheth Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

Manish Shetkar Vishwakarma Institute of Technology, Bibwewadi, Pune, India

Mrunal Shidore Vishwakarma Institute of Technology, Bibwewadi, Pune, India

Gitanjali R. Shinde Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU, Pune, India

Sahil Shitole Vishwakarma Institute of Technology, Pune, Maharashtra, India

M R. Shrihari SJCIT, Chickballapur, India

Shishir Singh Pimpri Chinchwad College of Engineering, Pune, India

Shishir Shekhar Singh School of Computer Science and Engineering), Galgotias University, Greater Noida, Uttar Pradesh, India

Sandip Singha Indian Institute of Information Technology, Kalyani, India

Kshitij Kumar Sinha School of Planning and Architecture, New Delhi, India;
Indira Gandhi Delhi Technical University for Women, New Delhi, Delhi, India

Atharva Sonawane Dr. Vishwanath Karad, MIT World Peace University, Pune, India

C. H. Sri Harsha B V Raju Institute of Technology, Electronics and Communication Engineering, Narsapur, Medak, India

Srinath K. Sathyabama Institute of Science and Technology, Chennai, India

Thummala Srithika VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India

Shreesh Srivastava Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India

K. N. Sudhakar Department of ISE, CMR Institute of Technology, Bengaluru, India

Gandi Vinay Suhaas Lovely Professional University, Phagwara, India

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

Ravi Thacker Symbiosis Institute of Technology, Pune, Maharashtra, India

Siddhant Thaware Vishwakarma Institute of Technology, Pune, Maharashtra, India

S. Tirumala Raghu Vardhan Lovely Professional University, Delhi, Punjab, India

Hrithik Tomar Medi-Caps University, Indore, India

Siddhesh Vaidya ApMoSys Technologies Pvt. Ltd., Navi Mumbai, India

N. Venugopal PES University, Bengaluru, India

Shanti Verma LJ University, Ahmedabad, India

S. Vinodh Kumar Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

Arvind Vishnubhatla Electronics and Communications Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

A. V. Vismaya Department of Computer Science, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Mysuru, India

Analysis of Smart Health Monitoring Systems

Mohit Kumar Goel, G. Akhil, B. Hari Kiran, A. Jitendra, S. Tirumala Raghu Vardhan, and Veeraboena Pavan

Abstract Monitoring our biosignals has become critical recently, especially since the COVID-19 pandemic. Although, due to a shortage of supplies, many do not have access to health kits. In the wake of progress and development in electronic devices available to the public in addition to increasing healthcare costs, resulting in the culmination of the e-health enterprise which is the integration of medical science with engineering theories and practices. The incorporation of technologies such as electronics and communications along with computer sciences combined with the healthcare industry inquired the possibilities to develop a system that utilizes all the available medical knowledge resources and produces accurate, reliable, affordable, and efficient healthcare services to not only aged people in need and patients with illness and disabilities but also to people working odd hours who lack time. This paper sheds light upon the existing techniques available for healthcare applications, and the benefits and importance of IoT-based medical systems and the benefits of IoT in healthcare are discussed. Our aim is to provide a systematic review of the latest research on IoT-based health care by reviewing this literature. The proposed project is to design and build an intelligent healthcare system that is based on IoT which functions by collecting the medical data of the patient from the multiple sensors and transmitting the data that has been processed and analyzed to both the patient and their doctor. It remotely monitors the patient's biosignals and parameters and diagnoses the diseases as early as possible.

Keywords Patient health m · Arduino · Sensors · Pulse · Oxygen saturation · Electrocardiogram (ECG) · Temperature · Cloud · Internet of things

M. K. Goel (✉) · G. Akhil · B. Hari Kiran · A. Jitendra · S. Tirumala Raghu Vardhan · V. Pavan
Lovely Professional University, Delhi 144411, Punjab, India
e-mail: mohitgoel4u@gmail.com

G. Akhil
e-mail: Akhil.11903936@lpu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_1

1 Introduction

Health is one of the most paramount aspects of human life if not the most supreme of all because only when one leads a healthy life, they will have a strong body and mind which will allow one to pursue their dreams, contribute to society, spend time with their loved ones, and support themselves and their family by earning a living. The population in India is expanding at a rapid pace, according to the United Nations "With a total population projected to reach 1.5 billion in 2030 and 1.66 billion in 2050, India is anticipated to exceed China as the world's most populous country by the year 2024 and exceed China's 1.317 billion population in the next 50 years. At present, both countries make up about 40% of the world's total population." And this boom in population will in turn increase the needs and health issues. Furthermore, the doctor-to-patient ratio is quite bad in many of the states across India, i.e., there is a shortage of doctors in many states across India compared to the population of that state. Couple that with a global recession and inflation, and many people are reluctant to go visit a doctor for their health check-ups. Thus, to overcome this problem, the most optimal solution is the use of advanced technologies like Smart Health Monitoring based on the Internet of Things (IoT). This project utilizes these following technology and multiple other technologies such as microcontroller namely Arduino, multiple sensors, namely: temperature sensor, pulse/heart rate sensor, electrocardiogram (EGC), oxygen saturation/oximeter (SpO_2), and finally a cloud server to store and retrieve the data collected by the sensors.

1.1 Objective

The key objectives of this system are to increase the access to health care for people by making the process of health check-ups and diagnostics portable by making the device a single unit consisting of all the sensors integrated into the device, highly user-friendly by making the interface that the user interacts intuitive and easy to understand so that everyone irrespective of their technological literacy will be able to use the device, cost-effective as one system can be used by multiple users, multiple time just by investing once into the device itself, convenient as the user does not have to take an appointment, wait in line or travel to the clinic/hospital for their check-up, and efficient as the check-up process is extremely quick and saves a lot of time.

1.2 Targeted Users

The main reason for making the system portable is so that it can be accessed by anyone from anywhere, which will be extremely beneficial for the elderly and handicapped people. The system being highly user-friendly will make it so that people in the age range of 15 and above can use it. The system being cost-effective makes it very appealing as an affordable means for health care to many families as it would be a one-time investment rather than spending a consultation fee for each person every time they visit a doctor. This will in turn positively impact the health of many people who do avoid going to the doctor until the situation is extremely dire to save money. The system is also targeted at working class people who do not have the time and makes it much more convenient and efficient.

2 Literature Review

Trivedi et al. [1] have proposed a health monitoring system for a device controlled medical parameter prediction by microcontrollers in mobile phones. The data gathered by the sensor will be sent to the Arduino UNO microcontroller board in analog form. The recorded analog data will be converted into digital data via an analog-to-digital converter. By using Bluetooth, the system will be able to transmit the data wirelessly to the mobile phone. The limitation is that the modules used by Bluetooth devices are not versatile and have a small range.

Kumar et al. [2] built a customized IoT security monitoring system. The framework is configured by segmenting it into 3 layers, namely the control layer, device layer, and transport layer. A DS18B20 temperature sensor has been used to monitor the body temperature of a specific location, and a pulse rate sensor has been used for the monitoring of the pulse rate of the patient. An Arduino is used to integrate these sensors and a Wi-Fi module has been used to interface it with the internet in the transport layer. Finally, the data collected will be transmitted to the server. The limitation here is caused due to the use of Arduino UNO which does not support all sensors.

Desai et al. [3] proposed the implementation of wireless sensor networks (WSNs) to be deployed in smart homes. The benefit of Wireless Sensor Networks is that is integrates services like health monitoring, traffic data, etc., with home automation which make it extremely efficient and convenient for public use. The limitation is that it may have security concerns.

Mehta [4] proposed a mobile voice health monitoring system. The system is powered by a wearable accelerometer for voice detection and a mobile phone to process the collected raw data into recognizable form. The limitation of this is the need to have an accelerometer attached to the neck of the user which is not comfortable, and it is not completely accurate.

Gao [5] designed multi-lead electrocardiogram (ECG) monitoring system based on mobile phones. The system uses a seven-lead sensor which is used to monitor the ECG signal and is interfaced with a mobile phone to process the collected data. They also developed an alarm to detect the abnormal ECG data to mitigate the risk of heart disease. The limitation is the continuous flow of huge amount of data which needs high computation power and absence of wireless capability and can only monitor ECG.

Kong [6] designed a wireless health service system based on Android mobile phones. The proposed system consists of medical sensors and Bluetooth module that combine the technology of data collection of the sensors with wireless communication using Bluetooth with an android mobile phone following which it will be sent to a web server for storing the data of the patient as well as their family members. This is achieved using HTTP web server interface and MySQL database technologies. The limitation is that it is not a stand-alone device and needs an android mobile phone to function.

Yedukondalu Udara et al. [7] proposed a health monitoring system that is powered by an Arduino Mega 2560 Microcontroller along with, ADS1292R ECG shield, LM35 temperature sensor, ESP8266 Wi-Fi controller, and a LCD. The data collected via these sensors is displayed on the LCD so that the system can even be used without an app locally and for more information the data sent to a web server can be viewed via an app. The limitation is that the EGC sensor used is costly and the ECG data cannot be displayed on LCD.

Shubham Banka et al. [8] have proposed a system that is powered by a Raspberry. Various biosignals such as blood pressure, body temperature, and heartbeat are monitored using sensors. The data collected is processed on the Raspberry Pi which runs on Linux operating system and the processed data is uploaded to a web database through the GSM module interfaced with the Raspberry Pi which make the monitoring of health parameters accessible remotely. The limitation is the high cost of the Raspberry Pi in contrast to Arduino.

Tamilselvi et al. [9] developed a system which monitors various attributes of a patient such as eye blink and body movement in addition to basic parameters such as heart rate, oxygen saturation, and body temperature. It is powered by Arduino UNO microcontroller along with a GSM module to transmit the data to a web server. The limitation of this is that the eye blink and body movement don't provide any significant medical insights and are not useful for health monitoring.

Banerjee et al. [10] have proposed a device that utilizes concept of photoplethysmography. Infrared light emitting diodes (LED) and photodiodes work as the sensor to detect the heartbeat/pulse on the artery of the fingertip. The pulse produced due to a heart bear creates a signal which is captured by the sensor and is sent to a counter via a circuit of timer and relay switching. The counter upon getting the signal counts the heartbeats and gives a digital output.

Gregoski et al. [11] developed a smartphone heart rate monitoring application. This system is a smartphone application that utilizes the flash on the rear camera along with the camera itself to detect pulse in the fingertip. The system is extremely versatile as no additional hardware is necessary and just a basic smartphone is the only thing needed. The limitation is that the flash and the camera are not designed for this application, and the output is not as accurate as the dedicated sensors.

3 Smart Remote Health Monitoring Systems Applications

Before IoT, communication between patients and doctors was limited to physical visits, telephone, and text messaging. Furthermore, there was no way doctors could monitor the health of patients continuously and make recommendations accordingly. IoT technology has made remote health monitoring possible in the healthcare sector, unlocking the ability for the patients to be healthy, safe and allowing medical practitioners to provide better care remotely. This in turn has resulted in increased patient satisfaction as patients are able to interact and engage with doctors more than ever before. Moreover, remote health monitoring of patients also has a major impact on reducing healthcare costs substantially.

3.1 Smart Health Monitoring Systems Applications

Smart health monitoring systems serve various applications, such as:

1. Remote Medical Access: Smart health monitoring systems based on the Internet of Things technology realizes the ability of medical diagnosis at the convenience of the patients wirelessly rather than the patient having to go physically. The patient's data is collected securely through various sensors following which it is analyzed using algorithms based on various parameters before sending the gathered results to the patients as well as doctors for optimal recommendations.
2. Real-time monitoring: IoT-based remote health monitoring systems use non-invasive sensors to collect accurate and extensive physiological information. Cloud-based servers facilitate analysis and supervise data storage.
3. Precautionary Role: Smart health monitoring systems use data collected via the sensors and make prediction of emergencies at an early stage and can alert the patient and their family. Early anomaly detection and health trend tracking are achieved through technologies like machine learning and prediction models using IoT.

3.2 Benefits of Smart Health Monitoring System

There are almost countless ways the smart health monitoring systems can benefit and enhance the medical care field, namely:

- The limitations and inconvenience of travel and inconvenience of long distances are eliminated.
- The ability to detect diseases early.
- Improvements in medication management.
- Affordable and convenient health check-up.
- Improved treatment outcomes.

- Reduced human errors.

3.3 Smart Health Monitoring System Architecture

Figure 1 illustrate the architecture of smart health monitoring system which is configured into three layers as follow:

- First and highest layer of the architecture is the application layer. It links the hardware with the Internet of Things network.
- Network layer is the second and the middle layer of the architecture. It allows the system to connect with various devices via various communication technologies such as GSM, Bluetooth, and Wi-Fi and authorizes it to connect to the Internet to connect to web servers and the cloud.
- The third and last layer of the architecture is the physical layer which is also known as the perception layer. It contains all of the physical objects implemented in the smart health monitoring systems such as sensors.

4 Conclusion

As the population increases, the doctor-to-patient ratio increases, healthcare facilities age and operating costs rise, there is a necessity to go beyond usual cost-cutting techniques while ensuring good patient outcomes. This paper focuses on remote health monitoring systems using IoT technology. The conventional practices that are available for the realization and implementation of a smart health monitoring system are surveyed and the challenges that are part of it are identified. This paper discusses the methods of health monitoring, where doctors are not required for health check-ups. This system mainly focuses on those people who are unable to consult doctors due to various reasons. Based on this research, we concluded that the best approach for an affordable, accurate, and reliable. Health monitoring system can be achieved using the following hardware: Arduino (as it is sufficient for the interfacing of various sensors as well as extremely affordable and easy to acquire as opposed to Raspberry Pi, DS18B20 for body temperature, MAX30100 for pulse rate and SpO2, AD8232 for ECG in addition with a Wi-Fi module. The measurements collected can then be sent to a cloud server and compared with reference values and processed after which the results of the analysis will be sent to the user and the doctors via an application.

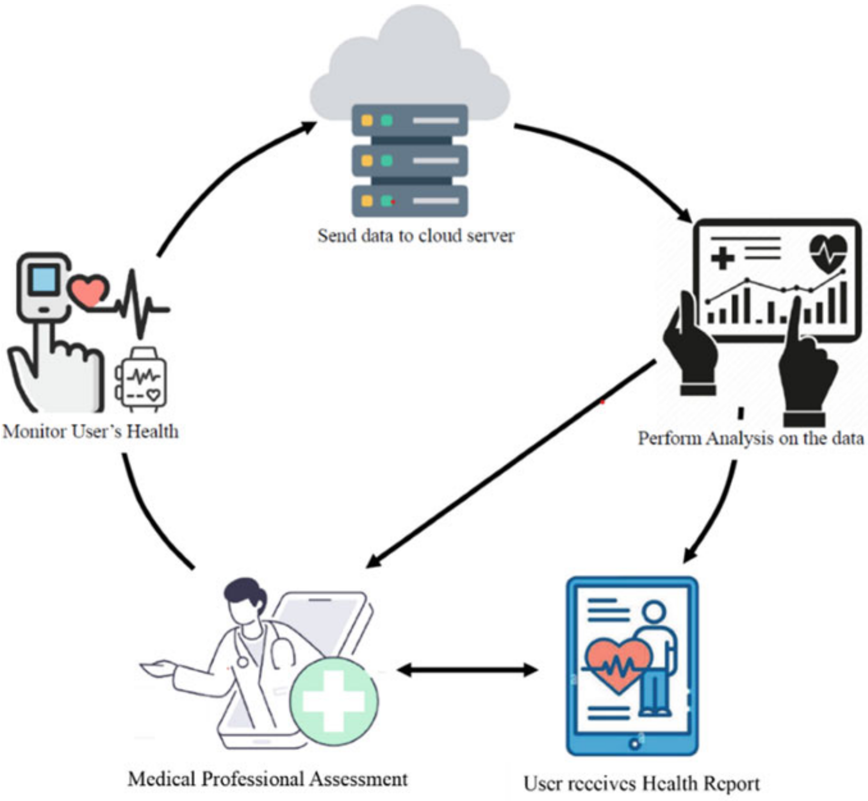


Fig. 1 Architecture of smart health monitoring system

Table 1 Comparison of methodologies used in the above explained literature survey

References	Techniques/methodology	Advantages	Disadvantages
[1]	Health monitoring system using Arduino and Bluetooth	Simple and cost efficient	Bluetooth is not versatile and has a small range
[2]	I0T monitoring system using DSI 8B20, pulse rate sensor, Arduino, and Wi-Fi module	System can be integrated with various platforms via Wi-Fi	Absence of sensors like oximeter and ECG
[4]	Voice health monitoring system via smartphone and accelerometer	Custom monitoring using just smartphone and an accelerometer	Not comfortable and poor accuracy
[5]	Electrocardiogram (ECG) monitoring system using mobile phones	Highly accurate monitoring	System has only ECG
[6]	Wireless health service system based on Android phone, various medical sensors and Bluetooth	Superior user interface using HTTP server and MySQL database	Works only on android mobile phone
[7]	Health monitoring system using Arduino Mega 2560, ADS ECG Shield, LM35, ESP8266, and LCD	Can be used as a stand-alone device via the use of LCD	ECG sensor used is expensive and ECG results cannot be displayed on LCD
[8]	Remote health monitoring system using Raspberry Pi and GSM	System can be used independently via GSM without the need of any Wi-Fi networks	High cost of the Raspberry Pi in contrast to Arduino
[9]	Health monitoring, blink detection and body movement system using Arduino UNO and GSM	Addition of new sensors to existing system	The eye blink and movement data have no correlation with health parameters
[10]	Smartphone heart rate monitoring application using camera flash	No additional hardware is required other than a smartphone	Accuracy is not on par with dedicated sensors

References

1. Trivedi S, Cheeran AN (2017) Android based health parameter monitoring. In: 2017 International conference on intelligent computing and control systems (ICICCS), IEEE, pp 1145–9. https://ieeexplore.ieee.org/document/8250646
2. Kumar SP, Samson VRR, Sai UB, Rao PLSDM, Eswar KK (2017) Smart health monitoring system of patient through IoT. In: 2017 International conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), IEEE, pp 551–6
3. Desai MR, Toravi S (2017) A smart sensor interface for smart homes and heart beat monitoring using WSN in IoT environment. In: 2017 International conference on current trends in computer, electrical, electronics and communication (CTCEEC), IEEE, pp 74–7. https://ieeexplore.ieee.org/document/8455124
4. Mehta DD, Zanartu M, Feng SW, Cheyne HA, Hillman RE (2012) Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. IEEE Trans Biomed Eng 59(11):3090–3096. https://doi.org/10.1109/TBME.2012.2207896

5. Gao H, Duan X, Guo X, Huang A, Jiao B (2013).Design and tests of a smartphones-based multi-lead ECG monitoring system. In: 2013 35th Annual international conference of the ieee engineering in medicine and biology society (EMBC), pp 2267–2270. https://doi.org/10.1109/EMBC.2013.6609989
6. Kong X, Fan B, Nie W, Ding Y (2016) Design on mobile health service system based on Android platform. In: 2016 IEEE advanced information management, communicates, electronic and automation control conference (IMCEC), pp 1683–1687. https://doi.org/10.1109/IMCEC.2016.7867504
7. Shubham B, Isha M, Saranya SS (2018) Int J Appl Eng Res ISSN 0973–4562 12(15): 11984–11989
8. Yedukondalu U, Srinivasarao U, Harish H, Hadimani H (2018) Int J Eng Manufact Sci. ISSN 2249–3115 8(1): 177–182
9. Tamilselvi V, Sribalaji S, Vigneshwaran P, Vinu P, GeethaRamani J (2020) IoT based health monitoring system. In: 2020 6th International conference on advanced computing and communication systems (ICACCS). IEEE, p 386–389
10. Banerjee S, Roy S (2016) Design of a photo plethysmography-based pulse rate detector. Int J Rec Trends Eng Res. 2:302–306
11. Gregoski MJ, Mueller M, Vertegel A, Shaporev A, Jackson BB, Frenzel RM, Sprehn SM, Treiber FA (2012) Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. Int J Telemed Appl 2012:1–7. https://doi.org/10.1155/2012/696324

Nerve Segmentation Using Multi-scale Attention Mechanism-U-Net

Ashis Das, Prabira Kumar Sethy, Raman Sahu, Santi Kumari Behera, Awantika Biswal, Manas Kumar Panigrahi, and Prafulla Kumar Bariha

Abstract In the medical field, surgical treatment is a major aspect of diagnosing any disease. To do the surgical procedure, it is very important to identify the nerve position. In most cases, ultrasound images are used to detect injury or abnormality due to certain diseases. Still, detection of the actual position of the nerve is difficult due to the presence of noise in the ultrasound images. This paper proposed a novel segmentation technique based on deep learning that is multi-scale attention mechanism-based U-Net (Ma-U-Net) segmentation. The modified U-Net segmentation technique successfully detects the nerve position of the neck using ultrasound images with an accuracy of 98.36%, a sensitivity of 98.26%, a specificity of 97.98%, and a Jaccard index of 0.96. This approach is helpful for medical practices to identify the nerve structure from ultrasound images of the neck so that unwanted side effect and pain should be avoided.

Keywords Nerve segmentation · U-Net · Multi-scale attention mechanism · Surgical procedure

A. Das (✉) · M. K. Panigrahi
Department of Computer Science and Engineering, SUIIT, Sambalpur University, Sambalpur 768019, India
e-mail: dasa7753912@gmail.com

P. K. Sethy
Department of Electronics, Sambalpur University, Sambalpur 768019, India

R. Sahu · A. Biswal
Department of Electronic Communication, SUIIT, Sambalpur University, Sambalpur 768019, India

S. K. Behera
Department of Computer Science Engineering, VSSUT, Burla 768018, India

P. K. Bariha
Department of Medicine, Veer Surendra Sai Institute of Medical Sciences and Research, Burla 768019, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_2

1 Introduction

While doing neurological practice, diseases of the peripheral nerves are common. They are important for many common musculoskeletal conditions that affect various parts of the body. Here's a brief overview of each condition are carpal tunnel syndrome (A condition that affects the hand and wrist), tendinitis (Inflammation of a tendon, which is the tissue that connects muscles to bones), rotator cuff injuries (affects the shoulder), epicondylitis (affects the elbow), trigger finger (A condition in which the finger or thumb gets stuck in a bent position and then suddenly pops straight, causing pain and discomfort), muscle strains (These occur when muscle fibers are stretched or torn due to overuse), and low back injuries (These can include a variety of conditions such as muscle strains), etc., which come under the section of musculoskeletal disorders in the field of orthopedic and rheumatology. There are many ways to find nerve lesions, but many of us know the traditional methods such as magnetic resonance imaging (MRI), computerized axial tomography (CAT) scan, and electrophysiological findings. These methods help to reflect the functional status of the nervous system, i.e., whether there is a nerve lesion or not, and it also tells us the exact position.

Some ways have come to solve this problem, such as computed tomography (CT) scan, computerized axial tomography (CAT) scan, and magnetic resonance imaging (MRI) which usually take ultrasonography visualizes as raw input and finds the place where the problem arises. On the other hand, ultrasonography has some limitations over it. First, it consumes more time, is labor-intensive, and requires a long process to implement successfully. In addition, in actual clinical settings, it is often difficult to identify peripheral nerves of the neck because the morphology and echotexture of the nerves vary among individuals and often present similar US features to those of other anatomic structures (especially tendons) [1]. Segmentation identifies and isolates specific systems or regions of interest within an image or dataset. Nerve segmentation involves identifying and separating nerve structures from surrounding tissues or structures within medical images such as MRIs or CT scans. U-Net is a convolutional neural network architecture which is developed at the Computer Science Department of the University of Freiburg [2]. Many recent reports, such as Zhao et al. [3], proposed a computer aid detection system to segment medical images. Here, the U-Net model is improved by taking advantage of the inception model and batch normalization by replacing the conventional layer. Doing so reduces the quality of parameters and enhances the accuracy without loss, which tells about how we can improve the model on U-Net segmentation. These models are commonly used for image segmentation tasks. In medical image analysis, U-Net has been widely used for tasks such as segmenting organs, tumors, and lesions from various medical images, including magnetic resonance imaging (MRI) and computed tomography (CT) scans. The U-Net architecture consists of a contracting path that captures context and a symmetric expanding path that enables precise localization. The contracting path consists of convolutional layers that extract high-level features and reduce the spatial dimensions of the input image. The expanding path uses deconvolutional layers to upsample the feature maps and recover the original spatial resolution of the input image. Nerve segmentation is a medical image-processing task that involves identifying the nerve structures in an image. Ma-U-Net, short for multi-agent U-Net, is a deep learning architecture for semantic segmentation tasks like nerve segmentation. It is an extension of the U-Net architecture, which uses convolutional neural networks to perform pixel-wise classification of the input image. The MA-U-Net architecture includes multiple agents that work in parallel, each responsible for a specific task, to improve the segmentation accuracy. It has been used for nerve segmentation in ultrasound images, MRI scans, and other medical imaging modalities. Panda et al. [4] evaluated a framework for robust and fully automated segmentation of the optic nerve, eye glows, and muscles. Here robust regularization process is employed to improve the voxel resolution. AI-Fahadami et al. [5] proposed a fully automated nerve segmentation and morphometric parameter quantification system for coronal confocal microscope images. The segmentation process involves image enhancement and noise removal using anisotropic diffusion filtering. Also, the morphological operation is applied to remove unwanted objects. Finally, edge detection is applied to detect nerve regions. Baby et al. [6] proposed automated segmentation methods of ultrasound images for nerve segmentation. Wu et al. [7] describe an innovative deep learning model that utilizes a global contextual model to automatically segment nerves from ultrasound images. Here, pixel-aware modeling is introduced to remove the speckle noise to improve the segmentation accuracy. Hafiane et al. [8] proposed a formulating a novel neural network model that incorporates spatiotemporal information to achieve precise nerve region segmentation. The methodology has 2 phases that are localization and the second is systemization. In the first phase, the CNN with special templet contingency detects the nerve zone. In the second phase, the region of interest is traced by utilizing the active contour model. Rubasinghe et al. [9] proposed deep probabilistic programming with U-Net to segment the nerve. Wang et al. [10] presented an optimized ResU-Net for nerve segmentation using ultrasound images. Initially, the application of median filtering is employed as a means to reduce the speckle noise. Incorporating Dense Atrous Convolution (DAC) and Residual Multi-kernel Pooling (RMP) modules into the ResU-Net architecture enhances the accuracy of segmentation approximation.

2 Material and Methodology

This section deals with the details of the dataset used and the adapted methodology.

2.1 About Dataset

We evaluated the efficacy of the Ma-U-Net segmentation method using a publicly accessible dataset specifically designed for nerve segmentation, which consists of 5635 raw nerve images and 5635 mask images for training. The images were obtained from ultrasound scans of the neck part of the human body. The images are resized to 576 × 480. The samples of neck ultrasound images and the mask of the nerve are presented in Figure 1.

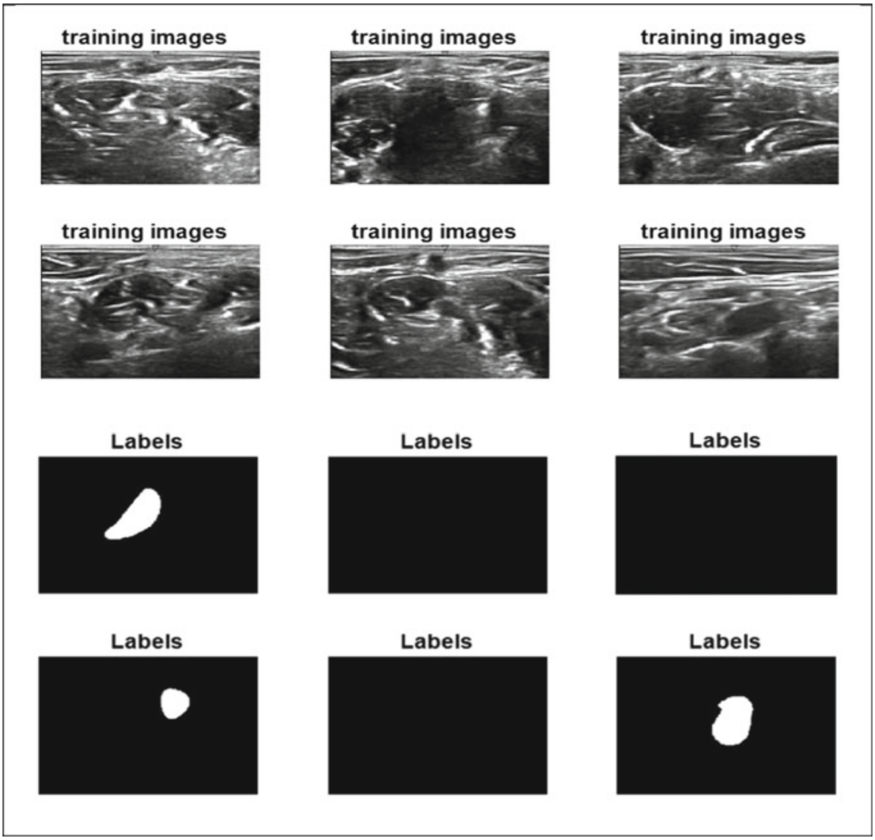


Fig. 1 Samples of train images and their ground truth

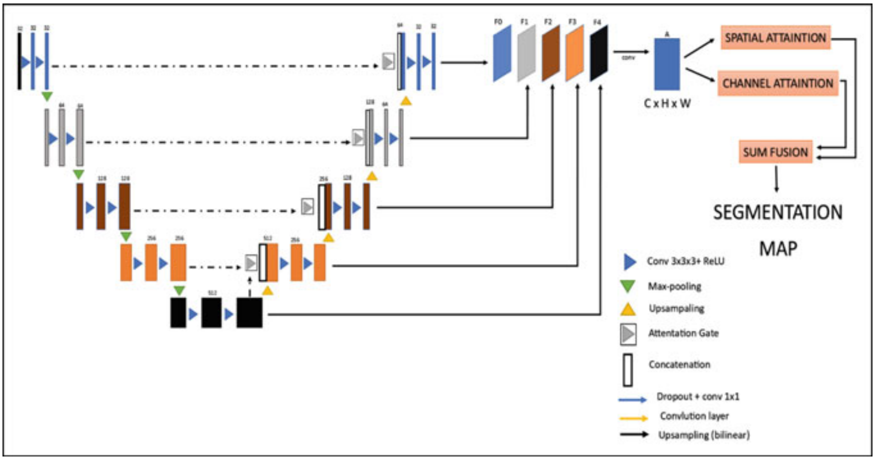


Fig. 2 Architecture of Ma-U-Net

2.2 Methodology

Medical imaging has revolutionized the diagnosis and treatment of neurological conditions, enabling non-invasive visualization of the brain, spinal cord, and peripheral nerves of the neck. Regardless, the interpretation of medical images can be complicated by the intricate and diverse nature of anatomical structures, as well as the potential for image artifacts and noise. Nerve segmentation is a critical step in analyzing medical images, as it isolates nerve structures from other structures, providing detailed information about the location, size, and extent of nerve damage or abnormality.

The Ma-U-Net segmentation method is a modified version of the popular U-Net architecture that includes a multi-scale attention mechanism (Ma) to enhance the feature representation of the network. By leveraging an attention mechanism, the network is able to dynamically allocate its attention to meaningful regions of the image, while suppressing irrelevant regions, leading to improved feature representation. The Ma-U-Net also includes residual connections to enable the network to capture fine-grained image details (Fig. 2).

3 Result and Discussion

The proposed model for nerve segmentation of the neck using ultrasound images is executed in HP Victus having windows 11, RAM 16GB with a GPU of 8GB in MATLAB 2021a. We trained and tested the Ma-U-Net segmentation method on the nerve segmentation dataset. The performance was evaluated in accuracy, sensitivity, specificity, Jaccard index, and dice coefficient. Here, 100 images are considered for testing. The dataset comprises both kinds of ultrasound images, i.e., the nerve's presence and the nerve's absence. The proposed modified U-Net segmentation technique with a multi-scale attention mechanism successfully detects the nerve position, which is better than the baseline U-Net segmentation technique. However, in the case of U-Net segmentation, some nerves are detected erroneously where there is no nerve. This problem is overcome with our proposed approach, and the measurement indices are higher. To show the genuineness of our approach, we presented some outputs with the presence of nerve and absence of nerve samples, represented in Figure 3.

4 Conclusion

This paper proposes using the Ma-U-Net segmentation method for nerve segmentation in medical images. According to the experimental outcomes, the Ma-U-Net method demonstrates superior performance in nerve segmentation when compared to other contemporary techniques. The modified U-Net segmentation technique successfully detects the nerve position of the neck using ultrasound images. Demonstrating a success rate of 98.36%, a sensitivity of 98.26%, a specificity of 97.98%, and a Jaccard similarity index of 0.96. The proposed method could be beneficial for diagnosing and treating neurological conditions and injuries. Despite this, further investigation is necessary to appraise the technique in a clinical setting.

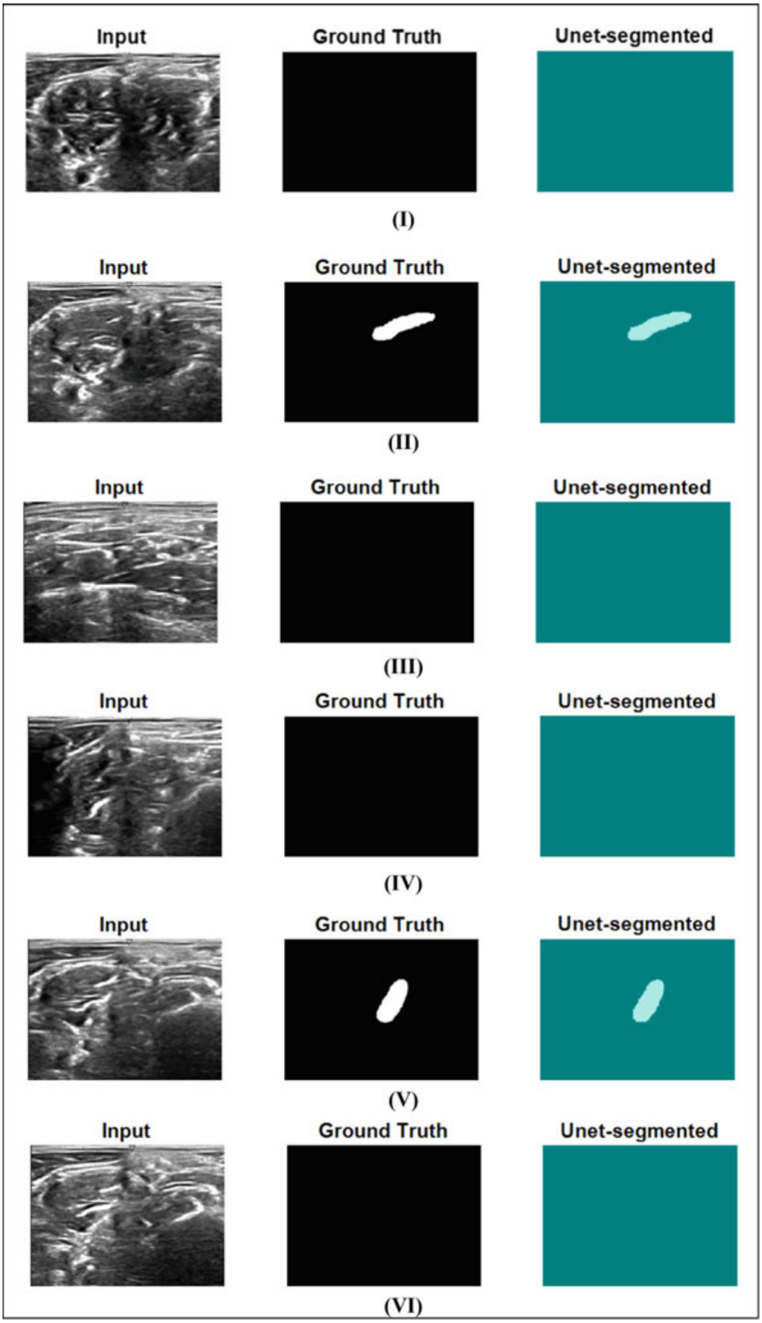


Fig. 3 Performance in the proposed model for nerve segmentation of neck using ultrasound images; have no nerve (I, III, IV, VI) and presence of nerve (II, V)

References

1. Kim BS, Yu M, Kim S, Yoon JS, Baek S (2022) Scale-attentional U-Net for the segmentation of the median nerve in ultrasound images. Ultrasonography 41(4):706–717. https://doi.org/10.14366/usg.21214. Epub 2022 Mar 15. PMID: 35754116; PMCID: PMC9532202
2. Ronneberger O, Fischer P, Brox T (2015). U-Net: convolutional networks for biomedical image segmentation. arXiv:1505.04597
3. Zhao H, Sun N (2017) Improved U-net model for nerve segmentation. In: Zhao Y, Kong X, Taubman D (eds) Image and graphics. ICIG 2017. Lecture notes in computer science, vol 10667. Springer, Cham. https://doi.org/10.1007/978-3-319-71589-6_43
4. Panda S et al (2014) Robust optic nerve segmentation on clinically acquired CT. In: *Medical Imaging 2014: Image Processing*, vol 9034. SPIE
5. Al-Fahdawi S et al (2016) A fully automatic nerve segmentation and morphometric parameter quantification system for early diagnosis of diabetic neuropathy in corneal images. *Comput Methods Progr Biomed* 135:151–166
6. Baby M, Jereesh AS (2017) Automatic nerve segmentation of ultrasound images. In: 2017 international conference of electronics, communication and aerospace technology (ICECA), vol 1. IEEE
7. Wu H et al (2021) Region-aware global context modeling for automatic nerve segmentation from ultrasound images. In: Proceedings of the AAAI conference on artificial intelligence, vol 35. no. 4
8. Hafiane A, Pierre V, Alain D (2017) Deep learning with spatiotemporal consistency for nerve segmentation in ultrasound images. *arXiv preprint* arXiv:1706.05870
9. Rubasinghe I, Meedeniya D (2019) Ultrasound nerve segmentation using deep probabilistic programming. J ICT Res Applicat 13(3):241–256
10. Wang R, Hui S, Meng Z. Ultrasound nerve segmentation of brachial plexus based on optimized ResU-Net. In: 2019 IEEE international conference on imaging systems and techniques (IST), IEEE

A Comparative Study of Register Allocation Algorithms

Gauri Amin, Gouravi Dorkhande, Mukta Ogale, Ankita Mane, and Chhaya Gosavi

Abstract Register allocation is one of the most important issues in compiler optimization because a smart allocator can improve the quality of a code by nearly 250%. It is a significant optimization that has an impact on how well-compiled code performs. Also, it is one of the most researched issues in compiler theory, and there are numerous different strategies available to resolve it. This study compares several register allocation strategies.

Keywords Basic blocks · Coalescing · Compiler · Control flow graph · Graph coloring · Interference graph · Live ranges · Live variables · Optimization · Spilling

1 Introduction

Memory hierarchies are used by computer architectures to store the data that are used by programs. In the pyramid of storage, registers are at the very because they give the CPU the quickest access to data. On a modern computer, reading and writing to registers uses a maximum of one cycle of the CPU clock. All of this speed has a price. One of the greatest questions for compiler writers is to create algorithms that maintain the most frequently used program variables in registers and relegate the least frequently used variables to memory if required because registers are so quick and scarce. Mapping variables to registers or memory is thus a difficulty with register allocation. The most crucial compiler optimization is this one. The problem then becomes how to allocate a finite number of CPU registers to an infinite number of programmer-used variables [1]. The process of choosing which values go into which registers and when during the course of a program is called as register allocation [1].

G. Amin (✉) · G. Dorkhande · M. Ogale · A. Mane · C. Gosavi
Computer Department, Cummins College of Engineering for Women, Pune, Maharashtra, India
e-mail: gauri10in@gmail.com

C. Gosavi
e-mail: chhaya.gosavi@cumminscollege.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_3

Thus, the important task is how to map program variables to memory addresses or machine registers. In the subsequent sections, we have introduced various register allocation techniques and their comparative study.

2 Register Allocation Algorithms

2.1 Graph Coloring

Graph coloring is a popular algorithm for register allocation. A graph that represents the intersection of variable live ranges in a program is known as the Interference Graph. An interference graph G includes a vertex for every variable v in a program P. In a graph, coloring is the allocation of colors to nodes, so that nodes connected by edges have distinct colors. If a graph has k colors, it is k-colorable. The number of colors is equal to the number of registers [2]. Graph nodes (temporaries) must be allotted colors (registers) in our problem. Register allocation is the assignment of values to registers during the execution of a program. This determines which values will be stored in each register at what time. A simple, organized technique for identifying registers and managing spills is graph coloring [3].

3 Chaitin-Briggs

Early works of Chaitin describe in detail graph coloring algorithm for register allocation optimization in a compiler. It provides evidence that graph coloring algorithms can be implemented in a system. The basic rule of the Chaitin's algorithm is the degree less than registers rule. When n node has less than K neighbors in a graph G with interference, it can be colored. The number K refers to the number of available colors. Node v can be expelled from G and pushed on a stack of colorable nodes in this case. This simplifies the graph. If the graph is still not empty and any node remains, then the spill code needs to be inserted [4, 5]. Phases in Chaitin-Briggs allocator are discovering live ranges in the program, building the interference graph, coalescing (eliminating copy instructions and copy propagation optimization), calculating spill cost (measure of speed of the program when mapping a variable to memory is done), simplifying (identifying the order for coloring), and coloring the graph [4].

This algorithm was further improved. In modified algorithm, new approach of late spilling decision was introduced and thus helped in avoiding unnecessary spilling. Also, it makes sure that the values will be spilled by the allocator in the least-priced manner. Approach of optimistic coloring was introduced. The limitation is that the cost of spilling is expensive. As the process is iterative, after spilling, the process would go back to the live variable analysis and further steps would be repeated [6]. Figure 1 illustrates the flow of the Chaitin-Briggs algorithm.

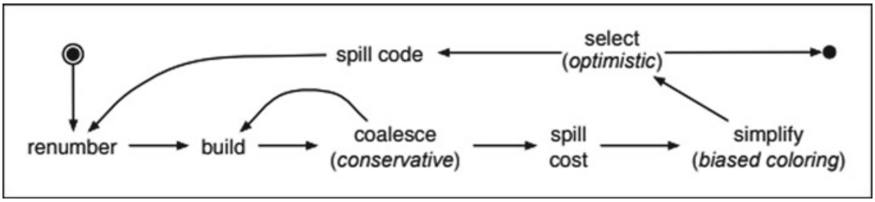


Fig. 1 Briggs graph coloring-based register allocator [4]

3.1 SSA-Based

SSA is an intermediate representation that implies each variable should only be defined once in the program. It has been independently demonstrated by several research groups that the static single assignment programs have chordal interference graphs. Graphs are chordal if every cycle with four or more vertices has an edge that connects two vertices on the cycle, that is, an edge that is not part of the cycle. Due to the local knowledge of the register pressure of SSA form programs, the register allocator can remove program live ranges till the pressure equals the availability of the registers. The variance between a number of live variables and a number of available registers is register pressure. As a result, the allocator can make spill decisions without assigning variables to registers. The following step assigns registers to variables, and the SSA features to ensure that spills will not occur again. After registers have been assigned variables, a third phase improves the initial register assignment to ensure that all variables related by copies have the same register. Figure 2 shows the flow of the SSA algorithm. Phases in SSA-based allocator-Build: Based on the liveness analysis, the interference graph is built, Spill: Live ranges should be removed if the number of available registers exceeds the register pressure, MCS: The Maximum Cardinality Search algorithm finds an optimal ordering of nodes in a graph using a greedy algorithm [7], Color: A trivial greedy coloring algorithm is used to assign registers to variables, Coalesce: Ensure that variables are linked by copy instructions given the same register by exchanging registers between variables [4].

Fig. 2 SSA-based register allocator [4]

3.2 Linear Scan

Compilers utilize various techniques to assign registers to program variables and one of those allocation techniques is linear scan. The fundamental concept of this technique is to linearly scan the program, keeping track of each variable's lifetime and assigning registers to those variables that are currently live. The algorithm scans all the live intervals in a single pass and then greedily allocates registers to variables in a program. The approach generates reasonably high-quality code and is easy to use. It is helpful in scenarios such as compilers which are just-in-time, dynamic compilation systems, interactive development environments when both code quality and compilation speed are crucial. The intermediate representation, which is utilized to identify live interval overlap, can capture the live intervals. Let us consider that N are the available registers and IntervalList is a list of live intervals, the linear scan approach distributes registers to a maximum number of intervals as it can, also it makes sure that there are no overlapping live intervals assigned to the same register. There must be at least n–N of the live intervals stored in memory if there are more variables (n) than N and the live intervals overlap at any time. The IntervalList list, which is ordered in ascending start point order, is where live intervals are kept. For quick scanning of live intervals, the algorithm hops from one start point to another. The algorithm scans IntervalList each time a new interval is found. It eliminates any expired intervals (intervals that do not overlap the new interval because their endpoint comes before the start point of the new interval) and makes room for allocation in the associated register. The size of the IntervalList list is limited to N at most. In the worst case, at the start of a new interval, IntervalList has length N and no intervals from the list expire. There is a need to spill one of the current live intervals in this case. The last-ending interval that is farthest from the present point can be spilled. Since IntervalList is sorted by increasing endpoint, it would be simple to locate this interval. There are two options for spilling the interval: the new interval or the most recent interval that ends later [8].

3.3 Tree

In Tree-based algorithms, the lifetime of each variable is represented by the color of each tree, which is ideal when variables are not spilled. The data flow between and within the trees is then connected after that. An interference graph can be created from any undirected graph, which means graph coloring is typically NP-complete. In contrast, if chords are present in the interference graph, the maximum cardinality search algorithm (MCS) can be used to color it in time proportional to its edges and nodes. Cycles with more than three nodes are chordal if they have a chord, which is an edge that does not form a cycle but connects two of the nodes in the cycle [9].

The main principle is to "treefy" the lifetimes of variables, which mandates that the interference graph be chordal and utilizes the MCS algorithm to allocate as few

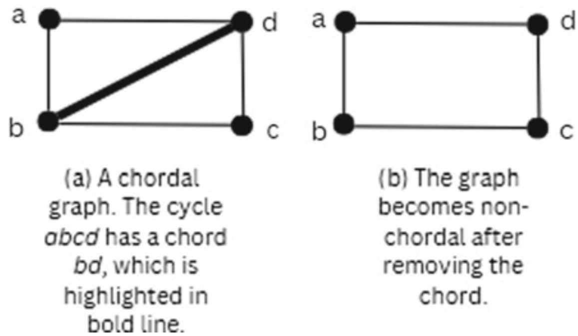


Fig. 3 Illustration of chordal graph [9]

registers per lifetime as possible. Figure 3 depicts the chordal graph. Three steps are involved: Tree formation: A set of trees is created using the lifetimes as a map. In theory, the total number of trees and the individual trees can be randomly defined. All that is required is that the lifetime of each tree covers the lifetime of its subtrees. Coloring: It is necessary that the interference graphs for the lifetimes of the trees are chordal. Thus, an MCS-based algorithm can color the lifetimes. Connecting data flow: When a variable's lifetime is assigned to different registers within or across trees, it may be necessary to reorder registers at some control flow edges to maintain dataflow [9].

3.4 Fusion-Based

Dealing with spilling and splitting at the same time is one of the difficulties for the register allocator. The purpose of this allocator is to relocate overhead operations to rarely processed portions of a program. Fusion-based global register allocation leverages the program structure to make spilling and splitting decisions. Initially, there is an interference graph for each region. Building up the interference graph in increments, i.e., for each region is the primary aspect of fusion-based register allocation. It is possible to define regions in compilers as individual blocks, superblocks, traces, or any other grouping [10]. The two concerns addressed by this framework [10]:

- Register pressure: As more values are allocated to the register, the register pressure increases.
- Sensitivity to compiler optimizations: Other compiler optimizations may not affect the allocator.

This framework follows a region-based approach. It offers a dynamic and adaptable method of allocating registers since the allocator does not command how these regions should be formed. The unique phases in this framework are [10]:

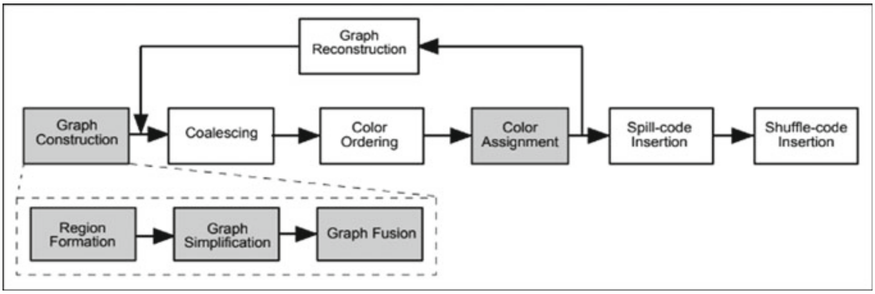


Fig. 4 Structure of fusion-based allocator [10]

- Region formation: During the region formation phase, a number of techniques can be used to form regions. Regions in compilers are defined as individual blocks, superblocks, traces, or any other grouping.
- Graph Simplification: This step decides how many live ranges should be spilled inside every region. There is no need for spill code in a region if an interference graph can be simplified. If the graph cannot be simplified, the choice of which live range to spill is postponed until the graph merging step.
- Graph fusion: By fusing interference graphs along each edge generated by the region generation phase. In this phase, the interference graph is fused along with each of the control flow edges which are set by the region formation phase. The invariant that the resulting interference graph is simplifiable which is maintained by powerful fusion operators. Fusion operators make the decision on how to split live ranges. Each region has a number of live ranges to spill and the location of the shuffle code in the merged graph is known, but no physical register commitments are made. We avoid premature coloring decisions with this invariant.
- Color assignment: This phase involves the assignment of physical register ranges and the generation of shuffle code. After the fusion of all interference graphs, the resulting interference graph can be colored. Decisions regarding spilling or splitting have been made already.

  The structure of the fusion-based allocator is shown in Fig. 4.

4 Comparison and Results

From the survey, the following comparative results are obtained as given in Table 1.

Table 1 Comparison of the algorithms

Algorithm	Time complexity	Splitting	Efficiency	Approach	Coalescing	Advantage	Limitation
Chaitin-Briggs	O (n^3)	Does not split variable	Relatively efficient for small to medium-sized programs	Global register allocation	Can be augmented with graph coalescing	Works best when code is compiled, optimized one time, then executed many times [11]. It produces good results in terms of minimizing spills, which can significantly impact program performance [5]	Its time complexity is high because of the spilling problem [11, 12]
SSA-based	O (n^2)	Does not split variable	Can simplify register allocation decisions	Transforms program into SSA form	Can be augmented with register coalescing	Simpler because spill decisions are made separately from assignment decisions [4]	Splits live ranges unnecessarily before coloring which results in needless code shuffle and performance degradation [10]
Linear scan	O (nlogn)	Does not split variable	Relatively efficient for large programs	Scans code linearly and assigns registers on-the-fly	Does not require coalescing	Faster than graph coloring, simpler to implement, as efficient as graph coloring [8]	Algorithm is unable to detect gaps in the live range of variables, control splits are concealed, has trouble spilling and splitting [8]

(continued)

Table 1 (continued)

Algorithm	Time complexity	Splitting	Efficiency	Approach	Coalescing	Advantage	Limitation
Tree	O (nlogn)	Does not split variable	Efficient for large programs with structured control flow	Constructs interference tree and assigns registers bottom-up	Can be augmented with tree coalescing	The method benefits from higher compiler throughput: Coloring can be done without the explicit construction of an interference graph. Lifetimes along each path of a tree are instead processed by coloring algorithm in chronological order [9]	Same as SSA [9]
Fusion-based	O (nlogn)	Can split variables into smaller units	Efficient for large programs with structured control flow	Fuses regions' interference graphs together and assignment of registers is done	Can be augmented with fusion-aware coalescing	In comparison to Chaitin-Briggs, this register allocation reduces data movement operations by up to 50% [10]	With biased coloring and optimistic coloring, further opportunities for improvement in code exist [10]

5 Conclusion

Numerous register allocation algorithms enable compiler optimization, as shown above. In essence, they are used in situations where compiler optimization becomes crucial. Each one is different in some or the other aspect and has an edge over the other. A comparative study was done in the above tables can help to determine and select the right algorithm for different purposes.

6 Future Scope

Register allocation algorithms can be further extended to portray how they can be used in applications like parallel and concurrent computing. Parallel computing is the simultaneous running of numerous calculations or processes. It is possible to break large problems into smaller ones, which can then be solved all at once. A form of computing known as concurrent computing involves running several computations simultaneously, during overlapping time periods, as opposed to sequentially, with one computation finishing before the next begins.

Acknowledgements We wish to express our sincere gratitude to the Principal, HOD and faculty of Cummins College of Engineering and Dr. Uday Khedkar Sir for their constant guidance and encouragement.

References

1. Kindson The Tech Pro (2015) Register Allocation Via Graph Coloration," *YouTube,* Aug. 19, 2015 [Video file]. Retrieved from https://www.youtube.com/watch?v=aMzcKTR5E1U&t=766s. Accessed on 31 Mar 2023
2. Register Allocation (via graph coloring). web.cecs.pdx.edu. Retrieved from http://web.cecs.pdx.edu/~mperkows/temp/register-allocation.pdf., last accessed 2023/03/31.
3. Bernice H (2023) Register allocation. Retrieved from http://technodocbox.com/. http://technodocbox.com/C_and_CPP/125816567-Lecture-26register-allocation.html. last accessed 2023/03/31.
4. Fernando Magno Quint˜ao Pereira F (2008) A survey on register allocation. In: A survey on register allocation, UCLA compilers group
5. Gregory JC, Mark AA, Ashok KC, John C, Martin EH, Peter WM (1981) Register allocation via coloring. Comput Lang 6:47–57
6. Preston B, Keith DC, Linda T (1994) Improvements to graph coloring register allocation. ACM Transact Programm Lang Syst VCI116(3):428–455
7. Anne Berry F, Jean B, Pinar H, Barry P (2004) Maximum cardinality search for computing minimal triangulations of graphs. Algorithmica 39(4):287–298
8. Massimiliano P, Vivek S (1999) Linear scan register allocation. ACM Transact Programm Lang Syst 21(5):895–913
9. Hongbo R (2019) Tree Register Allocation, Copyright ACM 978–1–60558–798–1/09/12

10. Guei-Yuan L, Thomas G, Ali-Reza AT (2000) Fusion-based register allocation. ACM Transact Programm Lang Syst 22(3)
11. Andrew P (2023) Algorithms for Register Allocation. courses.cs.washington.edu. Retrieved from https://courses.cs.washington.edu/courses/csep521/07wi/prj/pardoe.docx. Accessed on 31 Mar 2023
12. "Global Register Allocation via Graph Coloring." people.cs.umass.edu. Retrieved from https://people.cs.umass.edu/~moss/610slides/37.pdf. Accessed on 31 Mar 2023
13. Josef EJ (2015) Trace Register Allocation, SPLASH Companion'15, 5 ACM. 978–1–4503–3722–9/15/10 https://doi.org/10.1145/2814189.2814199, October 25–30

Brain Tumor Detection Using Machine Learning

Shishir Shekhar Singh, Rahil Ratan Kumar, and Sanjeev Kr. Punia

Abstract There are two types of brain tumors: benign and malignant. The quality of life and life expectancy of these patients are improved by early and timely disease detection and treatment plans. Utilizing deep neural networks is one of the most useful and significant techniques (DNN). In this study, brain magnetic resonance imaging (MRI) pictures were utilized to create a convolutional neural network (CNN) to identify a tumor. CNN was the first to use images. The classification accuracy of the soft max fully connected layer was 98.67%. Additionally, the decision tree (DT) classifier's accuracy is 94.24%, while the radial basis function (RBF) classifier's accuracy is 97.34%. We employ the standards of sensitivity, as well as the accuracy requirement, network performance is measured by specificity and precision. The network accuracy on the picture testing revealed that the soft max classifier has the highest accuracy in CNN, according to the data from the categorizers. This is a novel strategy for tumor detection from brain imaging that combines feature extraction methods with CNN. The method's predicted accuracy for the test data was 99.12%. The accuracy of the doctors' assistance in diagnosing the tumor and treating the patient rose as a result of the significance of the diagnosis provided by the doctor (Vanitha in JAMA 216:109, 228, 2002 [Vanitha CN, Malathy S, Dhanaraj RK, Nayyar A (2022) Optimized pol- lard route deviation and routeselection using Bayesian machine learning techniques in wireless sensor networks. Comput Netw 216:109,228]).

Keywords Brain tumor · Deep neural network · Convolutional neural network · MRI · And feature extraction

S. S. Singh (✉) · R. R. Kumar · S. Kr. Punia
School of Computer Science and Engineering), Galgotias University, Greater Noida, Uttar Pradesh, India
e-mail: Shishir221b@gmail.com

S. Kr. Punia
e-mail: sanjeev.punia@galgotiasuniversity.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_4

1 Introduction

There are two types of brain tumors: benign (noncancerous) and malignant (cancerous). The patient's condition may deteriorate as a result of the malignant tumors' fast spread to other brain areas. When the majority of the cells are worn out or damaged, they are destroyed and changed out with fresh ones. Problems may arise if damaged and old cells are not removed when producing new cells.

The growth or tumor is referred to as a mass of tissue that forms as extra cells are produced according to the kind, size, location, and shape of the brain tumor [2]. Since the tumor's size and resolution cannot be precisely measured, diagnosing brain cancers in their early stages is challenging (Fig. 1).

However, the likelihood of a patient being treated is quite high if the tumor is identified and treated early on in the process of tumor formation. As a result, early tumor detection is crucial for effective treatment. A medical examination using computer tomography or magnetic imaging is typically used to make the diagnosis. One of the most popular and crucial techniques for identifying and assessing the patient's brain is MRI imaging, which produces precise images of the brain. Medical detection systems (MDS) include MRI. The 2D and 3D pictures can be taken care of in different dimensions. The suggested method uses CNN to recognize and classify the tumor from brain pictures. The primary distinction between the neural network's primary channel. The advantage of the typical neural network is that it can automated and local feature extraction from each image. These networks are made up of neurons with weights as well as biases that can be taught. Considering the dataset's CNN results [3], it is necessary to enhance the strategy suggested. Algorithm for

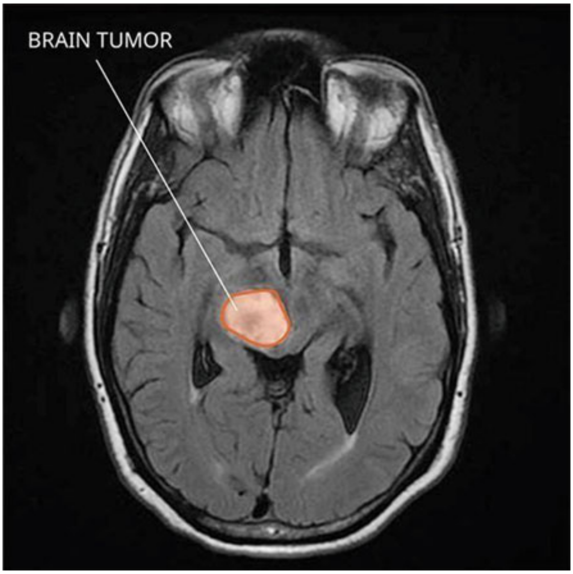


Fig. 1 A Malignant brain tumor

machine learning is employed to extract features. The method employed was the after applying the clustering method to the dataset, the photos used in relation to CNN. The outcomes demonstrated that the suggested technique has worked well. The goal of removing the property earlier Arbitrary Forest and Boosting. MSE by applying a punishment boundary, which relies upon the intricacy of the model.

Least Absolute Shrinkage and Selection Operator (LASSO) punishes the summation of the outright worth of the coefficients in Tree relapse. Dissimilar to the past ML draws near, Tree Regression is a nonparametric approach that doesn't need determination of a specific useful structure. The suggested method uses CNN to recognize and classify the tumor from brain pictures [4]. The primary distinction between the neural network's primary channel. The advantage of the typical neural network is that it can automated and local feature extraction from each image. These networks are made up of neurons with weights as well as biases that can be taught. Considering the dataset's CNN results, it is necessary to enhance the strategy suggested. Algorithm for machine learning is employed to extract features. The method employed was the after applying the clustering method to the dataset, the photos are used in relation to CNN. The outcomes demonstrated that the suggested technique has worked well.

The goal of removing the property earlier subtleties took a piece set ward on planning data as given. Consequently, business related information can be normally eliminated by the structure. We in that point propose a lodging cost guess model to help a house trader or a land master for better data thinking about the valuation with housing. Those assessments show that tie lose the faith assessment, in setting on accuracy, dependably beats substitute models in the execution of lodging cost measure. Note that contrast at the point of convergence of controlled and solo Taking in. In like manner, an illustrative model considers that endeavor for Taking in will see spam email versus the endeavor about reshaping ID. For the spam perceiving affirmation task, we consider a setting to which the student gets ready getting messages to which the etching spam/not-spam. This study includes AI calculations as an examination methodology that makes staying quote models [5]. We make a lodging cost measure model in perspective on AI assessment models for instance, controlled versus unsupervised, since taking in solidifies an association between those students and the climate, you quit presenting on that one could isolate accepting in endeavors as conveyed by those nature for that connection.

2 Information Pre-handling

MRI images are recognized and categorized using an automated process. This approach is based on the classification of each Super Pixel and the Super Pixel Technique. Each super pixel is divided into tumor and normal categories using the SVM classifier and the extremely randomized trees (ERT) classifier. Two datasets are used in this methodology: the BRATS 2012 dataset and 19 MRI FLAIR images [5]. The outcomes show that the good using the ERT classifier, this method's effectiveness. An automatic classification technique is utilized to identify a tumor created with a CNN with three little kernels [5]. The method simultaneously earned the top spot for the total, areas that are essential and improving in dice similarity, coefficient score at the 2013 BRATS Challenge (0.88, 0.83, 0.77). CNN is an AlexNet model normal tumors and MS.

CNN could provide correct information appropriately divide 98.67% of the photos into three classes [6]. In a multi-stage, Fuzzy C-Means (FCM) framework was proposed brain tumors from MRI images should be segmented. An effective and efficient technique that for classification and segmentation, CNNs are used. The suggested method that extracted characteristics using Image-Net. The outcomes acquired 84 and 97.5% accuracy for categorization in order to segment. Multi-phase MRI tumor pictures are shown in a comparison and study of grading has been conducted between fundamental neural networks and deep learning structures networks. The findings demonstrate that the network performance is based on CNN's increased sensitivity and specificity 18% less than neural networks, on average.

3 Straight Regression

For observing an association between two endless components, linear fall away from the faith is helpful. One variable is pointer or free, what's more the other variable can't try not to be variable reaction or subordinate. It searches for a relationship that is quantifiable yet not deterministic. It is said that relationship between two components is deterministic if the other can bestow one variable unequivocally.

4 Ridge Regression

Edge Regression is a multi-regression information investigating procedure experiencing multicollinearity. Most square measures are straightforward when multicollinearity happens, yet their movements are huge so they might be very a long way from the genuine worth. The edge break faith decreases standard goofs by adding a level of inclination to the apostatize measures. It is accepted that the net impact will be to give more reliable assessments [7]. Edge apostatize is changing the least squares methodology to permit to have lopsided assessors of the fall away from the faith coefficients in the lose the faith model. Edge for set a specific kind of requirements on limits.

5 Previous Analysis

The apostatize of Lasso changes from the relapse of the edge so much that purposes all around qualities instead of squares in the discipline work. prompts a discipline (or undefined block of how much all around likely gains of the assessments) which makes two or three assessments of the cutoff points end up being everything viewed as nothing [8]. The more basic the discipline applied to the assessments are also decreased to all things considered zero. Corpus for stop words is utilized to check for stops:

(a) **Case collapsing:** At this stage, all words are advanced to same protection, e.g., lower case or advanced.
(b) **Stemming:** Stemming is the most notable way to deal with conveying a root/base word's morphological assortments. Stemming programs are normally called stemming or stemming calculations.
(c) **Filling NaN Values:** Many of the factors had-1 characteristics that ought to be tended to. Taking into account what appeared to be OK, those values were done u also [8]. For example - 1 attributes for values, e.g., Back-street piled up with a string ("No Alley"), while Garage Yr Built-1 qualities were piled up with the middle to keep information away from being slanted.
(d) **Dummy Variables:** Categorical components are string which address a danger too. It is clear to make faker components which are numeric reliable for categorical factors which will help the models with Working on categorical components. Also, explicitly, we ran RF model fit on the arranging information in R inquisitively and gathered data about the fuse significance, considering everything [9]. Then, at that point, utilizing Caret's recursive part end (RFE) work, we coordinated all the part factors from the most essential to the least, and acted in reverse choice of parts by dropping 5 variables simultaneously and looking at their wise execution of cross-underwriting (CV) (Fig. 2).

To set up the dataset and make presumptions openly, I utilized Rope (least endlessly out shrinkage and affirmation supervisor) and Slope apostatizing models

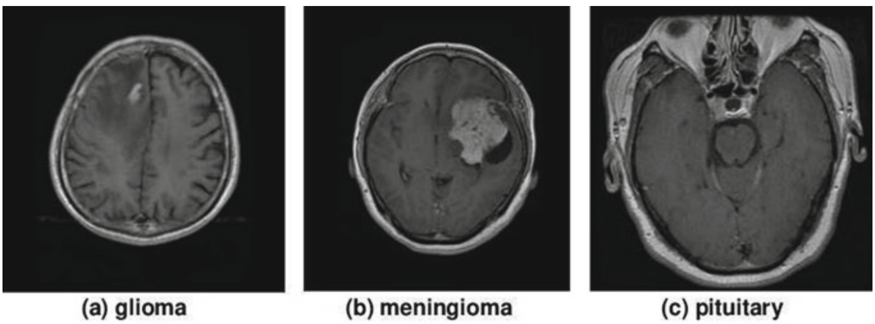


Fig. 2 Various stages of brain tumor during detection

Rope is a model of apostatize that pick and regularizes variable. The model LASSO utilizes a breaking point that rebukes such a colossal number of components for fitting. It awards variable coefficients to be lessened to 0, which basically accomplishes the model no influencing those components, in this manner diminishing dimensionality. Since there are various illustrative [8].

In like manner, HGG and LGG have various looks and models. To be amassed, it is expected to push view point toward the spot. Coordination gives an embracing sway relatively few limits for point making a difference. To ensure a strong agreement among fit and commonness, the going with advances can be taken to pick limits: (1) Setting regularization limits (lambda, alpha), (2) decreasing learning rate and those ideal limits again that the most renowned model used to predict house cost is by using various backslide examination.

The liberal expense model was normally used for gathering of critical variables along- side other backslide model, similar to assist vector with backsliding, different backslide examination and various models [10]. Meanwhile, research by picks XGBoost as the best model since it gives the least RMSE regard strangely, with various models in his audit. Such an examination is associated with the resulting investigation question of the audit.

6 Deep Learning

One of the most modern and practical varieties of machine learning is deep learning. In other terms, education is referred to as ingrained architecture. In reality, these architectures use the same old nerve networks that have evolved into DNN. Because these networks are data-driven and feature engineering is carried out automatically with no intervention from us, they are highly accurate and perform admirably in a variety of applications [3]. It is actually a deep learning of a collection of algorithms based on nerves that automatically learns features from our own input data.

7 CNN

The construction of the CNN is modeled after the biology of the cat's visual brain. The CNN is organized in layers and has a hierarchical structure. Input, output, convolutional, pooling, normalizing, and fully connected layers are also included in CNN. The number of layers used, the size and quantity of pictures, as well as the kind of activation functions used, are all varied in CNN. In CNNs, the parameters are picked by trial and error and experimentation.

8 Linear Regression

Multiple Linear Regression is a model used to choose the association between factors. To survey the association of the variables, the relationship coefficient or backslide condition can be used. Different backslide models can sort out which characteristics are the most indispensable to figure out the dependent variable [7].

Different backslide examination furthermore allows explicit expense assumptions by getting free and subordinate variable data. The power of the various backslide model ought to be noticeable when the value of the association among dependent and independent variables is assessed use different backslide exhibiting to depict improvements to a free component with a dependent variable. This model can be achieved including the house cost projection as discrete and subordinate variables like house costs, house size, property sort, number of rooms, and some more [6]. Appropriately, the house cost is set as a goal or dependence variable, while various characteristics are set as independent elements to conclude the standard variables by recognizing the association.

9 Vector Regression

Support Vector Regression is a perceptive model taking into account SVM, a mind network that by and large has three layers, a solid kind of directed learning. The model relies upon a subset of planning data. The advantages of assist vector with backsliding are that it is prepared for taking care of non-direct results, gives simply a solitary possible ideal plan, and prepared to overcome a little model learning issue. The likelihood to convey market estimates in a couple of business areas, including land, shows the way that this model can overcome the non-direct backslide issues and little model learning issues. Also, as this model didn't depend upon probability course assumptions, and the limit of arranging the information quality, either straight or non-immediate, this model was for the most part used at house cost showing. Support vector backslide offers colossal benefits in so many perspectives as this model can keep away from overfitting issues, while ensuring a single ideal course of action by restricting fundamental and trial risk [8]. In this field of study, support vector backslide is used to accumulate nuances on region, essential and locational credits.

10 Artificial Neural Network

Artificial neural network in 1958 spread the word about fake cerebrum network as ANN. Walter Pitts and Warren McCulloch appropriated a paper named "A Logical Calculus of Ideas Immanent in Nervous Activity" in the year 1943 which observes that a mind association may dishonestly be made, considering the work and development of a characteristic cerebrum association. In another assessment, as this model would as often as possible development learning, fake cerebrum networks are declared to be phony brain outlines. The fake cerebrum network model has perpetually been picked when a non-straight characteristic is involved [9]. The assessment of home estimation evaluation should in like manner include this model as a spatial naturally suspected at the expense of housing is furthermore non-direct. Thus, as in, their audit makes a fair result, subsequently it is promising to give an exact perceptive model utilizing the phony mind network computation. This structure, nevertheless has incredibly limited execution.

11 Gradient Boost

Gradient Boost was made by in 1999 and is a by and large used AI estimation taking into account its presentation, consistency and interpretability. Point helping conveys bleeding edge in various AI works out, similar to multistage request, click conjecture and situating. With the presence of huge data lately, tendency aiding faces new troubles, especially regarding the amicability among precision and execution. There are relatively few limits for point making a difference. To ensure a strong agreement among fit and commonness, the going with advances can be taken to pick limits: (1) Setting regularization limits (lambda, alpha), (2) decreasing learning rate and those ideal limits again that the most renowned model used to predict house cost is by using various backslide examination [9].

The liberal expense model was normally used for gathering of critical variables along-side other backslide model, similar to assist vector with backsliding, different backslide examination and various models. Meanwhile, research by picks XGBoost as the best model since it gives the least RMSE regard strangely, with various models in his audit. Such an examination is associated with the resulting investigation question of the audit.

12 Methodology

The 153 individuals with brain tumors and those with normal brain function who were referred to imaging centers due to headaches are represented in the data set images used in this work. 80 healthy people' brain scans were among the images gathered after the doctor had examined and diagnosed the patients. Include 1321 photos, of which 515 images are used for train data and 56 images are used for testing data. 73 patient tumors include a total of 571 pictures, of which 170 are for test data and 1151 are for train data. Of the total number of patients with brain tumor disease; their ages ranged from 8 to 66; 86 of them were women and 68 were men. 1892 photos, including 1666 images for train data and 226 images for test images, were gathered from a total of 153 patients [7]. The gathered photos were initially 512 512 pixels in size. A clustering technique is central clustering. This technique contains a duplicate procedure that attempts to get points as cluster centers—which are actually the same mean points belonging to each cluster—iteratively for a fixed number of clusters. Additionally, assign each sample of data to a cluster that has a minimum distance between it and its center. In the straightforward version of this procedure, the cluster centers are first chosen at random [10].

13 Future Scope

Future work on this study could be isolated into seven standard locales to additionally foster the result even further which ought to be conceivable by:

1. The used pre-taking care of procedures genuinely help in the assumption precision. Nevertheless, attempting various things with different blends of pre-taking care of procedures to achieve better gauge accuracy.
2. Use the available features and in case they could be joined as binning features has shown that the data got gotten to a higher level.
3. Setting up the datasets with different backslide methodologies like Elastic net backslide that joins both L1 andL2 guidelines.
4. The relationship has shown the relationship in the local data. Subsequently, attempting to overhaul the close by data is supposed to make rich with features that shift and can give a strong association relationship.
5. The results of this study have shown that ANN is leaned to overfitting. Regardless, ANN still a strong computation that has a lot of decisions that could, with the right procedures, give an unrivaled assumption accuracy. ANN has a lot of possible results that could incite a different outcome. For instance, investigating various roads in regards to the model while using blends of layers and neurons in excess of a couple of cycles to find what fits the estimation.
6. These predictions help the radiologist in making quick decisions. In the proposed work, a self-defined Artificial Neural Network (ANN) and Convolution Neural Network (CNN) are applied in detecting the presence of brain tumor and their performance is analyzed.
7. The accuracy of the doctors' assistance in diagnosing the tumor and treating the patient rose as a result of the significance of the diagnosis provided by the doctor.

When these algorithms are applied on the MRI images, the prediction of brain tumor is done very fast and a higher accuracy helps in providing the treatment to the patients.

References

1. Vanitha CN, Malathy S, Dhanaraj RK, Nayyar A (2022) Optimized pol- lard route deviation and routeselection using Bayesian machine learning techniques in wireless sensor networks. Comput Netw 216:109,228
2. Malathy S, Vanitha CN, Narayan N, Kumar R, Gokul R (2021) An Enhanced Handwritten Digit Recognition Using Convolutional Neural Network. In: Memon et al. (eds) 2021 third international conference on inventive research in computing applications (ICIRCA). 2021 Third International Conference on Inventive Re- search in Computing Applications (ICIRCA). IEEE. Lamport certificateless signcryption deep neural networks for data aggregation se- curity in wsn. Intell Automat Soft Comput 33(3):1835–1847
3. Saravanakumar Pichumani TVP, Sundararajan P, Rajesh Kumar D, Yun- young N, Seifedine K (2021) Ruzicka indexed regressive homomorphic ephemeral key benaloh cryptography for secure data aggregation in WSN. J Int Technol 22(6):1287–1297
4. Dhanaraj RK, Ramakrishnan V, Poongodi M, Krishnasamy A, Vijaya-kumar V (2021) Random forest bagging and x-means clustered antipattern detection from SQL query log for accessing secure mobile data. In: Jain DK (ed) Wireless communications and mobile computing 2021:1–9
5. Deaton A, Zaidi S (2002) Guidelines for constructing consumption aggregates for wel- fare analysis. WorldBank 1–107
6. Vinod HD (1978) A survey of ridge regression and relatedtechniques for improvements over ordinary least squares. In: Hrishikesh D (eds) Vinod source: the review of economics and statistics 60:121–131 Published by: The MIT Pr. Rev Econ Stat. (1978) 60(1):121–31
7. Varian HR (2014) Big data: new tricks for econometrics. JEcon Perspect 28(2):3–28
8. Mullainathan S, Spiess J (2017) Machine learning: an appliedeconometric approach. J Econ Perspect 31(2):87–106
9. Dhanaraj RK, Lalitha K, Anitha S, Khaitan S, Gupta P, Goyal MK (2021) Hybrid and dynamic cluster based data aggregation and routing for wire-less sensor networks. J Intell Fuzzy Syst
10. Krishnamoorthi S, Jayapaul P, Dhanaraj RK et al (2021) Design of pseudo-random number generator from turbulence padded chaotic map. Nonlinear Dyn

Restoration of Ancient Kannada Handwritten Palm Leaf Manuscripts with Modified Sauvola Method Using Integral Images

Parashuram Bannigidad and S. P. Sajjan

Abstract In the recent days, the research area of historical character recognition has got much attention toward ancient inscriptions, since they contain lots of unfolding knowledge in the field of science, literature, astronomy, medicine, etc. The materials used to write these inscriptions are paper, palm leaf, stone rocks, and temple walls, etc. These materials are now degrading in nature due to climatic conditions, ink bleeding, lack of attention, and unscientific storage. The purpose of this work is to digitization and restore medieval Kannada handwritten palm leaf scripts using a modified version of the Sauvola method and integral pictures. MSE, PSNR, Accuracy, Precision, Recall, and F1-measure are used to estimate performance, and the quality of the images is compared to those obtained manually by Epigraphists. The typical values for the mean squared error, the pseudo-stochastic noise ratio, the accuracy, the precision, the recall, and the F1-measure are 29.265, 76.828, 0.258, 0.085, 0.258, and 0.127. The better the picture quality, the higher the PSNR, the Accuracy, the Precision, and the F1-measure, and the lower the MSE and the Recall. The results are compared to those obtained by using the industry-standard methodologies of Sauvola and Niblack. The findings of the studies that compared several algorithms reveal that the proposed algorithm is superior than those that were considered as alternatives. The AMADI LONTARSET benchmark palm leaf dataset is also used to test the proposed method, demonstrating its comprehensiveness.

Keywords Palm leaf's · Image enhancement · Thresholding · Integral images · Sauvola method · Degraded document · PSNR · MSE · Document image binarization

P. Bannigidad · S. P. Sajjan (✉)
Department of Computer Science, Rani Channamma University, Belagavi 571159, India
e-mail: sajjanvsl@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_5

1 Introduction

In the recent days, the research area of handwritten character recognition has got much attention toward ancient inscriptions, since they contain lots of unfolding knowledge in the field of science, literature, astronomy, medicine, etc., the materials used to write these inscriptions are paper, palm leaf, stone rocks, and temple walls, etc. These materials are now degrading in nature due to climatic condition, ink bleeding, lack of attention, and unscientific storage. In digital image processing, the binarization of document image is often the first stage. Ancient documents are ruined, where extensive noise in the background or lots of changes exists. Hence, it is very difficult to categorize foreground and background pixels.

This study intends to apply a modified version of the Sauvola approach, which makes use of integral pictures, with the end goal of enhancing the quality of ancient Kannada handwritten palm leaf manuscripts that have already suffered from deterioration. Utilizing the integral pictures approach, local adaptive thresholding techniques pioneered by Shafaitet al. [1]. An image that contains noise that is not uniformly distributed cannot have that noise eliminated using a global threshold. In contrast to offering an adaptive solution for images with distinct background intensities, local thresholding was characterized by B. Gatos et al. as having an adaptive solution for images where the threshold varies based on the features of the local region [2]. W. Niblack came up with the evaluation known as the Niblack evaluation [3]. It uses neighbors as local thresholds. Sauvola and Pietikainen [4] proposed using two techniques to compute the threshold of each pixel based on the local mean and standard deviation (SD). Sauvola used these algorithms to calculate the threshold of each pixel. Neha Kundal and Anantdeep assessed the performance of a novel historical document restoration algorithm based on Sauvola thresholding, as well as a hybrid method merging both of these approaches with particular local filters. Their research was published in the journal computers in human behavior [5]. N. Venkata Rao and colleagues have suggested that obsolete document pictures be cleaned up using an Iterative Global Threshold that has been adjusted [6]. It has been proposed by Parashuram Bannigidad and Chandrashekar Gudada that degraded Kannada handwritten paper inscriptions can be repaired using image enhancement techniques (Hastaprati) [7]. Native threshold was proposed by B. Gatos et al. and E. Kavallieratos. It provides an adaptable solution for photos with changing background intensities, with the threshold value varying depending on the characteristics of the entire image [8, 9]. It was suggested by Parashuram Bannigidad and Chandrashekar Gudada that images of deteriorated and inconsistently lighted historical Kannada handwritten writings may be recreated [10]. The researchers Han, Yuehui, et al. described a combination method for binarizing the images of historical Tibetan document books [11]. Li, Deng, Yue Wu, and Yicong Zhou have evolved the Sauvola network, which is now known as Sauvolanet. Sauvolanet is a degraded document binarization adaptive learning system [12]. Bannigidad, Parashuram, and Chandrashekar Gudada have come up with an idea that they call "Text Line Segmentation using LBP Characteristics for Digitization and Recognition of Ancient Kannada

Handwritten Manuscripts” [13, 14]. Xiong, Wei, and others have provided a description of a framework that can improve the binarization of deteriorated historical document images [15]. Binarization of non-uniformly illuminated document images was accomplished by Yang, Xingxin, and Yi Wan by the use of the K-Means clustering technique [16].

In this paper, the efficient implementation of Sauvola method using integral images is used for binarization and the performance evaluation measures; MSE and PSNR, Accuracy, Precision, Recall, and F1-measure are used as quality measurements for degraded manuscripts. The results are compared with other standard methods in the literature; Sauvola and Niblack using our own dataset. Further, the proposed method is also used on AMADI LONTARSET, a standard palm leaf dataset, to evaluate and measure the performance of our algorithm with demonstrate that the proposed technique is exhaustiveness of the proposed technique.

2 Proposed Method

2.1 Implementation of Modified Sauvola Method Using Integral Images

In this paper, a new improved method has been developed to binarize degraded Kannada Handwritten Palm leaf manuscripts. Using adaptive thresholding methods for historical Kannada handwritten palm leaf manuscripts, Handwritten palm leaf manuscripts are converted into grayscale images apply to adaptive thresholding technique. The threshold is then considered using the Sauvola approach, and it takes into account the mean as well as the standard deviation of the window pixels surrounding it. In contrast to averaging the mean and standard deviation of all of the pixels contained inside a window, integral images can be utilized to determine mean and standard deviation with as few as four separate mathematical procedures. Without sacrificing the high quality of the original Sauvola method, this drastically reduces the runtime of the technique and makes it much more efficient [1].

2.1.1 Sauvola’s Method

As a local adaptive thresholding method, the Sauvola procedure takes grayscale images and converts them into binary. The average intensity and standard deviation (SD) of a local window surrounding the pixel are used to calculate a local threshold for that pixel. The method’s goal is to get rid of the noise in an image without losing any of the original image’s details or texture.

The Sauvola method uses the following formula to compute the local threshold for each pixel:

$$T(x,\ y) = m(x,\ y) * \left[1 + k * \left(\frac{\text{stddev}(x, y)}{R}\right) - 1\right] \tag{1}$$

where $T(x, y)$ is the pixel's location-specific threshold, mean(x, y) is the pixel's local window's mean intensity value, stddev(x, y) is the pixel's local window's standard deviation, k is a user-defined parameter that regulates the threshold sensitivity, and R is the pixel values' dynamic range.

Once the local threshold is calculated for each pixel, the image is binarized by setting the intensity of each pixel to 0 (black) if it is below the local threshold or to 255 (white) if it is above the local threshold.

The Sauvola method is a simple and effective technique for binarizing grayscale images. However, it has some limitations, such as being sensitive to the size and shape of the local window and not being able to handle images with varying illumination or contrast. To address these limitations, modifications of the Sauvola method have been proposed, such as using integral images to speed up the computation and adapting the method to handle images with non-uniform illumination.

2.1.2 Integral Images for Computing Local Means and Variances

Using the integral image allows for a fast and accurate calculation of the total sum of values (the values of each pixel) in an image that has been provided. The intensity of a pixel in an integral picture of an input image, denoted by the value $g(i, j)$, is defined as the sum of the intensities of all the pixels in the original image that are located above and to the left of the place in question. As a consequence of this, the magnitude of the intensity at the position (x, y) can be expressed as follows:

$$I(x,\ y) = \sum_{i=0}^{x}\sum_{j=0}^{y} g(i,\ j) \tag{2}$$

In OpenCV, the integral image of any grayscale image may be easily generated in a single pass using only one line of code. This is possible because integral images are linear. After we have obtained the integral image, we can compute the local mean $m(x, y)$ for any window size by doing two additions and two subtractions instead of totaling all of the pixel values that are contained within that window:

$$\begin{aligned} m(x, y) = (&I(x + w/2,\ y + w/2) + I(x - w/2,\ y - w/2) \\ &-I(x + w/2,\ y - w/2) + I(x - w/2,\ y + w/2))/w^2 \end{aligned} \tag{3}$$

Similarly, we can compute the variance as:

$$S^2(x,\ y) = \frac{1}{w^2}\sum_{1=x-\frac{w}{2}}^{x+\frac{w}{2}}\ \sum_{j=y-\frac{w}{2}}^{y+w/2} g^2(i,\ j) - m^2(x,\ y) \tag{4}$$

We can now calculate the standard deviation by calculating the square root of the variance. We can easily substitute the mean and standard deviation values into the original Sauvola formula to obtain the threshold effectively and independently of the local window size. This strategy reduces the computational complexity from $O(W^2 \times N^2)$ to $O(N^2)$, resulting in a substantially faster overall procedure. It is recommended that 32-bit integers be used for the implementation since the values of the squared integral image can get fairly large, which can lead to overflow problems. This is an essential piece of guidance for the implementation process.

The updated Sauvola approach employing the integral images algorithm is applied first to degraded palm leaf images. If the damaged palm leaf image still contains noise, it is detected and reprocessed independently. The following steps comprise the proposed algorithm:

Algorithm 1. Applying modified Sauvola method using integral images

The modified Sauvola method using integral images can be implemented using the following algorithm:

Step 1: Compute the integral image I of the input grayscale image

Step 2: For each pixel in the input image, compute the local mean and variance using the integral image:

Step 3: Using this formula, we can determine the threshold value for each individual pixel:

$$\text{threshold} = \text{mean} * (1 + k * ((\text{var}/R) * *0.5/128 - 1))$$

where k is a user-defined parameter (typically set to 0.5–0.7), and R is the dynamic range of the image (the difference between the maximum and minimum pixel intensities)

Step 4: Binarize the input image by setting each pixel to black or white based on whether its intensity is below or above the threshold, respectively

Step 5: Output the binarized image

Step 6. In order to determine whether or not the approach is effective, it is recommended that a binarized image be created and various geometric feature values, such as Precision, Recall, F-Measure, MSE, and PSNR, be computed

End Procedure

Overall, the modified Sauvola method using integral images is an efficient and effective technique for document image binarization. It can produce high-quality binarized images with a low computational cost. However, the optimal values of the various parameters (window size, k, etc.) may change based on the particular qualities of the image that was used as the source. As a result, it is possible that in order to get the best results, it will be necessary to experiment with alternative values for the parameters.

3 Experimental Results and Discussion

The historical Kannada handwritten palm leaf manuscripts are collected from e-Sahithya Documentation Forum, Bangalore. The implementation is done on a windows system containing Intel i5 processor 2.30Ghz speed, 8 GB RAM, 4 GB

GPU (NVIDIA GeForce GTX 1050 Ti), on the system using Anaconda3 Distribution, Spider, Python 2.7. The camera captured mediaval Kannada handwritten palm leaf manuscripts are shown in Fig. 1. Figure 1a shows a typical noisy document with non-uniformly distributed noise. The image's background is golden brown in color. Computed histogram of the original noisy document is shown in Fig. 1b. The anticipated result was then compared to the Sauvola and Niblack which are shown in Fig. 1c, d. All the methods produce some noise in the manuscript while maintaining image clarity, the resulting proposed method of modified Sauvola method using integral images produced clear image as same as ground truth image, shown in Fig. 1e.

Standard palm leaf datasets are used to test and implement the suggested technique, such as the AMADI LONTARSET dataset and the results are shown in Fig. 2. A

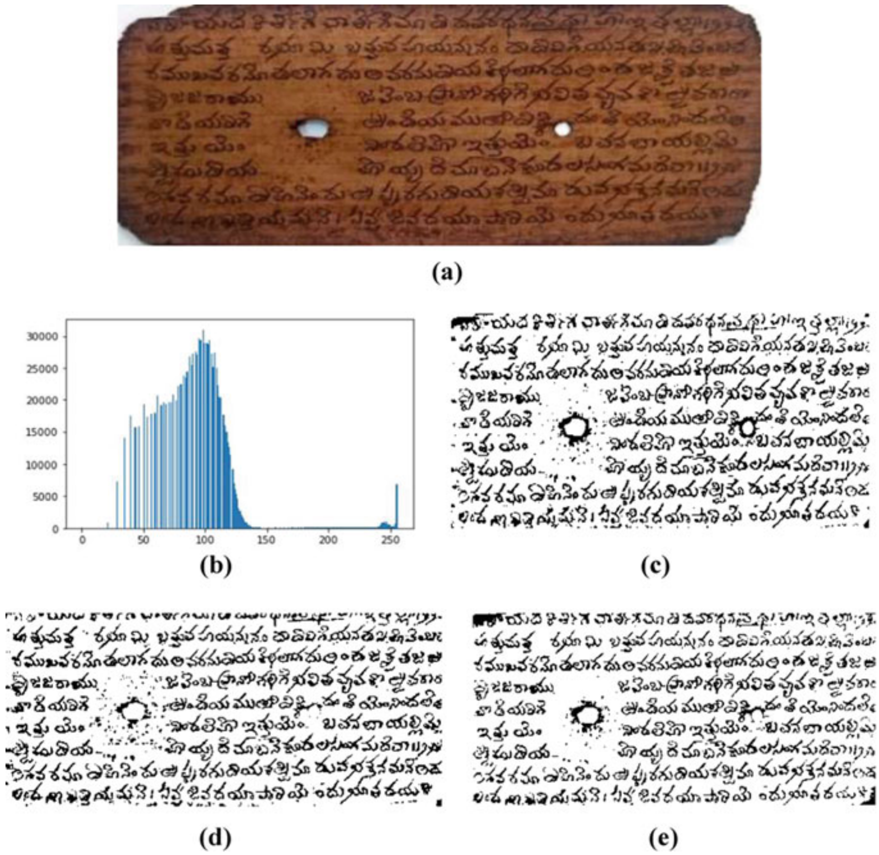


Fig. 1 **a** Sample images of original historical Kannada handwritten palm leaf. **b** Histogram of original image. **c** Sauvola threshold applied on gray images. **d** Niblack threshold applied on gray images. **e** Improved method of Sauvola method using integral images

(a) (b)

Fig. 2 **a** Sample palm leaf handwritten images of LONTARSET dataset. **b** Proposed method of Sauvola method using integral images

typical noisy document AMADI LONTARSET dataset image in Fig. 2a the resultant proposed image is presented in Fig. 2b.

Table 1 displays the results of applying various metrics to each image, including precision, recall, F-Measure, mean squared error, and peak signal-to-noise ratio. The effectiveness of the proposed method is assessed by languages experts and epigraphers. Precision is 0.838%, recall is 40.2%, F-Measure is 0.1108, mean squared error is 5.442, and peak signal-to-noise ratio is 5.442. The literature implies that better picture quality is associated with more precision, F-Measure, PSNR, and less recall, MSE.

Table 1 Proposed results are compared with other standard methods

Binarization methods	Performance evaluation approaches	Performance evaluation of degraded Kannada Palm leaf images					
		Image 1	Image 2	Image 3	Image 4	Image 5	Avg
Souvola method	PSNR	30.01	28.61	29.172	29.908	28.581	29.256
	MSE	64.86	89.533	78.671	66.405	90.142	77.922
	Accuracy	0.226	0.212	0.274	0.358	0.221	0.258
	Precision	0.062	0.056	0.091	0.151	0.061	0.084
	Recall	0.226	0.212	0.274	0.358	0.221	0.258
	F-measure	0.098	0.089	0.137	0.213	0.096	0.127
Niblack method	PSNR	30.013	28.609	29.17	29.899	28.579	29.254
	MSE	64.817	89.554	78.701	66.539	90.183	77.959
	Accuracy	0.197	0.185	0.228	0.279	0.193	0.216
	Precision	0.055	0.049	0.077	0.125	0.053	0.072
	Recall	0.226	0.212	0.274	0.358	0.221	0.258
	F-measure	0.086	0.077	0.116	0.173	0.083	0.107
Proposed method	PSNR	30.015	28.63	29.173	29.916	28.591	29.265
	MSE	64.792	84.475	78.661	66.282	89.931	76.828

(continued)

Table 1 (continued)

Binarization methods	Performance evaluation approaches	Performance evaluation of degraded Kannada Palm leaf images					
		Image 1	Image 2	Image 3	Image 4	Image 5	Avg
	Accuracy	0.226	0.212	0.274	0.358	0.221	0.258
	Precision	0.06	0.057	0.091	0.153	0.062	0.085
	Recall	0.226	0.212	0.274	0.358	0.221	0.258
	F-measure	**0.096**	**0.09**	**0.137**	**0.214**	**0.097**	**0.127**

Bold indicates the performance evaluation approach F1 Measure gives better results

4 Conclusion

The digitization and restoration of Kannada handwritten palm leaf manuscripts have a significant role to understand the ancient history and cultural customs and also helps in understanding and identifying the age of palm leaf. In this research, a degraded image of a Kannada handwritten palm leaf is used to test out the modified Sauvola method using integral images, and metrics like mean squared error (MSE), root mean squared error (RMSE), precision, recall, and F1-measure were extracted to assess the effectiveness of the methods developed. The typical values for the mean squared error, the pseudo-stochastic noise ratio, the accuracy, the precision, the recall, and the F1-measure are 29.265, 76.828, 0.258, 0.085, 0.258, and 0.127. The better the picture quality, the higher the PSNR, the accuracy, the precision, and the F1-measure, and the lower the MSE and the recall. When compared to other conventional procedures such as Sauvola and Niblack, the achieved results are encouraging. The proposed method, which makes use of integral pictures processed with the Sauvola method, is able to successfully eliminate the effects of uneven lighting and improve accuracy. The suggested approach has also been successfully applied to and tested on the AMADI LONTARSET dataset, another popular palm leaf benchmark. Manuscripts written in Kannada and preserved on palm leaf may one day be properly cataloged and acknowledged.

References

1. Shafait F, Keysers D, Breuel TM (2008) Efficient implementation of local adaptive thresholding techniques using integral images. In: Document recognition and retrieval XV. vol 6815. SPIE
2. Gatos B, Pratikakis I, Perantoni SJ (2006) Adaptive degraded document image binarization, pattern recognition, vol 39, pp 317–327
3. Niblack W, An introduction to digital image processing, 1st edn. Prentice Hall, pp 115–116
4. Sauvola J, Pietikainen M (2000) Adaptive document image binarization. Pattern Recogn 33:225–236
5. Kundal N, Anantdeep (2015) Performance evaluation of novel historical documents restoration algorithm. IJCSET 5(7):278–282

6. Rao NV, Rao AS, Balaji S, Reddy L (2011) Cleaning of ancient document images using modified iterative global threshold. Int J Comput Sci 6(8):409–417
7. Bannigidad P, Gudada C (2017) Restoration of degraded Kannada handwritten paper inscriptions (Hastaprati) using image enhancement techniques. In: Proceeding of 2017 international conference on computer communication and informatics (ICCCI), pp 1–6
8. Gatos B, Pratikakis I, Perantoni SJ (2006) Adaptive degraded document image binarization. Pattern Recogn 39:317–327
9. Kavallieratou E, Binarization A (2005) Algorithm Specialized On Document Images And Photos. In: Proceeding of 8th international conference on document analysis and recognition, pp 463–467
10. Bannigidad P, Gudada C (2018) Restoration of degraded non-uniformally illuminated historical Kannada handwritten document images. Int J Comput Eng Appl 12:1–13
11. Han Y, Wang W, Liu H, Wang Y (2019) A Combined approach for the binarization of historical Tibetan document images. Int J Pattern Recogn Artif Intell 33(14):1–21
12. Li D, Wu Y, Zhou Y (2021) Sauvolanet: learning adaptive sauvola network for degraded document binarization. In: Proceeding of international conference on document analysis and recognition, vol 12824. Springer, Cham, pp 1–36
13. Bannigidad P, Gudada C (2019) Digitization and recognition of historical Kannada handwritten manuscripts using text line segmentation with LBP features. 6(4):657–664
14. Bannigidad P, Gudada C (2019) Age-type identification and recognition of historical Kannada handwritten document images using HOG feature descriptors. In: Proceeding of computing, communication and signal processing, vol 810. Springer, Singapore, pp 1001–1010
15. Xiong W, Zhou L, Yue L, Li L, Wang S (2021) An enhanced binarization framework for degraded historical document images. EURASIP J Image Video Process 1:1–24
16. Yang X, Wan Y (2021) Non-uniform illumination document image binarization using K-Means clustering algorithm. In: 2021 IEEE 9th international conference on information, communication and networks (ICICN), pp 506–510

Conversion of Gujarati Alphabet to Gujarati Sign Language Using Synthetic Animation

Nasrin Aasofwala, Shanti Verma, and Kalyani Patel

Abstract The deaf people are using sign language for communication. Mostly information is available in written form or spoken language. As the deaf people cannot understand easily, avatars are useful to generate sign from text. Synthetic animation we can achieved through a notation system like HamNoSys, Stokoe, SignWriting, etc. The Hamburg notation system for sign languages (HamNoSys) is an alphabetical system for the primarily phonetic description of signs. However, there are no research on creation HamNoySys Notation for Gujarati Sign Language Alphabets. In this paper, avatar-based Guajarati dictionary has proposed for speech to text to Gujarati Sign Language System. This paper represents a creation of HamNoSys of Guajarati alphabets and convert them into HamNosys codes that codes will help to generate SiGML–XML code for performing the sign by Avatar.

Keywords Synthetic animation · Gujarati sign language · HamNoSys · Avatar · SiGML

1 Introduction

1.1 HamNoSys Symbols

The Hamburg notation system for sign languages (HamNoSys) is an alphabetical system for the primarily phonetic description of signs [1]. HamNoSys notations include Handshape, Orientation, Location and Actions (Fig. 1).

Handshape: The symbols for basic forms of finger and thumb position and bending make up the description of a handshape. In addition, there are variations. Based on

N. Aasofwala (✉) · S. Verma
LJ University, Ahmedabad, India
e-mail: nasrin.mscit16@gmail.com

K. Patel
Gujarat University, Ahmedabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_6

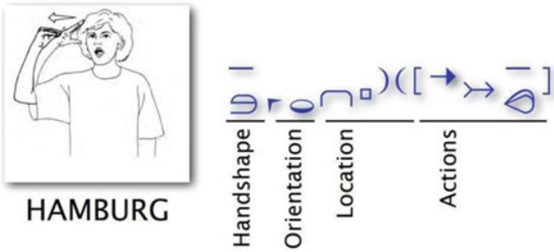


Fig. 1 HamNoSys notation

this broad–with, a description with regard to if there are fingers involved, the shape of finger by finger is possible to specify. Whenever it is essential, forms in the middle can be characterized as well. Following picture describes the sample open handshape symbol.

Hand Orientation: HamNoSys determines the angle of the hand by 2 parts: (a) Extended finger direction (b) Palm orientation (Fig. 2).

(a) Extended Finger Direction

Extended finger direction has 3 parts: [1. Signer's View, 2. Bird's View, 3. View from the right] (Fig. 3).

(a) Palm Orientation

8 symbols are required for palm orientation. Each symbol meaning is defined corresponding to the extended finger direction (Fig. 4).

Locations: Location parameters, like hand orientation, are divided into two parts: The first determines the x and y coordinates within the frontal plane, while the second determines the z co-ordinate. Following picture describes head location (Fig. 5).

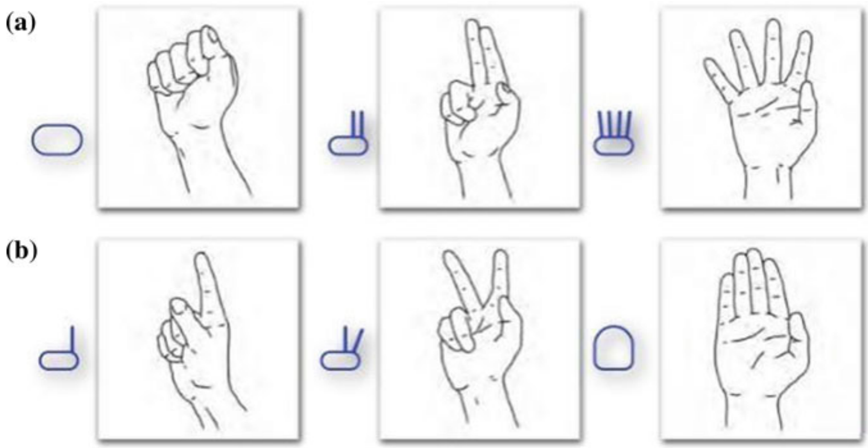


Fig. 2 Open handshape notation

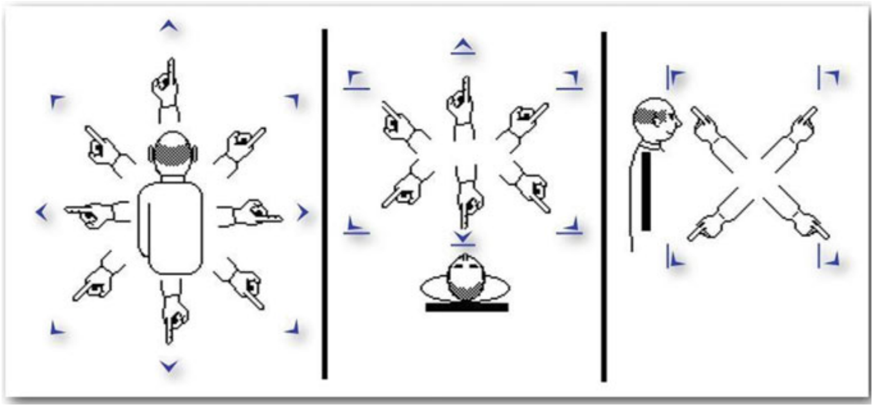


Fig. 3 Extended finger direction

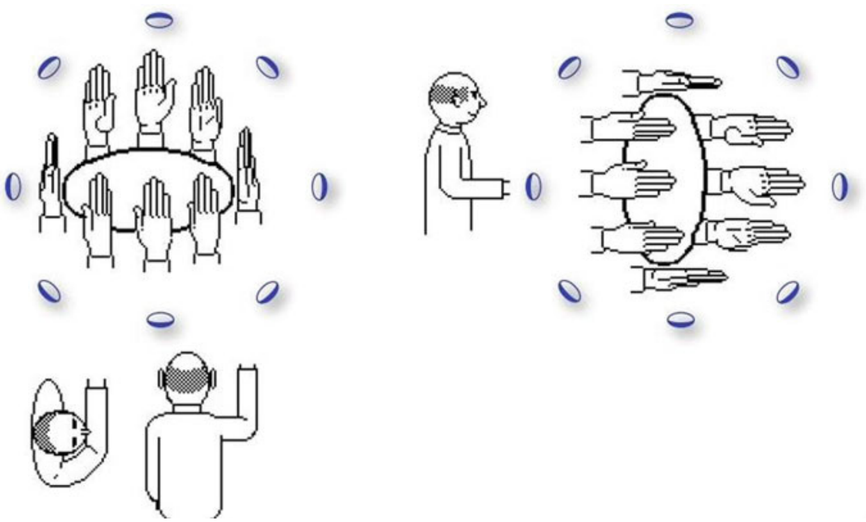


Fig. 4 Palm orientation

Actions: Path movements (movements that change the location of the hand) and in place movements of the hands and non-manual movements make up actions. The combinations can be done in any order or at the same time. Straight, curved and zigzag lines, circles and similar patterns are the building blocks of direction movement in HamNoSys (Fig. 6).

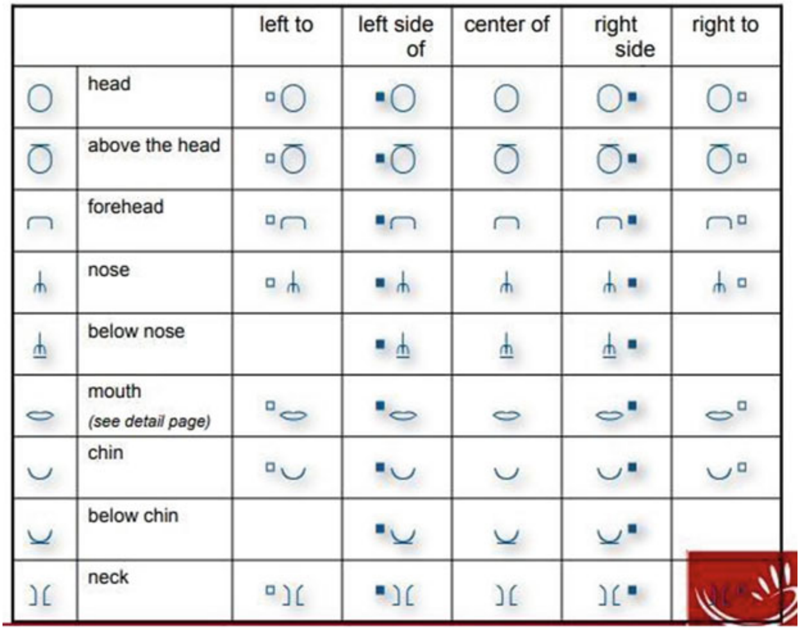


Fig. 5 Head location

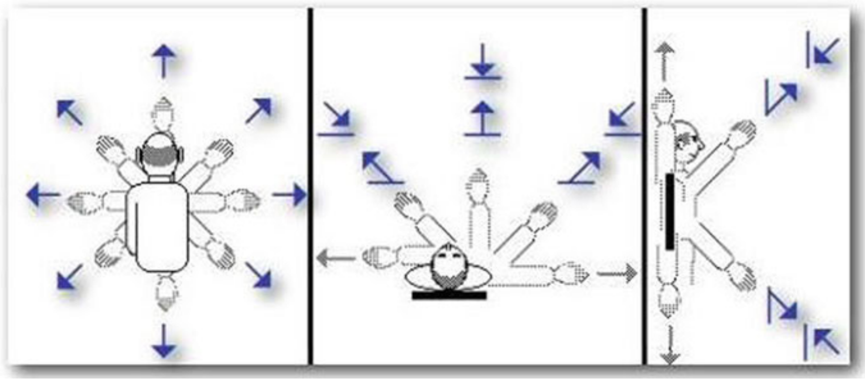


Fig. 6 Straight movements

2 Literature Review

See (Table 1).

3 Architecture

See (Fig. 7).

The above architecture is explained in the following steps.

Step 1: Input Speech by Microphone and recognize the speech by speech_recognition Python API.

```
r=sr.Recognizer()
with sr.Microphone() as source:
  print("Talk")
  audio_text=r.listen(source,timeout=20)
```

Step 2: Convert Speech to Gujarati Text with google translate Python API.

```
temp=r.recognize_google(audio_text)
print("Text:"+temp)
translator=Translator()
translation=translator.translate(temp,src='en',dest='gu')
print(translation.text)
```

Step 3: Gujarati Sign Language for Gujarati Text (Fig. 8).

Step 4: Creating a HamNoSys Notation of Gujarati Sign.

Table 1 Related work of different sign language

Paper No.	Language	Sign Language	Description
[1]	Malayalam	Indian sign language	This paper includes to accept a Malayalam text and output is a ISL. This will first convert in HamNoSys and then sign animation. It is used as a sign language tutoring system
[2, 3]	English	Indian sign language	This paper includes simple words conversion from English to corresponding HamNoSys. Then converting in SiGML for animation. It includes one and two handed symbols
[4]	German	Swiss German Sign Language	It convert the written German train information into DSGS gloss. Signed the information from the gloss transcription in front of camera. Add the non-manual feature details and played the avatar according to that
[5]	Arabic	Arabic sign language	It is a machine translated system which convert text to sign language in Arabic. It performs a morphology and syntax analysis of the text convert in a video sequence frames that played by a 3D avatar
[6]	English	British sign language	It is development of grammar for synthetic animation. It takes input as a English text -> CMU parser-> DRS creation -> HPSG creation- > SL generation -> HamNoSys -> animation
[7]	Hindi	Indian sign language	American and British sign language HamNoSys notation are available in e-sign editor but ISL words are not available. It designed an automation software which generate ISL word from the HamNoSys Notation Further it will convert in SiGML which will play avatar on JA SiGML URL APP
[8]	English	British Sign Language	This translation system of ViSiCAST text to signed animation will take input as an English text and convert in three national sign languages, Dutch Sign Language, German Sign Language and British Sign Language
[10]	English or Hindi	Indian sign language	ISL HamNoSys dictionary is developed using for Hindi and English Language. Dictionary can be used as a database of signs. HamNoSys converted into SiGML files manually. System takes sentences as a input and generate a 3D avatar
[12]	English or Urdu	Pakistan sign language	The Urdu or English text convert in equivalent HamNoSys, apply grammar rules of Pakistan Sign Language and then it convert in sign animations using an avatar. It also added a PSL -> NL translate from video as input. The video is passed to video to HamNoSys generator which generated the corresponding HamNoSys

(continued)

For development of HamNoSys, notation tools can be used as (Figs. 9 and 10):

Table 1 (continued)

Paper No.	Language	Sign Language	Description
[13]	English	Indian sign language	This paper convert English text to SiGML code file. First, it converts text to HamNoSys then further xml file which is in SiGML extension It is a mapping between English text and HamNoSys Notation where the words are used of daily life
[14]	English	South African sign language	Natural language in South Africa such as English, Afrikaans or Xhosa. The main objective to develop to SASL system is that most systems are developed for American and British sign language and second is all other systems are expensive. They are using the XTAG system for parsing an English text. Further the English text syntax to SASL syntax from syntax tree and to the graphics module. Graphical module can draw the avatar which perform a sign
[15]	English speech	Indian sign language	It will convert English speech to ISL. Objective of the software is to decrease the gap between normal people and deaf people. It uses the NLP, Google cloud speech API and sign language database. It provides on average 77% accuracy. It will take 0.85 processing time
[16]	English	American sign language	Translation from English sentence to intermediate representation as gloss notation. Synchronous Tree Adjoining Grammar (STAG) is used for mapping between English to ASL
[17]	Englishor Chinese	Chinese or American sign language	Speech input is translated into text using Google's speech recognition library. He then added annotations, segmentation and grammatical conversion using Stanford University's Natural Language Processing (CoreNLP) tool. Gesture of each word from the database and associate them into a sentence. 3D virtual man shows the sign language of this sentence
[18]	English	Indian sign language	The goal of the paper is to develop an ISL dictionary using synthetic animation. It will take a commonly used word to convert into HamNoSys. Further, it will converted into SiGML for synthetic animation. They have implemented on total 1478 words which are categorized into adjective, verb, noun, etc

(1) https://www.sign-lang.uni-hamburg.de/hamnosys/input/

(2) eSign Editor Software

Step 5: Convert HamNoSys symbol to Hexadecimal codes (Fig. 11).

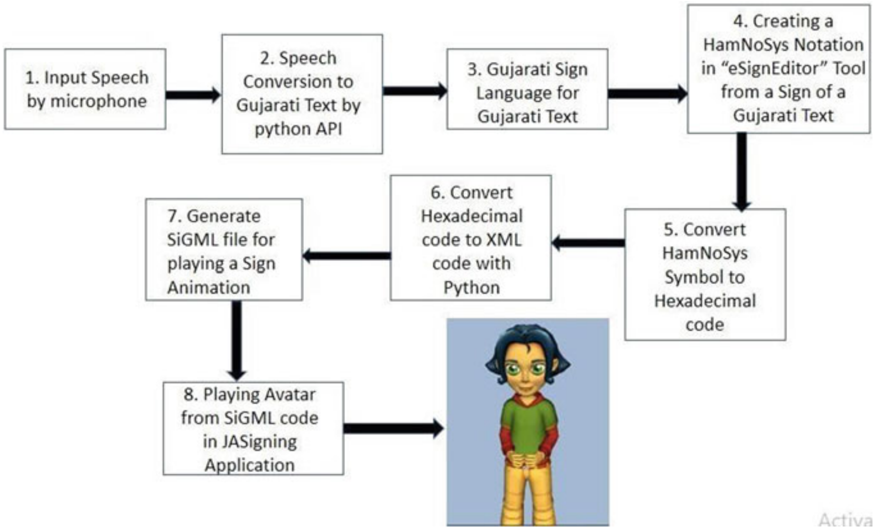


Fig. 7 Architecture of synthetic animation of Gujarati sign language alphabets

Fig. 8 "Ka" ક Gujarati alphabet

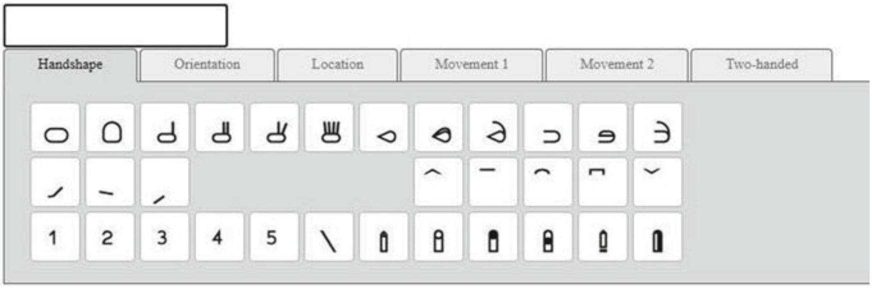


Fig. 9 HamNoSys input palette

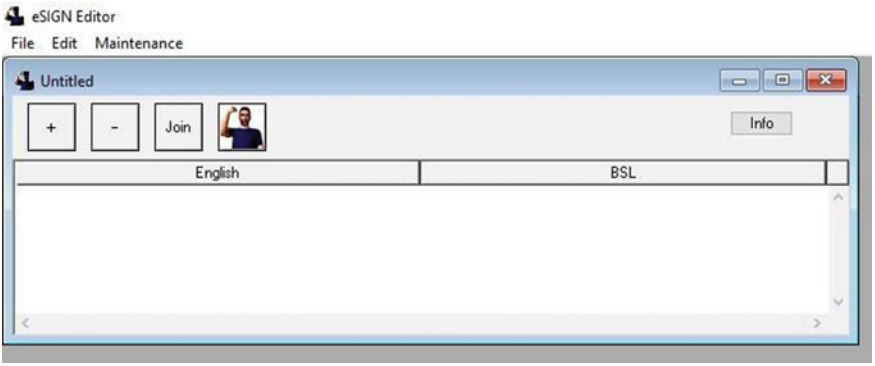


Fig. 10 eSign editor software

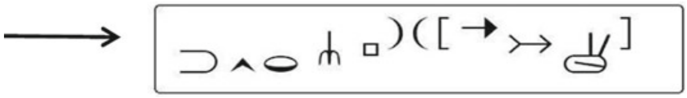
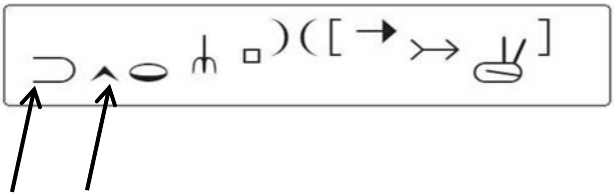


Fig. 11 HamNoSys for "Ka" alphabet



E009 E020 E03C E045 E058 E0D0 E0E2 E082 E0AA E004 E00D E0E3

Each hamNoSys symbol has Hexadecimal code which will help to generate a SiGML code. The following table shows the basic handshape symbol and code. All hamNoSys symbol code is presented in hamNoSys package paper [21].

Step 6: Generate the SiGML xml format from hexadecimal code argument and Python code.

```
<?xml version="1.0" encoding="UTF-8"?>
<sigml>
    <hns_sign gloss="Ka">
       <hamnosys_nonmanual/>
       <hamnosys_manual>
           <hamcee12/>
           <hamextfingeru/>
           <hampalmd/>
           <hamnose/>
           <hamlrbeside/>
           <hamclose/>
           <hamparbegin/>
           <hammover/>
           <hamreplace/>
           <hamfinger23spread/>
           <hamthumbacrossmod/>
           <hamparend/>
       </hamnosys_manual>
    </hns_sign>
</sigml>
```

Step 7: Play an avatar on https://vhg.cmp.uea.ac.uk/tech/jas/vhg2022/SiGML-Player-Orig-gui.html URL with the SiGML code (Fig. 12).

The above sign is generated for “Ka” Gujarati Symbol. The code is pasted in the SiGML text box and click the “Play SiGML Text” Button. If SiGML code is properly it will play avatar otherwise it will shows Invalid SiGML (Table 2).

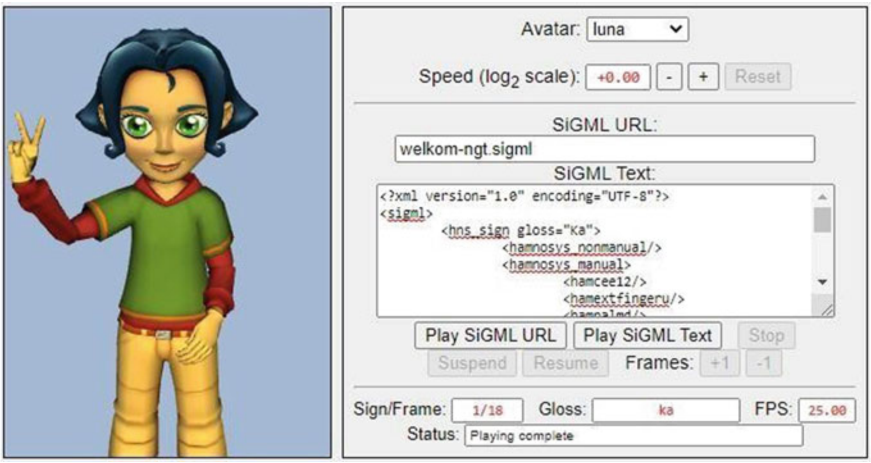
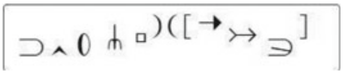
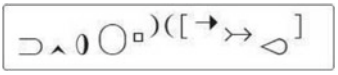


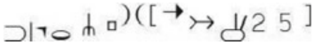
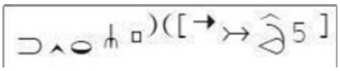
Fig. 12 Avatar sign for “Ka” alphabet

Table 2 Gujarati alphabet's HamNoSys notation and synthetic animation

No.	Gujarati alphabet	HamNoSys	Avatar
1	ક Ka		
2	ખ Kha		
3	ગ Ga		
4	ઘ Gha		

(continued)

Table 2 (continued)

No.	Gujarati alphabet	HamNoSys	Avatar
5	ચ Ca		
6	છ Cha		

4 Conclusion

This paper presents the synthetic animation of Gujarati Sign Language Alphabets. These synthetic animation is used to convert from speech to sign language which will be helpful to deaf-dumb community of Gujarat region. This system will take speech as an input then converted into a Gujarati text, further it will converted a HamNoSys notation. HamNoSys notation is used to generate xml format SiGML code that will be used to play a sign by Avatar. This paper also includes literature review of different spoken and sign languages. HamNosys Notation and Avatar is developed for 34 Gujarati consonant alphabets for creating Gujarati Sign Language (GSL) synthetic animation dictionary. In future these work will be extended to convert Gujarati vowel alphabets series to develop Gujarati Sign Language dictionary.

References

1. Nair MS, Nimitha AP, Idicula SM (2016) Conversion of Malayalam text to Indian sign language using synthetic animation. In: 2016 international conference on next generation intelligent systems (ICNGIS). IEEE

2. Kaur S, Singh M (2015) Indian sign language animation generation system. In: 2015 1st International conference on next generation computing technologies (NGCT). IEEE
3. Kaur R, Kumar P (2014) HamNoSys generation system for sign language. In: 2014 international conference on advances in computing, communications and informatics (ICACCI). IEEE
4. Ebling S, Glauert J (2016) Building a Swiss German sign language avatar with JASigning and evaluating it among the deaf community. Univ Access Inf Soc 15(4):577–587
5. Brour M, Benabbou A (2019) ATLASLang MTS 1: Arabic text language into Arabic sign language machine translation system. Procedia Comput Sci 148:236–245
6. Marshall I, Safar E (2005) Grammar development for sign language avatar-based synthesis. In: Proceedings HCII
7. Kaur K, Kumar P (2016) HamNoSys to SiGML conversion system for sign language automation. Procedia Comput Sci 89:794–803
8. Elliott R et al (2008) Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. Univ Access Inf Soc 6(4):375–391
9. Koller O, Bowden R, Ney H (2016) Automatic alignment of hamnosys subunits for continuous sign language recognition. In: LREC 2016 proceedings, pp 121–128
10. Kumar P, Kaur S (2018) Online multilingual dictionary using Hamburg notation for avatar-based Indian sign language generation system. Int J Cogn Lang Sci 12(8):1117–1124
11. Ebling S (2013) Evaluating a Swiss German sign language avatar among the deaf community
12. Khan NS et al (2020) Speak Pakistan: challenges in developing Pakistan sign language using information technology. South Asian Stud 30(2)
13. Varghese M, Nambiar SK (2018) English to SiGML conversion for sign language generation. In: 2018 international conference on circuits and systems in digital enterprise technology (ICCSDET). IEEE
14. Van Zijl L, Barker D (2003) South African sign language machine translation system. In: Proceedings of the 2nd international conference on Computer graphics, virtual Reality, visualisation and interaction in Africa
15. Patel BD et al (2020) ES2ISL: an advancement in speech to sign language translation using 3D avatar animator. In: 2020 IEEE Canadian conference on electrical and computer engineering (CCECE). IEEE
16. Zhao L et al (2000) A machine translation system from English to American Sign Language. In: Conference of the association for machine translation in the Americas. Springer, Berlin, Heidelberg
17. Kang Z (2019) Spoken language to sign language translation system based on HamNoSys. In: Proceedings of the 2019 international symposium on signal processing systems
18. Goyal L, Goyal V (2016) Development of Indian sign language dictionary using synthetic animations. Indian J Sci Technol 9(32):1–5
19. Aasofwala N, Verma S (2021) Survey on Gujarat's deaf and mute school's special educators to understand the education of deaf community in this pandemic. Towards Excellence 13(4)
20. Aasofwala N, Verma S, Patel K (2021) A novel speech to sign communication model for Gujarati language. In: 2021 third international conference on inventive research in computing applications (ICIRCA). IEEE
21. Schulder M, Hanke T (2022) The hamnosys package

Wearable Device to Monitor the Health Parameters of a Cyclist

Sampada Joshi, Harshada Chavan, Poorva Darade, Isha Nikam, and Varsha Pimprale

Abstract Cycling is an increasingly popular form of exercise and transportation, providing numerous health benefits. However, monitoring the health parameters of a cyclist is essential to avoid potential health risks and ensure their safety. Wearable devices have emerged as a promising solution for monitoring health parameters during cycling. This research paper aims to explore the design and development of a wearable device specifically tailored to monitor the health parameters of a cyclist. The device is equipped with sensors to track heart rate, spO_2, body temperature, distance covered, speed, and cadence, and it connects to a mobile application via Bluetooth. The mobile application analyses the data collected by the device and provides insights into the user's overall health and fitness. Additionally, the device is designed to perform various functionalities like routine planning, deciding daily and weekly goals of the user, providing regular water intake reminders, GPS tracking, and alerting the user in case of abnormal fluctuations in health vitals. The algorithm used for heart rate monitoring is based on photoplethysmography (PPG), while the algorithm for distance tracking is based on the accelerometer and gyroscope sensors in the device. The GPS sensor is used to track the cyclist's speed, and the accelerometer and gyroscope sensors are used for cadence tracking. The data collected from these algorithms is used to calculate the number of calories burned by the cyclist. In conclusion, this research paper presents a comprehensive and innovative solution for monitoring the health parameters of a cyclist using a wearable device. The device's accuracy and cyclist-specific features make it a promising tool for cyclists to improve their overall health and fitness.

Keywords Wearable device · Health care · Internet of Things · Machine learning · Sensors · Monitoring health parameters

S. Joshi (✉) · H. Chavan · P. Darade · I. Nikam · V. Pimprale
MKSSS's Cummins College of Engineering for Women, Pune, India
e-mail: sampada.joshi@cumminscollege.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_7

1 Introduction

Users are inclined more towards regulated, well-monitored, and personalized health-care routines in the medical and health management system. The growth of mobile technology has encouraged the use of more portable and feasible gadgets. These developments encourage focus on wellness, health, fitness, and prevention of several diseases, wearable technology is developing into an exceptional health monitoring tool for medical care services. They can monitor an elderly person's health parameters like blood pressure, body temperature, heart rate, pulse rate and respiration in addition to predict the risk such as deteriorating health, worsening diseases, and other life-threatening conditions including blood pressure increase or respiratory issues. The main aim of the project is to create a mobile application and a wearable gadget to track a cyclist's health parameters.

1.1 Research Paper Study

See (Table 1).

Table 1 Research paper study

Authors	Article	Findings	Learning
Lin Lu, MD, Jiayao Zhang, MD, Yi Xie, MD, Fei Gao, MD, Song Xu, MD, Xinghuo Wu, MD, and Zhewei Ye, Prof Dr, MD	Wearable Health Devices in Health Care: Narrative Systematic Review, November 2020	Article aims to review the application and use of current wearable devices available in healthcare domain and highlights the gaps for further research	Develop low-consumption and high-integration sensor technology to improve information accuracy, extend battery life, and user experience [1]
Stephanie Soon, Hafdis Svavarsdottir, Candice Downey, David George Jayne	Wearable devices for remote vital signs monitoring in the outpatient setting: an overview of the field, January 2020	Article presents the combined information on all wearable devise present in the market for outpatient vital signs monitoring	ECG monitoring, reminders of water intake, GPS monitoring was not included [2]
Sergei Kumaniaev, Pavel Korolev	The Wearable Electronic Customizable Device for Monitoring Activity Indicators, May 2022	Article proposes a wearable device for monitoring human activity parameters with a mobile application which collects the data and displays it	High accuracy and low energy consumption of device will contribute to high user satisfaction [3]

(continued)

Table 1 (continued)

Authors	Article	Findings	Learning
Mário Meireles and Paulo J. G. Ribeiro	Digital Platform/ Mobile App to Boost Cycling for the Promotion of Sustainable Mobility in Mid-Sized Starter Cycling Cities, March 2020	Article proposes analysis of different types of digital platforms was done to find the main features that favoured the also it mentioned what more information can be added for promotion of the usage of bicycles in cities	Features like login using social networks, streets identification by type of infrastructure and safety, gamification, parking spots are preferred by the users [4]
Sarmila Yesm, Sadia Sobhan Misty, Kashshaf Labe, and Md. Nasfikur Rahman Khan	A Cost-effective IoT Based Fitness Tracker For Physically Challenged People, July 2021	Article proposes an AI system that observes health parameters using IOT technology for disabled people Capable of extracting data such as BP, body temperature, ECG and monitoring gas and alcohol levels in blood from patients regularly and taking actions	Tracker can have 24/7 service, emergency ambulance service. Add tracker system with the Android application system. This application is used for accessing large database of patients with the help of medical experts [5]

2 Modules

1. Vitals Monitoring

Health indicators like body temperature, breathing rate, oxygen saturation level, and pulse rate will be displayed on the screen periodically. Moreover, red alarms will be displayed when any variations in the aforementioned criteria are noticed. Additionally, it will provide frequent reminders to drink water and reward the user with heart points when the specified task is accomplished.

2. Daily Routine Recommendation

The feature assists the user in determining the ideal routine and duration for achieving their fitness objectives.

3. Navigation Recommendation

The user is assisted in choosing the best path to take to get to the destination by this feature, which also provides directions. The navigation panel shows the distance travelled, the amount of time needed, the current weather, and the severity of traffic congestion.

4. Music Streaming

This module will give the user access to additional entertainment features by letting him or her browse and listen to music of their choosing. The user will have a better experience as a result, which will encourage them to work harder towards achieving their daily goals.

5. Activity Monitoring

The user can keep track of cycling metrics including total distance travelled, speed, cadence, and amount of power used when pedalling.

6. Nutrition Tracking

After the daily goal has been reached, the application shows the number of calories burned. Also, it shows the user's daily progress and sends reminders to drink water at regular intervals.

7. Routine Planner

A complete, detailed graph of the workout will be displayed if the user wants to create a personalized training plan. Based on prior cycling parameters, the application will recommend a distance to ride to burn a given number of calories.

3 Design and Development

The wearable device we have designed is specifically for cyclists and monitors various health parameters. The device is equipped with sensors to track heart rate, distance, speed, and cadence and sends data to the ThingSpeak cloud for analysis. The device is worn on the wrist and is connected to a mobile app via Wi-Fi. The mobile app analyses the data collected by the device and provides insights into the user's health. The device is designed to be lightweight and comfortable to wear during long cycling sessions. The sensors are accurate and provide real-time data to the user (Fig. 1).

4 Algorithms

There are several algorithms that can be used in a wearable device to monitor the health parameters of a cyclist. Here are some of the common algorithms used in such devices:

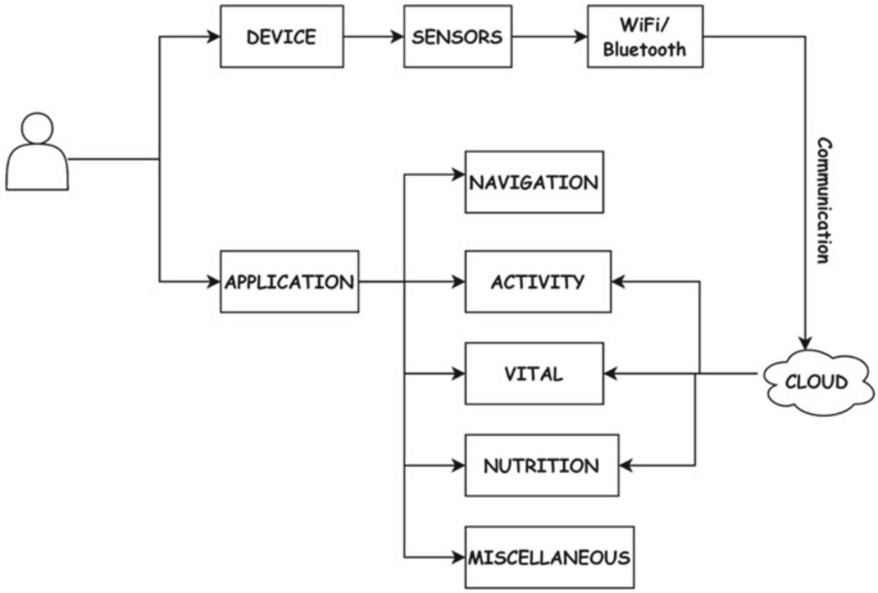


Fig. 1 Data flow diagram

1. Heart rate monitoring: Heart rate monitoring is a crucial parameter for a cyclist as it helps to track the intensity of the exercise and prevent any potential health risks. The algorithm used in heart rate monitoring is based on photoplethysmography (PPG), which is a non-invasive method that measures changes that occur in volume of blood.
2. Distance tracking: Distance tracking is an important parameter to monitor as it helps the cyclist to track their progress and set goals. The algorithm used for distance tracking is based on the accelerometer and gyroscope sensors in the wearable device, which track the motion of the cyclist and calculate the distance covered.
3. Speed tracking: Speed tracking is another important parameter for a cyclist, as it helps to track the intensity of the exercise and set goals. The algorithm used for speed tracking is based on the GPS sensor in the wearable device, which tracks the location of the cyclist and calculates the speed based on the distance covered over time.
4. Cadence tracking: Cadence tracking is the number of revolutions per minute (RPM) of the cyclist's pedal. The algorithm used for cadence tracking is based on the accelerometer and gyroscope sensors in the wearable device, which track the motion of the cyclist and calculate the number of pedal revolutions per minute.
5. Calorie tracking: Calorie tracking is a crucial parameter for a cyclist as it helps to track the number of calories burned during the exercise. The algorithm used for calorie tracking is based on the heart rate monitoring and distance tracking

algorithms. The data collected from these algorithms is used to calculate the number of calories burned by the cyclist.

In conclusion, the algorithms used in a wearable device to monitor the health parameters of a cyclist are based on various sensors such as PPG, GPS, accelerometer, and gyroscope. These algorithms provide accurate data on parameters such as heart rate, distance, speed, cadence, and calories burned, which can help the cyclist to improve their overall health and fitness. The product as of now is tested for heart rate, speed, temperature, and SpO_2 with MAX30102 and MPU6050.

5 Product Comparisons

See (Table 2).

The product designed in the project is named Fitzee.

6 Results

The accuracy of the device measures up to 97%. It was found that the readings on the available smartwatches was up to 78% accurate. Temperature feature is not available on other smart watches. Constant measuring of vitals while smartwatches take up to 10–12 s to show a single reading. The main gap that the product addresses is the accuracy of the vitals readings is more than the famous available smart watches (Figs. 2 and 3).

7 Challenges Faced

The selection of the microcontroller board for the product was the main challenge keeping into account the power consumed, compatibility with the sensors and compact design. The next was the connectivity module used either Bluetooth or Wi-Fi chip to connect to the mobile device. NodeMCU had a built-in Wi-Fi chip which made it easier to connect to the Internet, ThingSpeak server, and mobile application keeping the design compact. The choice and placement of the sensors on the body part of the cyclist for accurate readings was a challenge. The overall budget of the product must not exceed the price of the available products in the market.

Table 2 Product comparisons

Features	Google fitbit	Strava app	Hammer smart watch	Apple watch series 8	Fitzee
Battery charging speed	Fully charged in 1–2 h	Operates with mobile battery life	Fully charged in 1–2 h	Fully charged in 2.5 h	Fully charged in 30 min and operates for 6 h which is more than the average cycling time
Speed parameter of cyclist	The application expects the user to enter the speed and keeps track of it	The speed of the cyclist is real-time displayed	No mention of the speed of the cyclist/user	The speed is calculated using built-in GPS	Speed is calculated using an accelerometer sensor giving the upper hand in accuracy
Daily exercise routine	No such feature available	No such feature available	This device tracks your favourite exercise routines and gets live stats on your wrist. It also finds out if your body is ready to workout or it needs rest	It recommends workouts based on things you typically do with the workout app on apple watch as well as your favourite workout apps that work with health apps [6]	This device will suggest a daily workout routine programme based on your past workout routines, your health parameter readings, and your past health-related diseases
SpO_2 tracking	No such feature available	No such feature available	SpO_2 sensor is built-in feature of this device. It helps to track blood oxygen level in human body	This device can take blood oxygen measurement at any time with the blood oxygen app	This device also monitors blood oxygen level using SpO_2 sensors
Vitals fluctuation notifications	No such notifications displayed	In case of irregular heartbeat, it will send alert notification to the user	No such notifications displayed	In this device, the IRNF analyses pulse rate data collected by the photoplethysmography (PPG) sensor and it sends an alert notification to the user when such episode is detected	This device detects irregular heartbeat or any other fluctuations in any other health parameters. It will send an alert notification to the user

Fig. 2 MAX30102 sensor readings

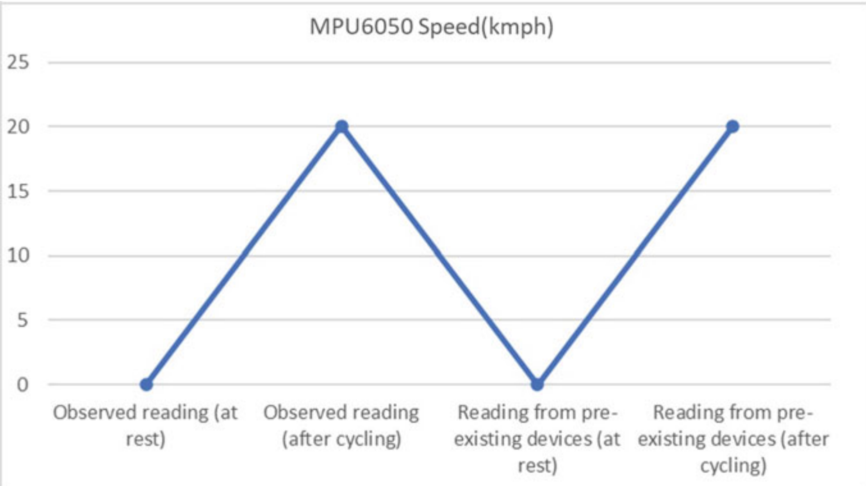


Fig. 3 MPU6050 sensor readings

8 Conclusion

Wearable health devices are an advanced development in the medical field and still a lot of research and study is going on keeping the goal to integrate these devices in the healthcare system. The concept behind Wearable Health Devices is integration of several different biosensors, smart processing and alerts, to support healthcare applications, while interacting with people in healthcare management, using the technology that is still in research and not available in market for use. The existing devices present in the market are highly expensive and are working based on wireless technology that is not available everywhere in many countries. Wearable Health Devices contribute to lowering healthcare management expenses by introducing healthcare methods that are preventive, cutting equipment cost and labour costs, and removing different services. The technology which is based on the concept of wireless sensor is combined with the human health monitoring device that is based on the Internet of Things to test and predict readings of the health-related parameters. The test results are processed and analysed using the data collected from the device. Observations are made that the health monitoring system for humans is designed using IoT which is efficient and stable compared to other technologies and has provides accurate data, real-time monitoring and alerts and emergency notifications. The human health monitoring system based on the Internet of Things designed in this study has completed collecting the health parameters of the user's such as blood pressure, pulse, heart rate, and oxygen levels [7]. After long-term collection of data, factors that are related to predicting high risk health conditions that must be explored in future. This provides a scientific and effective platform for prevention of high risk diseases.

9 Future Scope

The device will be able to take necessary decisions by itself by incurring machine learning and AI. The personalized routine for every cyclist will be suggested which will be based on Body Mass Index calculated by taking weight and height inputs from the user. Record vital data for clinic diagnosis and prediction by medical professionals [8].

References

1. Lu L, Zhang J, Xie Y, Gao F, Xu S, Wu X, Ye Z (2020) Wearable health devices in health care: narrative systematic review
2. Soon S, Svavarsdottir H (2020) Wearable devices for remote vital signs monitoring in the outpatient setting: an overview of the field
3. Kumaniaev S, Korolev P (2022) The wearable electronic customizable device for monitoring activity indicators
4. Meireles M, Ribeiro PJ (2020) digital platform/mobile app to boost cycling for the promotion of sustainable mobility in mid-sized starter cycling cities
5. Yesm S, Misty SS, Labe K, Khan MN (2021) A cost-effective IoT based fitness tracker for physically challenged people
6. Brand-AppleFitnessAnnounceApple. https://www.apple.com/ca/apple-fitness-plus
7. Ru L, Zhang B, Duan J, Ru G, Sharma A, Dhiman G, Gaba GS, Jaha ES, Masud M (2021) A detailed research on human health monitoring system based on Internet of Things
8. Dias D, Paulo Silva Cunha J (2018) Wearable health devices—vital sign monitoring, systems and technologies

A Machine Learning-Based Detection of IoT Cyberattacks in Smart City Application

Santosh Gore, Anuradha S. Deshpande, Nitin Mahankale, Sandip Singha, and Dheeraj B. Lokhande

Abstract The creation of smart cities using IoT-enabled technology has resulted in considerable improvements in operational efficiency and service quality. However, as smart city networks expand, the possibility of cybersecurity attacks and threats increases, which can cause IoT devices to malfunction and potentially threaten people's lives. This study investigates the use of various machine learning algorithms like LR, DT, SVM, RF, ANN, and KNN in smart city networks to sense assaults and anomalies. Ensemble approaches like boosting, bagging, as well as stacking are also used to improve the detection system's effectiveness. The suggested technique effectively identifies cyberattacks, according to experimental data, and the ensemble technique like stacking beats previous models with enhanced accuracy as well as other performance metrics. This study emphasizes the potential of employing machine learning algorithms and ensemble approaches to identify cyberattacks in IoT-based smart cities, as well as the necessity of the selection of relevant features with cross-validation techniques for increased accuracy and performance.

Keywords IoT · Network security · Smart City · Machine learning · Cyberattacks · Ensemble model

S. Gore (✉)
Sai Info Solution, Nashik, India
e-mail: gore.santosh87@gmail.com

A. S. Deshpande
JSPM's Imperial College of Engineering and Research, Pune, India
e-mail: panuradha2010@gmail.com

N. Mahankale
Symbiosis Centre for Management Studies, Pune, India
e-mail: nitin.mahankale@scmspune.ac.in

S. Singha
Indian Institute of Information Technology, Kalyani, India
e-mail: sandip_phd21@iiitkalyani.ac.in

D. B. Lokhande
Shree Siddheshwar Women's College of Engineering, Solapur, India
e-mail: dheerajlokhande1985@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_8

1 Introduction

The Internet of Things (IoT) links diverse devices including smart home sensors, environmental sensors, surveillance equipment, medical devices, and industrial robots to allow for seamless information exchange [1]. It has grown in popularity around the world, with over 27 billion connected devices in 2017 and a projected increase to 125 billion by 2030. However, the employment of many facilities, tools, and protocols complicates the maintenance of future IoT structures, resulting in system vulnerabilities [2]. Extensively employed IoT devices in applications of smart city, causes unauthorized cyber-attacks, such as the Miria botnet assault, which affect citizens' privacy. Symantec predicted a 600% increase in IoT platform threats in 2019, highlighting the considerable security concerns provided by smart city applications including zero-day assaults [3]. Intelligent cyber-attack detection is important before smart city activities are disrupted. Thus, IDS (intrusion detection systems) are optimized for IoT networks. The information gathered by IoT devices is analysed and stored on cloud servers. However, the increased number of IoT devices has caused cloud congestion and delay difficulties, which can be remedied via fog computing [4]. Fog computing distributes computational workload across fog layer devices, lowering energy usage, network traffic, and latency. Also, the fog layer cyber-attacks detection benefits early detection of attacks without disruption to the usual flow of urban life. Machine learning approaches can detect attacks in real-time in less processing time [5]. Thus, we investigate a machine learning-based technique for detecting attacks and abnormalities in IoT networks, addressing the problem of recognizing compromised IoT devices. The model is implemented with training to learn from data in dispersed fog networks and sense attacks and anomalies. We use an ensemble technique [6] like stacking, boosting, and bagging, to improve performance metrics like accuracy as well as F1-score, precision, and recall.

2 Related Work

Contextual evidence on machine learning (ML) techniques for identifying irregularities and intrusions in IoT network traffic [7] is provided in the related studies. It analyses significant studies on the subject and emphasises the significance of securing sensitive information in IoT networks from cyber threats. Benkhelifa [8] evaluated actual issues in cyber-attacks intended for the Internet of Things (IoT) and examined advancements and issues in intrusion recognition methods for IoT in a recent publication [9]. Pahl and Aubet [10] employed the ML technique to forecast IoT service behaviour by monitoring the service communication that takes place in disseminated multi-dimensional microservices of IoT sites. Haq et al. introduced an ensemble technique in [11] incorporating a decision tree, Naive Bayes, and Bayesian Net classifier. When tested using the tenfold cross-validation approach, the ensemble true positive rate generated is 98%. Gaikwad et al. [12] proposed a

Table 1 Techniques of intrusion and anomaly detection over IoT network

References	Classification technique	Dataset	Technique	Accuracy in %
Pahl et al. [10]	Multi-class	Own	K-means	96.3
Diro et al. [5]	Multi-class	NSL-KDD [15]	Neural network	98.27
Khraisat et al. [16]	Multi-class	Bot-IoT [17]	C5 classifier One class SVM	99.97
Alrashdi et al. [18]	Binary	UNSW-NB15	RF	99.34
Zhou et al. [19]	Multi-class	NSL-KDD [15] AWID [20] CIC-IDS2017 [21]	Ensemble	99.8 99.5 99.9

bagging ensemble approach with REPTree as the basic classifier. On the NSL-KDD dataset, their model attained an accuracy of 81.29%. Jabbar et al. [13], developed a method of an ensemble comprised of KNN and ADTree (alternating decision tree) with performance assessment revealing a higher detection rate of 99.8% achieved. Bagging and boosting ensemble approaches were presented in [14], with random forest and decision tree as foundation classifiers. Research on the dataset of NSL-KDD showed better results using bagging including decision trees. Table 1 reviews noteworthy research explaining intrusion as well as anomaly recognition in networks utilizing ML systems, along with their ensemble methodologies in some works.

Research Gap. The existing strategies used for the detection of intrusion threats across IoT networks in smart cities are not specifically stated in the information presented. The research does mention, however, that previous works have mainly employed signature-based algorithms for detecting attacks and abnormalities, which have substantial outlays and are susceptible to identified threats. As a result, the research investigates the viability of ensemble-based learning versus a model using only a classifier for detecting cyberattacks over the IoT network in smart city networks. Furthermore, the research covers a multi-class classification setting, and incorporation of cross-validation with feature selection, which are rare ML methodologies in the literature for this domain.

3 Proposed Model for Cyberattacks Detection

The figure depicts our suggested methodology for monitoring network traffic transient through every fog node. Because fog nodes adjacent to IoT devices are better positioned to detect cyber-attacks on fog nodes rather than the cloud centre. The method also allows for the early detection of assaults and the warning of IoT and network administrators, allowing them to assess and improve their systems.

We chose an anomaly-based NIDS for this study among HIDS and NIDS. To track and identify malicious activities, HIDS (host-based) requires software installation on each network-connected device [22], which is impractical for most resource-constrained IoT devices with limited capabilities like lock doors, smart lamps and watches. Signature-based NIDS [23], instead, incurs substantial processing costs in database storage of attacks and is unable to identify a novel attack in traffic of a latent network, making anomaly-based NIDS the best option. We use this NIDS data to construct ML models using an ensemble technique for detecting aberrant activity in IoT fog networks (Fig. 1).

IDS Dataset. The UNSW-NB15 dataset [24] is a useful and recent IDS dataset with contemporary threats. The dataset was established in 2015 to detect and classify regular and malicious network traffic. The raw network packets are produced in the cyber range lab of the Australian Centre for Cyber Security (ACCS) through the IXIA PerfectStorm tool [25]. Dataset cleaning, and visualization with feature engineering, as well as vectorization, have all been performed. For our research, we selected 175,341 samples at random from a pool of over 2.54 million including 56,000 benign and 119,241 attack samples. The dataset is separated into a training set and test set of 140,272 and 35,069 samples respectively, with both sets comprising the same proportion of attack and benign samples as the original (Table 2).

Data Pre-processing. Selecting relevant features is crucial for improving the efficacy of the model, as it helps reduce over-fitting, improve accuracy, and reduce training time [26]. In this work, we employed feature selection using the information gain ratio method, which selects only the most significant features for prediction. We selected the top 25 features built on their information gain ratio for given datasets.

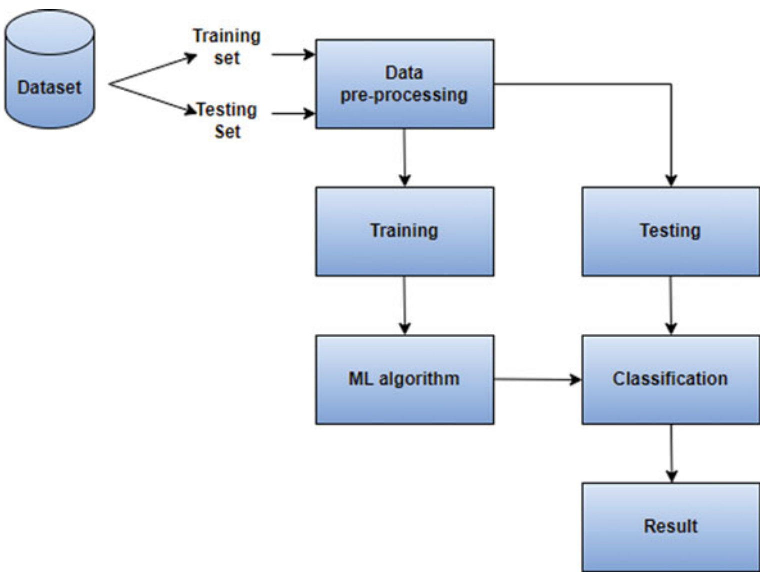


Fig. 1 Architecture of proposed cyberattack detection model

Table 2 Distribution of samples as normal and attack [24]

S. No.	Class	Details	Samples
1	Fuzzers	The network suspended by randomly generated data	18,184
2	Backdoors	Security is circumvented stealthily	1746
3	Exploits	Occur due to security flaws and vulnerabilities	33,393
4	Worms	Replicates to spread over all connected devices	130
5	Shellcode	Small code segment exploiting a vulnerability	1133

We only considered features with a higher ratio greater than 0.5 in the UNSW-NB15 dataset for better classification of benign and malicious classes.

Machine Learning Model. In our study, we utilized popular machine-learning techniques such as LR, SVM, DT, RF, KNN, and ANN. They are commonly employed for developing IDS systems. Ensemble approaches are commonly applied in machine learning, where multiple base models are integrated to create a single superior predictive model [27]. By combining numerous models, ensemble methods aim to improve accuracy and overcome weaknesses in individual models. In the literature, three types of ensemble approaches are commonly used. Bagging is known as a parallel ensemble method that generates basic learners in parallel to enhance the ML algorithm's accuracy. While the Boosting ensemble technique is a sequential strategy that aims to decrease variation and bias in the supervised type of ML algorithms by sequentially generating base learners. The ensemble stacking technique is a learning strategy that creates a new dataset by combining predictions from multiple basic classification models and then uses another classifier to solve the problem [28].

4 Results

For the UNSW-BC15 dataset, the accuracy of the ML algorithms is obtained as shown in Table 3. It shows poor performance of SVM whereas DT and RF achieved better performance than others. However, the performance of the stacking ensemble illustrates efficient performance with RF and ANN as base classifiers and DT as a meta-classifier. Bagging and boosting techniques use RF and DT as base learners.

5 Discussion

To detect IoT cyberattacks in smart city applications, the researchers used a variety of machine learning algorithms, including DT, KNN, LR, ANN, SVM, RF, bagging, boosting, and stacking. The results demonstrated that machine learning techniques can be successful in detecting cyberattacks, with DT and RF being more effective for the majority of attacks. According to the study, combining machine learning

Table 3 Accuracy and performance of ML model in IoT cyberattack detection

Sr. No	Algorithm	Accuracy (%)	Precision	Recall	F1 Score
1	LR	72.32	0.72	0.72	0.71
2	RF	78.89	0.78	0.79	0.78
3	DT	71.49	0.7	0.71	0.70
4	SVM	81.77	0.82	0.82	0.82
5	ANN	80.69	0.81	0.81	0.80
6	KNN	78.23	0.79	0.78	0.78
7	Stacking	83.84	0.83	0.83	0.83
8	Boosting	83.30	0.83	0.83	0.81
9	Bagging	82.36	0.82	0.82	0.81

Table 4 Comparison of proposed model performance with other model [29]

Algorithm	Accuracy		Precision		Recall		F1 Score	
	Our	[29]	Our	[29]	Our	[29]	Our	[29]
LR	0.723	0.538	0.72	0.414	0.72	0.538	0.71	0.397
RF	0.789	0.755	0.78	0.755	0.79	0.755	0.78	0.724
DT	0.715	0.733	0.7	0.721	0.71	0.733	0.7	0.705
SVM	0.818	0.581	0.82	0.586	0.82	0.581	0.82	0.496
KNN	0.782	0.622	0.79	0.578	0.78	0.622	0.78	0.576
Boosting	0.833	0.608	0.83	0.502	0.83	0.608	0.81	0.526

approaches, such as stacking, can improve detection accuracy, and a multi-class classification strategy can provide significant insights into the nature and severity of various sorts of attacks.

Comparing the performance of the suggested model with other models gives enhanced outcomes of our model as shown in Table 4. It shows better performance of proposed model as related to existing model in [29]. Also, Fig. 2 illustrates the comparison between performances of our model and other model graphically.

The study's findings emphasize the difficulties in detecting and preventing cyberattacks in IoT networks, as well as the significance of establishing robust and effective IDS for smart cities. Since smart cities rely largely on IoT devices to maintain critical infrastructure, any hacks on these systems might be disastrous. Hence, real-time recognition and anticipation of cyberattacks are crucial for ensuring the security and integrity of these systems.

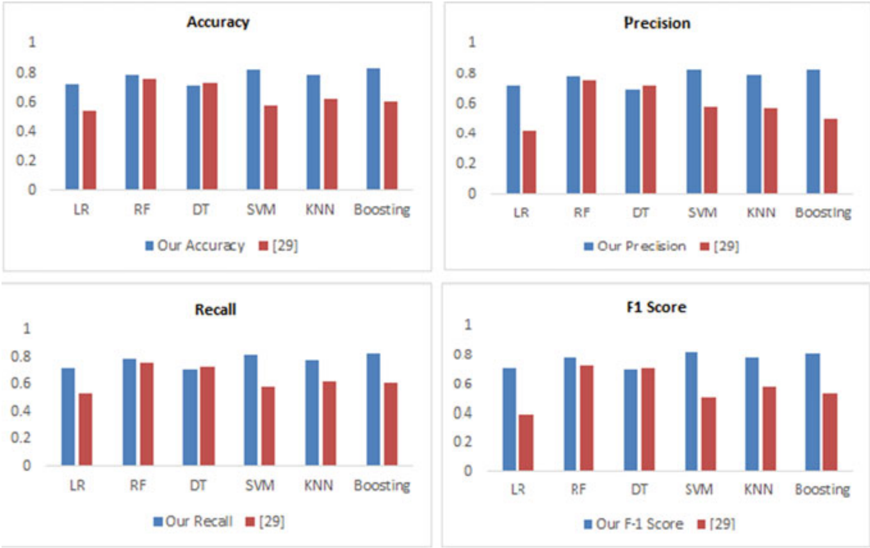


Fig. 2 Comparing proposed and existing model—Accuracy, Precision, Recall, F1 Score

6 Conclusion

We developed an ensemble-based learning strategy to recognize IoT cyberattacks using a single classifier for smart city applications. A recent database of IoT attacks is used in experiments illustrating outperforms stacking technique in detecting attacks from a dataset with benign samples. The proposed method additionally includes the most powerful feature selection approaches prior to developing the model resulting in higher detection accuracy. Furthermore, important metrics like accuracy as well as F1-score precision, and recall, of our ensemble technique with stacking outperform other previous ensemble models in categorizing attack types. Our findings indicate that the employment of machine learning techniques, particularly ensemble learning with stacking, is able to augment the security of IoT-based smart city applications significantly. Since smart city automation become too common, they are more susceptible to cyber threats, which can have serious economic, social, and health consequences. As a result, our findings shed light on the prospective of machine learning to improve the security of these systems and protect vital infrastructure. In the future, we intend to investigate deep learning techniques to increase IoT threat detection performance even further. Future studies will provide more information in this regard.

References

1. Chowdhury A, Karmakar G, Kamruzzaman J (2019) The co-evolution of cloud and IoT applications. igi-global.com, pp 213–234
2. Pahl MO, Aubet FX, Liebald S (2018) Graph-based IoT microservice security. In: IEEE/IFIP network operations and management symposium: cognitive management in a cyber world, NOMS 2018, pp 1–3
3. Symantec (2020) Internet security threat report. Symantec Tempe, AZ, USA, p 1
4. Yousefpour A, Ishigaki G, Jue JP (2017) Fog computing: towards minimizing delay in the Internet of Things. In: Proceedings of IEEE 1st international conference on edge computing EDGE, pp 17–24
5. Diro AA, Chilamkurti N (2018) Distributed attack detection scheme using deep learning approach for Internet of Things. Futur Gener Comput Syst 82:761–768
6. Ludwig SA (2019) Applying a neural network ensemble to intrusion detection. J Artif Intell Soft Comput Res 9(3):177–188
7. Zhang K, Zhu Y, Maharjan S, Zhang Y (2019) Edge intelligence and blockchain empowered 5G beyond for the Industrial Internet of Things. IEEE Netw 33(5):12–19
8. Benkhelifa E, Welsh T, Hamouda W (2018) A critical review of practices and challenges in intrusion detection systems for IoT: toward universal and resilient systems. IEEE Commun Surv Tutorials 20(4):3496–3509
9. Zhang K, Leng S, Peng X, Pan L, Maharjan S, Zhang Y (2019) Artificial intelligence inspired transmission scheduling in cognitive vehicular communications and networks. IEEE Internet Things J 6(2):1987–1997
10. Pahl M, Aubet F (2018) All eyes on you: distributed multi-dimensional IoT microservice anomaly detection. In: 14th International conference on network and service management (CNSM). IEEE, pp 72–80
11. Haq NF, Onik AR, Shah FM (2015) An ensemble framework of anomaly detection using hybridized feature selection approach (HFSA). In: IntelliSys 2015—Proceedings of 2015 SAI intelligent systems conference, pp 989–995
12. Gaikwad DP, Thool RC (2015) Intrusion detection system using bagging ensemble method of machine learning. In Proceedings—1st international conference on computing, communication, control and automation, ICCUBEA, pp 291–295
13. Jabbar MA, Aluvalu R, Sai Satyanarayana Reddy S (2017) Cluster based ensemble classification for intrusion detection system. In: ACM international conference proceeding series, F1283, pp 253–257
14. Pham NT, Foo E, Suriadi S, Jeffrey H, Lahza HFM (2018) Improving performance of intrusion detection system using ensemble methods and feature selection. In: ACM International conference proceeding series
15. Tavallaee M, Bagheri E, Lu W, Ghorbani AA (2009) A detailed analysis of the KDD CUP 99 data set. In: IEEE symposium on computational intelligence for security and defense applications, CISDA
16. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J, Alazab A (2019) A novel ensemble of hybrid intrusion detection system for detecting internet of things attacks. Electronics 8(11)
17. Koroniotis N, Moustafa N, Sitnikova E, Turnbull B (2019) Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset. Futur Gener Comput Syst 100:779–796
18. Alrashdi I, Alqazzaz A, Aloufi E, Alharthi R, Zohdy M, Ming H (2019) AD-IoT: Anomaly detection of IoT cyberattacks in smart city using machine learning. In: 2019 IEEE 9th annual computing and communication workshop and conference CCWC 2019, pp 305–310
19. Zhou Y, Cheng G, Jiang S, Dai M (2020) Building an efficient intrusion detection system based on feature selection and ensemble classifier. Comput Netw 174
20. Kolias C, Kambourakis G, Stavrou A, Gritzalis S (2016) Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset. IEEE Commun Surv Tutorials 18(1):184–208

21. Sharafaldin I, Lashkari AH, Ghorbani AA (2018) Toward generating a new intrusion detection dataset and intrusion traffic characterization. In: Proceedings of the 4th ICISSP, pp 108–116
22. Besharati E, Naderan M, Namjoo E (2019) LR-HIDS: logistic regression host-based intrusion detection system for cloud environments. J Ambient Intell Humaniz Comput 10(9):3669–3692. https://doi.org/10.1007/s12652-018-1093-8
23. Zarpelão BB, Miani RS, Kawakani CT, de Alvarenga SC (2017) A survey of intrusion detection in Internet of Things. J Netw Comput Appl 84:25–37
24. Moustafa N, Slay J (2016) The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. Inf Secur J 25(1–3):18–31
25. Moustafa N (2017) Designing an online and reliable statistical anomaly detection framework for dealing with large high-speed network traffic. In: Ph.D. Thesis, University New South Wales, Canberra, Australia
26. Mohammed MA (2021) A multi-agent feature selection and hybrid classification model for Parkinson's disease diagnosis. ACM Trans Multimed Comput Commun Appl 17(2s)
27. Hansen LK, Salamon P (1990) Neural network ensembles. IEEE Trans Pattern Anal Mach Intell 12(10):993–1001
28. Odegua RO (2020) An empirical study of ensemble techniques (bagging, boosting and stacking). Deep Learn. IndabaX 12(10):1683
29. Vinayakumar R, Alazab M, Soman KP, Poornachandran P, Al-Nemrat A, Venkatraman S (2019) Deep learning approach for intelligent intrusion detection system. IEEE Access 7:41525–41550

A Comprehensive Survey on Security Aspects of Using Blockchain Technology for Digital Degree Verification Solutions

Kamiya Sharma and Monica Gahlawat

Abstract Education plays a critical role in a person's life, and obtaining a degree certificate along with knowledge can promise a successful career. However, when students apply for jobs, organizations manually verify their certificates, which is a tedious and time-consuming process. Unfortunately, there are instances when students present counterfeit certificates, making it challenging for organizations to identify them. The verification process for academic credentials can be complex and time-consuming. Moreover, managing digital documents can also pose its own set of challenges. Digital documents can be vulnerable to hacking and data breaches, and ensuring that the information in these documents is accurate and up-to-date can be challenging. To address these challenges, universities and other educational institutions are increasingly turning to various technologies for verifying academic credentials. Blockchain technology is not governed by a central authority due to its decentralized nature, which streamlines the verification process and improves the security and accuracy of academic credentials [1]. Additionally, this study focuses on security aspects with current blockchain-based systems for digital degree verification.

Keywords Digital degree · Blockchain technology · Security attacks · Privacy · Smart contracts · Distributed ledger

1 Introduction

Various industries have shifted from the traditional approach to network-based applications for storing digital documents. This simply means that document management via the web has become an essential component of running any organization. In academics, the primary challenge is to secure digital credentials, especially degree

K. Sharma · M. Gahlawat (✉)
L.J. Institute of Computer Applications, LJ University, Ahmedabad, India
e-mail: Monica.gahlawat@ljku.edu.in

K. Sharma
e-mail: kamiya.sharma_ljmca@ljku.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_9

Table 1 A glance at prominent fake degree cases [2–8]

Country	Prominent fake degree cases
U.S.A	Ranks first in housing fraudulent degrees
U.K	The BBC claimed that 40 websites were shut down for selling fake degrees to the UK in 2015
Liberia	The American non-profit news agency discovered 120 linked institutes managed by the American scam that distributes fake degrees in a variety of professions
Armenian	In 2018, 40 Israeli doctors were detained on suspicion of obtaining forged Armenian diplomas
Vietnam	According to The Asian Reporter, more than 1700 police officers were promoted in 2005 using forged degrees
Singapore	According to a 2019 survey, Singapore ranks first in Asia Pacific for bogus degree certifications
China	In 2008, Shenzhen addressed the issue of fake degrees on a daily basis, with at least 500,000 persons utilizing fake degree certificates in China
India	Manav Bharti University in Himachal Pradesh issued 36,000 fake degrees across different states in 2021

certificates that are crucial for students when applying for jobs. The management of student mark sheets is a challenging task for universities due to the considerable limitations of centralized degree management having a single point of failure. The current degree issuance mechanism is not trustworthy or verifiable, which makes it easy for counterfeit degrees to be produced. Table 1 provides a summary of prominent fake degree cases.

This research paper is structured into several sections. Section 1 provides an introduction and an overview of counterfeit degree fraud around the globe. Section 2 presents the architecture and features of blockchain. Section 3 provides a study of the literature survey related to security parameters and attacks involving different blockchain-based digital degree solutions. Finally, Sect. 4 concludes the paper with future recommendations.

2 Taxonomy of Blockchain Technology

Blockchain is a secure and decentralized method of sharing data over a network through an immutable ledger. A single block carries a hash value of the previous node, timestamp, and nonce, it is helpful for keeping documents safe. The cryptographic **hash** ensures that the contents cannot be tampered with, and the **timestamp** records when the block was validated. The chain is secure because the preceding block hash value connects each block to the previous block. A **nonce** is used to validate the hash and thwart replay attacks. However, the selection of a blockchain network for storing digital certificates impacts the system's efficiency and security. Permissionless networks are not suitable as they allow any peer entity to join and

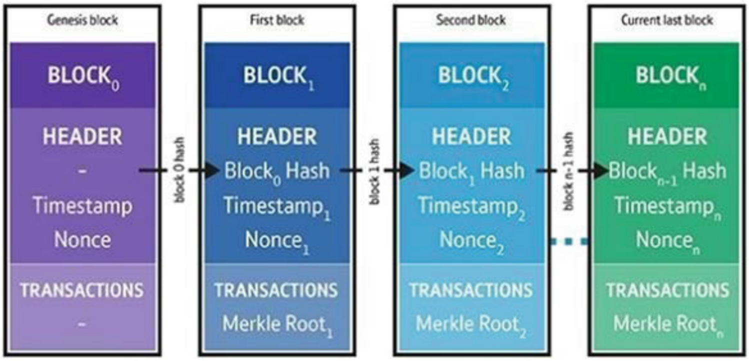


Fig. 1 Block structure of blockchain [9]

store a complete copy of the transactions globally. A block structure of the blockchain is explored in Fig 1.

The National Internet Emergency Response Center (CERT) of China has made public a list of 247 blockchain security vulnerabilities on its website. While public blockchain allows any entity to participate and validate transactions, private blockchain requires authentication of every participating peer entity. A permissioned blockchain requires the consent of a central authority, and its security depends on the honesty of its participants. Permissioned blockchain are the preferred choice for organizations such as those in the education [10, 11] government, and medical sectors to secure their data. Alternatively, organizations can opt for a consortium/hybrid blockchain network, which incorporates both public and private blockchain but carries the risk of attack. Blockchain technology's distinctive properties, including its decentralized architecture, consensus, immutability, and anonymity, can be used to reduce fraud and raise public trust in services through consensus algorithms that validate transactions without a third-party intermediary. The features of blockchain are summarized in Fig. 2.

3 Study of Literature Survey

Various organizations are embracing Blockchain technologies for digital certificate verification. However, current research has proposed various applications of this technology that are not yet a comprehensive solution to handling digital degree credentials, as they do not highlight security parameters such as authentication, privacy, integrity, confidentiality, provenance, and non-repudiation. **Authentication** is the primary challenge when dealing with digital documents, as users must navigate various security obstacles such as encryption, digital signatures, and public–private key cryptography mechanisms (Fig. 3).

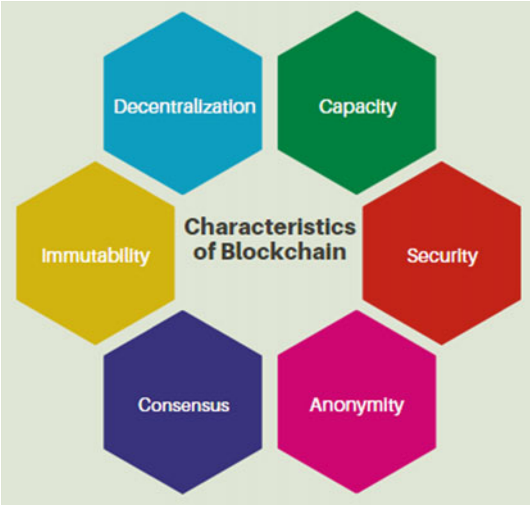


Fig. 2 Blockchain features

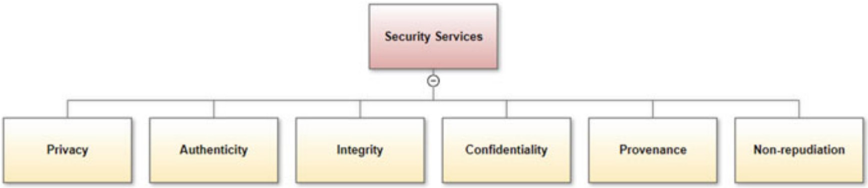


Fig. 3 Security services

Data privacy is a critical aspect of data control, which is achieved through access control mechanisms and public-key cryptography. While public keys are visible, they cannot guarantee document privacy. On the other hand, **data integrity** guarantees that data is correct, comprehensive, and dependable and can be preserved via private keys, public-key cryptography, and hashes. However, to enhance security, improvements are still needed in blockchain algorithms. Hashing techniques and public-key cryptography are cryptographic techniques that ensure the **confidentiality** and integrity of data, while **non-repudiation** is a technique used to ensure that the parties involved in a transaction cannot later deny their involvement or the authenticity of the transaction. By using **data provenance** techniques, organizations can ensure the integrity and traceability of their data, which is particularly important in the education sector, where data accuracy is critical. Table 2 explains security parameters and their key components.

In this literature survey, we have explored and listed some of the shortcomings of blockchain-based degree storage and verification on security parameters.

Table 2 Security parameters with key components

Security service parameters	Key components
Privacy	Public-key cryptography, encryption mechanism, access control
Authenticity	Public-key cryptography, digital signature, encryption mechanism
Integrity	Secret key, public and private key
Confidentiality	Encryption mechanism
Provenance	Data provenance mechanisms
Non-repudiation	Digital signature and access control

3.1 KMI

KMI has been working on developing an open badge infrastructure for the UK higher education sector, which would enable universities and other educational institutions to issue and verify open badges for their students. The Knowledge Media Institute (KMI) has collaborated with various institutions, including the Ghent and Texas universities, on blockchain projects related to education and credentialing. Digital badges are often kept as digital tokens in a public (permissionless) blockchain, and individuals can manage documents using a private key that may be provided by educational institutions and other organizations to recognize achievements and skills. However, KMI application is convoluted and makes the individual's personal data visible on a public blockchain. As a result, there is a security risk because there is no system in place to protect the student's identity. As a result, the data is vulnerable to various sorts of privacy assaults, such as phishing attacks, Trojan attacks, and sniffer attacks, among others.

3.2 RecordsKeeper

RecordsKeeper [12] was established in 2016 by Toshendra Sharma and Rohendra Singh as a solution centered on public blockchain. In this solution, academic credentials are issued by educational institutions with a receipt, which serves as a Transaction ID for students. These receipts are sent to third parties as evidence of the validity of certificates. However, RecordsKeeper does not have an encryption mechanism for user records. If students forget to encrypt their documents, their identity is exposed and the data is visible to all platform users. In a public blockchain, transferring ownership rights to a third party allows for document tampering, which leads to various data confidentiality and integrity attacks such as session hijacking attacks, man-in-the-middle attacks, and deliberate attacks.

3.3 Blockcerts

In terms of data verifiability, Blockcerts [13, 14] based on self-sovereign identity provide more security than KMI. However, there are some concerns about privacy and the potential risks associated with lost private keys. The degree certificate holder has complete control over their data with Blockcerts and can choose whom to share their certificate with. This is a significant advantage over KMI, which relies on a centralized authority to verify and share certificates. However, privacy may be a concern for some individuals. While students can choose to share their certificates with one employer, it is not possible to hide the certificate from another employer. This could potentially limit job opportunities for individuals who may not want certain employers to see their certificates. In terms of security, multi-signature authorization is used during the authentication process to increase security. However, all parties involved in the process need to remember their private keys. There are some potential privacy concerns and risks associated with lost private keys that need to be addressed. Overall, it leads to various attacks such as brute-force attacks, double spending attacks, credential stuffing attacks, DoS attacks, and more.

3.4 EduCTX

EduCTX is a platform that employs Blockchain technology to maintain study credits via the European Credit Transfer and Accumulation System (ECTS). It is necessary for both students and verification agencies to preserve digital identities. It handles ERC20 tokens, which are used to build and issue smart contracts on the Ethereum blockchain, and these tokens store a student's academic credentials. Students acquire tokens that correspond to successfully completed courses using the EduCTX wallet. To ensure that students cannot share tokens with other blockchain entities, multi-signature assignments are employed for security.

A third party sends a message to the student during verification, which the student signs using their private key to guarantee its legitimacy. Students are responsible for maintaining their cryptographic keys. If they fail to do so, there is a risk of various types of attacks such as credential stuffing attacks, selfish mining attacks, denial of service attacks, and others.

3.5 UNIC

The University of Nicosia (UNIC) [15, 16] in Cyprus became the world's first university to offer academic digital diplomas based on Bitcoin's Blockchain technology. The institution provides a software solution that allows a third party to validate these certificates' authenticity. The digital certificates contain a unique digital signature (using SHA-256 algorithm) that is linked to the student's public Bitcoin address and signed by the student's private key, making it easy for employers or any other interested entity to validate the certificate's legitimacy. The student then shares this key with a third party to grant access to the documents. However, it should be noted that students may be unable to authorize an employer to obtain their certificates.

UNIC uses a smart contract to enforce integrity and provenance. Because the smart contract is publicly visible on the blockchain, anyone can inspect and verify the integrity of the contract and the certificate it contains. This eliminates the privacy of provenance data and may make the system vulnerable to various types of attacks, such as sniffing attacks, Distributed denial of service attacks, and others.

3.6 UZHBC

The University of Zurich (UZH) in Switzerland has developed a blockchain-based diploma verification system called UZHBC (University of Zurich Blockchain Credentials). UZHBC uses a smart contract and the public blockchain to regulate the access and confirm the legitimacy of digital degrees. Users can enter their credentials by uploading a PDF file, which is then hashed (using SHA-256 hashing calculation) and stored within the smart contract. If the hashed value of the diploma matches the hashes recorded within the smart contract, the certificate is considered authentic, and the system certifies its authenticity. However, if there is no match, the system generates feedback indicating that the certificate is not authentic.

However, it should be noted that the UZHBC system does not use digital signatures for authenticity verification. To verify the legitimacy of a UZHBC certificate, anyone can access the Ethereum blockchain and use the certificate's unique identifier to look up the certificate on the blockchain. It could compromise the integrity and authenticity of the documents, such as man-in-the-middle attacks, deliberate attacks, Sybil attacks, and other types of authentication attacks.

3.7 SmartCert

SmartCert [17] is a platform that uses RFID and Blockchain technology to verify academic credentials. Universities provide students with certificates encoded as RFID tags, which include the student's name, graduation date, degree type, photograph, and fingerprints. The institution has digitally signed the certificates, and the RFID tag can be scanned by the verifier to confirm the student's identification. This provides transparency in the recruitment process, but it also comes with security risks. The most significant risks to an RFID network are authorization and integrity attacks like phishing attacks, denial of service attacks, eavesdropping attacks, and sniffing attacks.

3.8 BCDiploma

BCDiploma [18] is a safe online certification platform that stores digital credentials in an encrypted manner. BCDiploma encrypts the certificate data using student, institute, and solution provider key, guaranteeing that it is secure and cannot be accessed by unauthorized persons. The encrypted certificates can be posted several times on social networking platforms such as LinkedIn and Twitter. However, BCDiploma is not completely decentralized since the certificate data is encrypted on the blockchain, the process of retrieving and verifying the certificate requires interaction with a centralized system that manages the keys and provides access to the encrypted data. The integrity of the data can be compromised if the primary server goes down, which can also result in Denial of Service attacks, selfish mining attacks, man-in-the-middle attacks and others.

3.9 SuperCert

SuperCert [19] from NITI Aayog illustrates the promise of Blockchain technology that ensures the integrity, and authenticity of certification procedures, making it easier for individuals to promote their talents and achievements and for employers to check the authenticity of certifications. The University provides a hashed version of certificates that are digitally encrypted and signed by the university. SuperCert permits verification through the use of a public key of students and universities. Using a student's and an institution's public key for verification might be challenging when many key pairs are associated with the same block. This is because determining which key pair to use for verification can be difficult. This can potentially result in privacy and authorization attacks such as sniffing, phishing attacks, and others.

Table 3 examined the security aspects of different degree verification solutions based on Blockchain technology. 'Yes' indicates that the corresponding security parameter is adequately addressed, while 'No' indicates that there is a security gap or the parameter is not addressed.

As the world becomes increasingly digital, online platforms are replacing traditional methods for exchanging information. Traditional approaches to document verification are considered unreliable, leading many organizations to explore digital verification options. Blockchain's ability to provide a decentralized, secure, and unchangeable platform is an attractive option for organizations looking to store and manage vast amounts of digital data. Researchers have proposed various solutions for using blockchain to manage digital documents, and Table 4 summarizes previous research on this topic, highlighting significant improvements in security and notable challenges.

Table 3 Comparison of blockchain-based degree verifications solutions [3, 11, 17, 20–23]

	Solutions								
Security Property Violated	KMI	RecordsKeeper	BlockCerts	EduCTX	UNIC	UZHBC	SmartCert	BCDiploma	SuperCert
Confidentiality	No	No	No	Yes	Yes	Yes	No	No	Yes
Privacy	No	Yes	No	Yes	No	No	No	Yes	No
Integrity	Yes	No	Yes	Yes	No	No	Yes	No	Yes
Non-repudiation	Yes	Yes	Yes	No	Yes	No	No	Yes	Yes
Authenticity	No	No	No	Yes	Yes	No	Yes	Yes	No
Data Provenance	Yes	Yes	No	No	No	Yes	Yes	No	Yes

Table 4 Comparison of existing digital degree verification solutions benefits and problems [2, 5, 11, 21, 24–26]

Author detail	Year	Study focus	Benefits	Problems Identified
Anant et al	2021	Using ethereum technology to store education certifications on a distributed ledger	The legitimacy is ensured by smart contract technology Student identities are safeguarded using cryptography	No definite way to verify an employer's legitimacy Absence of a consensus-building process
Gilda et al	2018	Resolve the nested authorization and data consent issue	Student data consent and authorization using Hyperledger Fabric and Composer	The certificate is susceptible to manipulation
Rangi et al	2020	This paper examines Blockchain technology through the lens of the proof-of-concept Supercert for certificate verification solutions	Permissioned architecture reduces the possibility of intelligent impersonation and fake degree verification	Fundamental security precautions are required Vulnerable to proofing attacks
Saleh et al	2020	In this study, researchers examine several security topics in existing blockchain solutions for degree verification	The security requirements for an educational credential relying on Hyperledger Fabric technology	Absence of a consensus process Omits to emphasize the validity of the certificates
Rodelio et al	2018	Provide a blockchain solution for digital degree credentials verification	Privacy for recipients is ensured via the Hyperledger (permissioned) blockchain	Employers are unable to authorize Absence of a consensus process
Jerinas et al	2018	An Ethereum blockchain-based system for verifying and issuing diplomas	The hash stored in smart contracts is used to verify the diploma's authenticity	No defined procedure for protecting students' privacy
Tarek et al	2019	A blockchain smartphone application that stores and verifies digital certificates	The certificate's veracity and legitimacy are confirmed by the Ethereum platform	No attention paid to students' privacy
Turkanović et al	2018	Infrastructure based on ARK technology secures the transfer of student education credits across institutes	The Consortium blockchain employs the Delegated Proof-of-Stake algorithm for permission	The security of certificates is dependent on a single authority's potentially insecure private key

4 Conclusion

This article summarizes multiple research by several authors, bringing together information from well-known sources to investigate the uses of Blockchain technology in academia. One major challenge in the current system is the prevalence of fake degree certificates, highlighted by the 2016 hack of Lynda.com which affected 9.5 million user accounts. Blockchain technology offers a solution to this dilemma by creating a safe and permanent record of educational accomplishments and qualifications. Despite the potential benefits, there are still issues, especially in terms of security flaws and potential attacks [27, 28]. Future research areas might focus on creating more secure verification systems for the education industry and tackling these problems in real-time settings.

References

1. https://acloudguru.com/blog/engineering/Blockchain-and-serverless-processingsimilarities-differences-and-how-they-fit-together. Last accessed 05 Dec 2022
2. Saleh OS, Ghazali O, Rana ME (2020) Blockchain based framework for educational certificates verification. Studies, Planning and Follow-up Directorate. Ministry of Higher Education and Scientific Research, Baghdad, Iraq. School
3. Gilda S, Mehrotra M (2018) Blockchain for student data privacy and consent. In: 2018 International conference on computer communication and informatics (ICCCI). IEEE
4. Gresch J et al (2018) The proposal of a blockchain-based architecture for transparent certificate handling. In: International conference on business information systems. Springer, Cham
5. Tariq A, Haq HB, Ali ST (2022) Cerberus: a blockchain-based accreditation and degree verification system. IEEE Trans Comput Soc Syst
6. Shukla A et al (2020) Academic credential verification technique using blockchain. Int J Adv Sci Technol 29:4244–4254
7. Domingue J (2017) Blockchains as a component of the next generation internet
8. Ahmad S, Qing YX, Frauds avoidance of digital educational certificates using blockchain
9. https://www.unic.ac.cy/verify. Last accessed 21 Feb 2023
10. Arenas R, Fernandez P (2018) CredenceLedger: a permissioned blockchain for verifiable academic credentials. In: 2018 IEEE international conference on engineering, technology, and innovation (ICE/ITMC). IEEE
11. Rangi PK, Aithal PS (2020) A study on blockchain technology as a dominant feature to mitigate reputational risk for Indian academic institutions and universities. Int J Appl Eng Manag Lett (IJAEML) 4(2):275–284
12. Sayed RH (2018) Potential of blockchain technology to solve fake diploma problem, thesis. Upload and verify academic certifications over RecordsKeeper
13. https://www.blockcerts.org/about.html. Last accessed 10 Oct 2022
14. https://www.blockcerts.org/guide/. Last accessed 10 Jan 2023
15. Academic certificates on the blockchain, UNIC blockchain initiative 2018. (Online). Ac cessed 08 Mar 2021
16. University of Nicosia. Academic Certificates on the Blockchain. https://www.unic.ac.cy/iff/Blockchain-certificates/. Last accessed 05 Dec 2022
17. Wibowo T, Christian Y (2021) Usage of blockchain to ensure audit data integrity. Equity 24(1):47–58
18. https://www.bcdiploma.com/en/howItWorks. Last accessed 22 May 2022

19. https://niti.gov.in/sites/default/files/202001/Blockchain_The_India_Strategy_Part_I.pdf. Last accessed 15 Nov 2022
20. World Education Services + Reviews. Diploma mills: 9 strategies for tackling one of higher education's most wicked problems (Internet). Cited 05 May 2020. Available from: https://wenr.wes.org/2017/12/diploma-mills-9-strategies-for-tackling-one-ofhigher-educations-most-wicked-problems
21. Turkanović M, Hölbl M, Košič K, Heričko M, Kamišalić A (2018) EduCTX: a block-chain based higher education credit platform. IEEE Access 6(1):5112–5127
22. Kolvenbach S et al (2018) Blockchain 4 education. In: Proceedings of 16th European conference on computer-supported cooperative work-panels, posters and demos. European Society for Socially Embedded Technologies (EUSSET)
23. (2018) Challenges, and future directions. Int J Intell Syst Appl 10(6):40–48 https://www.ledgerinsights.com/china-launches-national-database-Blockchain-security-vulnerabilities/
24. G R, Sharma MKS, Pathak P (2017) Blockkchain imperative for educational
25. Jat AS (2020) Comparative analysis of existing blockchain based certification dapps
26. Alkhalifah A et al (2020) A taxonomy of blockchain threats and vulnerabilities. In: Blockchain for cybersecurity and privacy. CRC Press, pp 3–28
27. Rathod N, Motwani D (2018) Security threats on Blockchain and its countermeasures. Int. Res. J. Eng. Technol 5(11):1636–1642
28. Jayalakshmi S (2021) The impact of the blockchain on academic certificate verification system-review. EAI Endorsed Trans Energy Web
29. Gresch J et al (2018) The proposal of a blockchain-based architecture for transparent certificate handling. In: International conference on business information systems. Springer, Cham
30. https://www.apriorit.com/dev-blog/578-Blockchain-attack-vectors. Last accessed 25 Nov 2022
31. Kanan T, Obaidat AT, Al-Lahham M (2019) SmartCert blockchain imperative for educational certificates. In: 2019 IEEE Jordan international joint conference on electrical engineering and information technology (JEEIT). IEEE
32. Grolleau G, Lakhal T, Mzoughi N (2008) An introduction to the economics of fake degrees. J Econ Issues 42(3):673–693

Augmented Reality-Based Education Ecosystem for Specially Abled Children

S. Vinodh Kumar, P. Kumar, V. Roshini, M. Shalini, and D. Sharath Srinivas

Abstract Children who are mentally unstable or who have small or major physical organ impairments are referred to as "specially abled" children to help them integrate into life's normal flow and give them a nurturing environment in which they can thrive. Carrying hefty books and unappealing texts affects their level of concentration to a greater extent. An application must be devised to satisfy the ongoing needs of students who need special care in education. Any student requiring special care in education can learn by hearing and seeing things. Students can listen to sound through which they will be able to listen to audio of instructors coaching the subject. In this application, which is built using Android Studio, when one scans a photo in a book, a 3D animation is projected through the mobile phone in 3D. An improvised learning environment can be provided through the application. The concept of AR is so new and powerful for learning. The AR application's main motive is "ANYONE CAN LEARN ANYWHERE ANYTIME". The app is built offline, to be accessed by anyone at any time. For future developments, the online mode can be added to have more access to learning content.

Keywords Augmented reality · Education · Mobile application

S. Vinodh Kumar (✉) · P. Kumar · V. Roshini · M. Shalini · D. Sharath Srinivas
Department of Computer Science and Engineering, Rajalakshmi Engineering College, Thandalam, Chennai, India
e-mail: vinodkumar.s@rajalakshmi.edu.in

P. Kumar
e-mail: kumar@rajalakshmi.edu.in

V. Roshini
e-mail: roshini.v.2019.cse@rajalakshmi.edu.in

M. Shalini
e-mail: shalini.m.2019.cse@rajalakshmi.edu.in

D. Sharath Srinivas
e-mail: sharathsrinivas.d.2019.cse@rajalakshmi.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_10

1 Introduction

Education ecosystem for specially abled students should meet their educational needs through an improvised system. Specially abled children find it extremely difficult to focus on a particular task as they are easily distracted. According to the types of ailments that exist in the kids, the educational ecosystem's needs must be met, and provisions must be made to help pupils feel special, cared for, and inspired, as they develop. It is crucial to create a nurturing atmosphere where Specially Abled Children can live and develop to help them assimilate into daily life. Specially Abled Children may have mild to severe physical or mental impairments. This approach helps create an environment in the educational ecosystem that caters to the different ailments that kids have, and it plans to take care of them and help them grow up with the right attitude. The system demands a visually appealing mode of education and less distractive environment, thereby enhancing the existing education system. According to UNESCO's report on children with disabilities, 27% of impaired children between the ages of 5 and 19 do not have access to education. Additionally, it claims that the education system excludes roughly 3/4 of children with exceptional needs under the age of 5. AR provides an enhanced learning platform due to its interactivity and cutting-edge spatial visualization are appealing characteristics.

2 Literature Survey

The promise of Augmented Reality (AR) to enhance inclusive education has yet to be completely realized. This in-depth research explores the state of employing AR technology as a medium of instruction that considers the student requirements, especially those who are specially abled. The sample size was small (22% of individuals), hence the results addressing the overall impact of AR might not be adequate [1]. This study looked for trends in earlier research on educational augmented reality (AR). In light of this, 105 papers from the ScienceDirect, EBSCOhost, and ERIC indexes were analyzed. Analysis revealed that throughout time, there have been more studies on AR in education. In those articles, quantitative methods were commonly applied, and it was found that educational AR was frequently employed in engineering education, medical training, and science [2]. The system was created with the aid of an Android application. Children and others with visual impairments face challenges from birth. That can be the main reason why people are unable to use devices in their daily lives. Adapting the technology could be difficult for the visually impaired in the initial stages. Blind students who are learning need some tactile, audio, or friend assistance [3]. This study focuses on using augmented reality to bridge the gap between academic and practical knowledge in youth education. An AR system enhances learning by creating a composite view of reality and virtuality. An application that uses a single image target improves kids' understanding of letters and sounds, as well as their ability to memorize and pronounce the English alphabet. However, it's limited to digital platforms and cannot be used in textbooks [4]. A study aims to explore how well applications in AR can enable people with special needs to learn and develop new skills. In compliance with PRISMA guidelines, analyzing the overall impact of AR on people with different disabilities was conducted. This study provides crucial information about the relative effectiveness of AR in helping people with special needs develop their academic and daily living abilities [5]. In higher education settings, students must utilize their own devices to keep track of attendance and participate in class activities using online learning management systems and other teaching and learning resources including schedules and virtual learning environments. To precisely grasp how AR improves kids' learning, more research is necessary. The results might not be applicable to all higher education special education settings [6]. The study aimed to improve vowel understanding among children over four years old by creating an educational mobile app using Extreme programming. This appraised students' performance in learning the alphabet, pronunciation, and numerals. However, it's limited to vowels and digits and challenging to upgrade [7]. This study sought to assess and understand how a mobile application that incorporates AR affects the desire to learn among health science students at the undergraduate level at the University of Cape Town. ARCS model acted as the framework for analyzing the impacts of student motivation with respect to AR's contribution to the field of education [8] .An overview of the application of augmented reality technology in higher education literature published between 2011 and 2020, with a focus on research indexed in eight international databases. The findings may not be applicable to all higher level special education settings. The majority of the research analyzed involved only a small number of participants in their investigations, and this was not implemented [9]. The research on augmented reality applications in education is still in an early stage, and there is a lack of research on the effects and implications of augmented reality in the field of education. The combination of real and virtual objects may cause confusion as students may face difficulty navigating between fantasy and reality [10, 11].

3 Proposed System Architecture

Since it is difficult to hold a specially abled kid's attention for a longer time, we cannot force them to study like normal people. This AR-incorporated application uses 3D models and sounds to make it visually attractive enabling them to be more involved and learn the topic more quickly. This application also requires less assistance to teach them how to use the application. As this is a mobile application which works in offline mode, it can be used to educate anyone, anywhere at any time.

The architecture diagram, as seen in Fig. 1, explains the entire working of the application. Android Studio is used to facilitate user interaction through a user-friendly interface. Using this UI, the student can use the program to log in or register and access the necessary contents. The logged-in credentials are now stored in Firebase. The administrator has full access to the database as well as the ability to add,

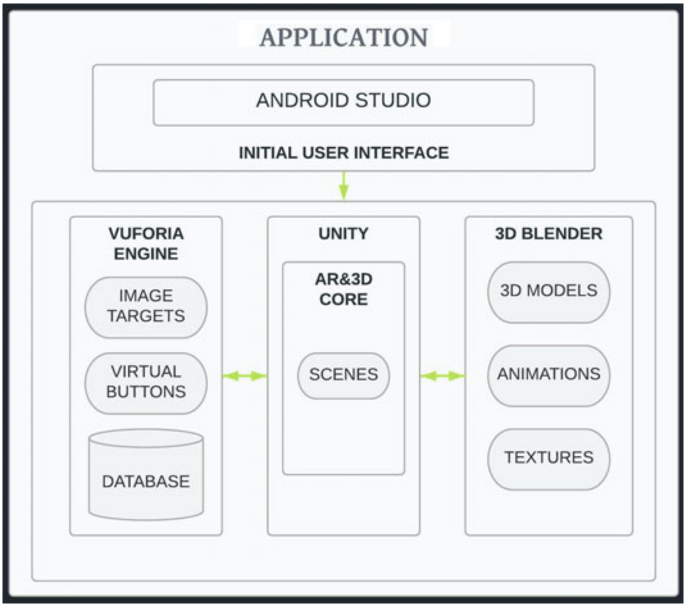


Fig. 1 Architecture diagram

delete, and change application material. Scenes are created using Unity which has 3D models displayed for the corresponding image target. The 3D models are developed using Blender. Suitable animations and textures are added to the models and are later incorporated into a single application. Image targets are created using the Vuforia engine which when scanned shows the AR object in the real world alongside the predefined animations.

4 Implementation and Results

The application utilizes Unity, Blender, and Android Studio for the experiment setup. A camera-equipped smartphone or tablet can show augmented reality visuals and recognize letters of the alphabet. The Android application that incorporates Augmented Reality and 3D models consists of 6 modules to facilitate the academic education of the specially abled. The application is also provided with a layer of authentication and security using a login and register module where the student will be able to log in and use the application. The aim is to develop modules that will make up the Android application that combines 3D models and augmented reality to aid in the academic education of those with special needs. A login and registration module that the student may use to log in and utilize the application adds an additional degree of authentication and protection to the application. The modules include AR Numbers, Human Body and Organ System, Geometrical Shapes, and AR Alphabets in English, Tamil, and Hindi.

4.1 Creating a Scene Using Unity

Unity provides a 2D and 3D platform for game developers to create applications. In this application, as seen in Fig. 2 3D models are imported first. In order to create an AR-based object, a new scene has to be created where the 3D models are imported. Target images are imported by Unity to activate the corresponding AR object. The Vuforia package enables the setting of the complete scene by importing target images and enables the AR camera. The Unity application is then built and incorporated into a stand-alone Android application.

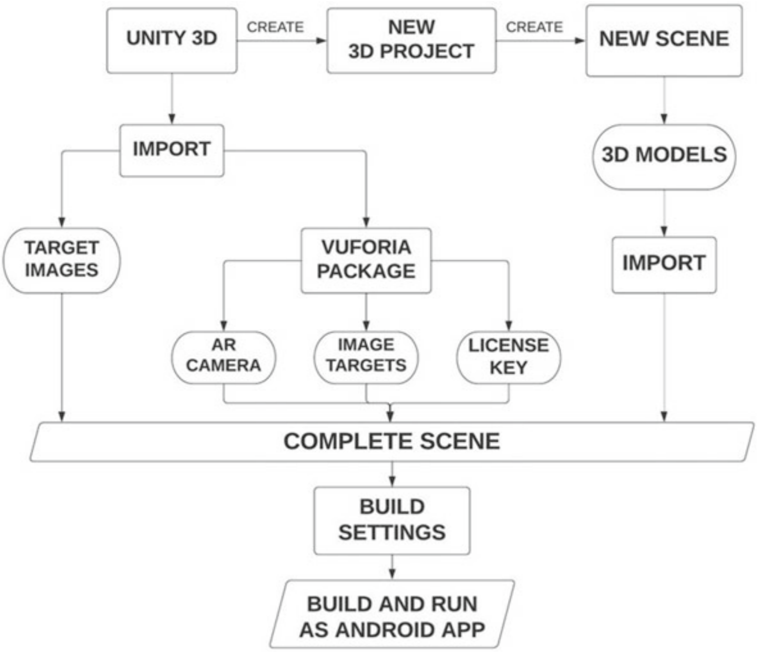


Fig. 2 Unity workflow

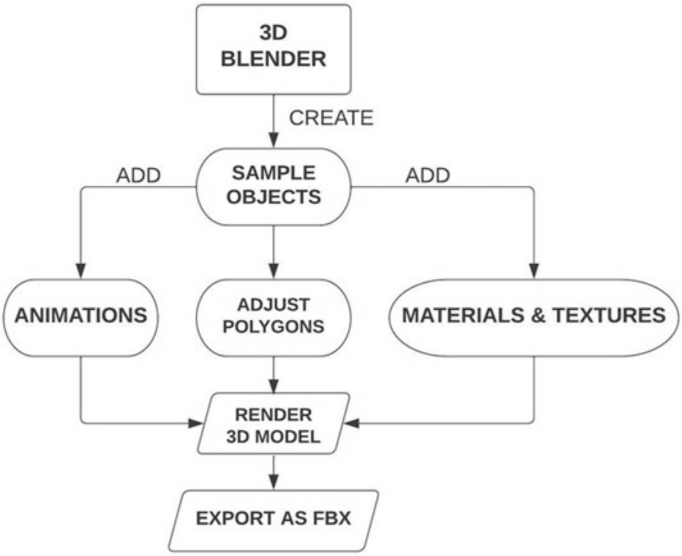


Fig. 3 Blender workflow

4.2 Building 3D Objects Using Blender

In this application, as seen in Fig. 3, Blender is used to create a 3D model by adding the required materials and textures, adjusting shapes, and adding any necessary animations. The final product is then rendered and exported in FBX file format.

4.3 Creating Image Targets Using Vuforia Engine

In this application, as seen in Fig. 4, with the SDK, one can add advanced computer vision functionality to one's application, allowing it to recognize images, objects, and spaces with intuitive options to configure your app to interact with the real world. To enable Vuforia, access the Project Settings in Unity3D, and open the Player menu. Then, select the specific mobile device for which you are developing the app. In the XR Settings section, choose Vuforia from the list of available Virtual Reality SDKs, and check the Vuforia Augmented Reality Supported option. With the Vuforia Developer Portal, the target images are included in the Vuforia database. Thereafter, the database is downloaded and immediately uploaded to Unity.

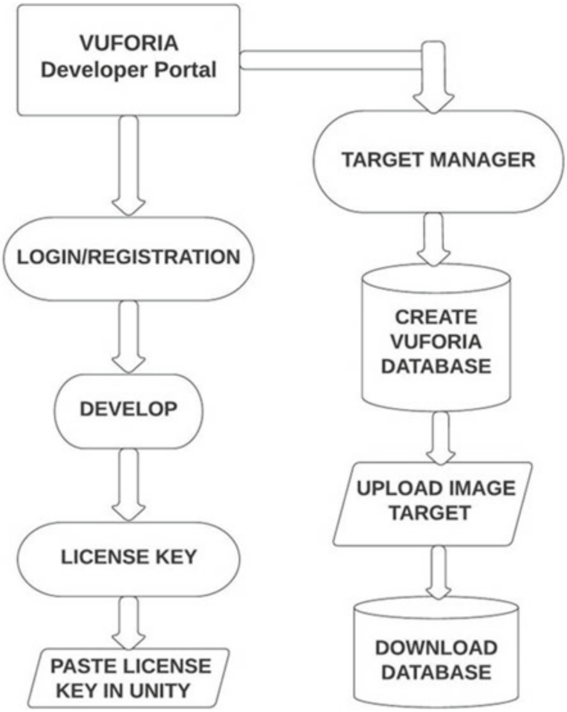


Fig. 4 Vuforia engine workflow

4.4 Integrate Through Android Studio

In this application, as seen in Fig. 5, the initial UI is designed with necessary buttons that enable the user to navigate through the entire application. The AR application is developed using Unity packages and the unity.apk is imported into the Android Studio and is assigned as a function for each button as a module. Thus, each module is integrated into the android application and it is exported as a single file incorporating all the applications.

4.5 Results and Discussion

The appropriate 3D model, which is preprogrammed to be placed on the plane as seen in Fig. 6, is activated once the picture target is scanned using the phone's camera. The respective word is then displayed on the canvas while its audio is playing in the background. This has been implemented for all 26 alphabets, which depict the word, audio, and the 3D model. As a demonstration, the letter O has been selected here, and when scanned with the phone's camera, a 3D representation of an owl and the word owl are presented on the screen.

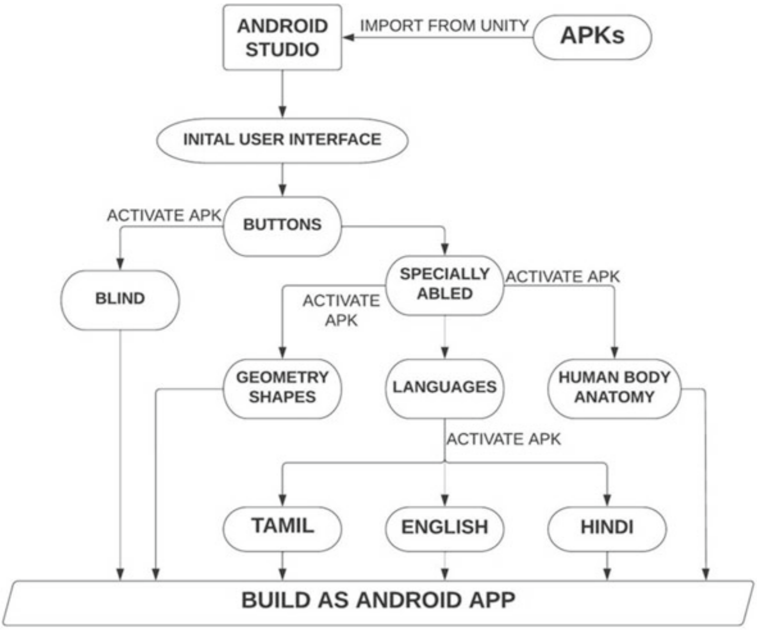
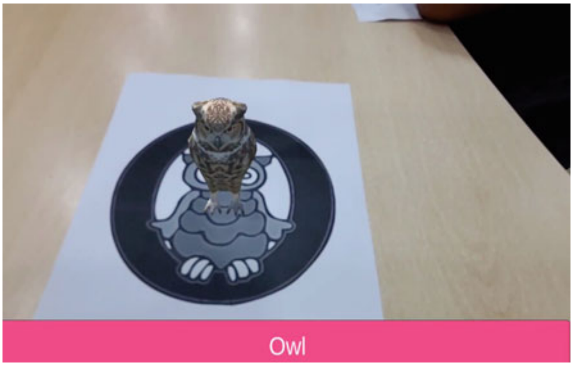


Fig. 5 Android studio workflow

Fig. 6 AR English



Similar approach is used to design other modules which include, AR Tamil, AR Hindi, AR Geometry, AR Numbers etc. added according to the user requirement as seen in Figs. 7, 8, 9 and 10.

Fig. 7 AR Hindi

Fig. 8 AR Numbers

Fig. 9 AR Tamil

5 Conclusion and Future Enhancement

After carefully examining the benefits and drawbacks of the literature review that has been conducted, an Augmented Reality-based Smart Education System for people with special needs is planned to be designed. The goal is to overcome the limitations

Fig. 10 AR Geometry

by developing an augmented reality application that focuses on both physically and mentally challenged individuals and is integrated with Android Studio to act as a single module that can be accessed by anyone, at any time, demonstrating no difference and, as a result, promoting equality in the field of education. This can create a revolution in the education sector for both normal and a special child promoting inclusivity in the long run.

References

1. Strange A (2021) This is how the augmented reality world of the metaverse could be opened to the visually impaired
2. Tan BL, Guan FY, Leung IMW (2022) A gamified augmented reality vocational training program for adults with intellectual and developmental disabilities: a pilot study on acceptability and effectiveness
3. Cieza E, Lujana D (2018) Educational mobile application of augmented reality based on markers to improve the learning of vowel usage and numbers for children of a kindergarten in Trujillo
4. Warnars HLHS, Nicholas N, Raihan M, Ramadhan A, Mantoro T, Adnan WAW (2021) Mobile application for the blind and their family
5. Hossain MF, Barman S, Haque AKMB (2019) Augmented reality for education; AR children's book
6. Quintero JH, Baldiris S, Rubira-García R, Chaves JR (2019) Augmented Reality in educational inclusion. A systematic review on the last decade
7. Quintero J, Baldris S, Velez G (2019) Augmented reality in educational inclusion. a systematic review on the last decade. Antretter T, Blohm I, Grichnik D, Wincent J (2019) Predicting new venture survival: a Twitter-based machine learning approach to measuring online legitimacy. J Bus Venturing Insights 11:e00109
8. Masmuzidin MZ, Aziz NAA, Suhaimi S (2022) A systematic review of the design of augmented reality applications for young children
9. Jdaitawi MT, Kan'an AF (2021) A decade of research on the effectiveness of augmented reality on students with special disability in higher education
10. Khan T, Johnston K, Ophoff J (2019) The impact of an augmented reality application on learning motivation of students

11. Sathya A, Priya L, Kumar P (2022) A virtual reality based interactive Android educational application. ECS Trans 107(1):18235–18245. https://doi.org/10.1149/10701.18235ecst

Intelligent Filter for Internet Traffic Monitoring and Analysis of Possible Cyber Threats

Hari Pranav Arun Kumar, J. S. Gokul, and K. N. Sudhakar

Abstract Internet and its services have attracted various domains to evolve over it. Gradually all devices are having Internet-dependent component integrated within its functionality. As Internet and its existence is public in nature, keeping the identity of the communicating device secure is one of the major challenges. In specific to Internet of Things if we are dealing with tiny components like sensors which are Internet driven they are highly prone to attacks from different threat mechanisms. To counter these attacks, we have many solutions which are cloud driven and in specific developed over time-bound subscription model. But they are not practical as it is very expensive and suited for large enterprises. Hence this paper proposes a cost-effective solution which can be used in home networks to help identify potential threats.

Keywords IDS · Suricata · Zeek · Docker · Elastic search

1 Introduction

With more devices being connected to the Internet via IoT they are prone to attacks from different threats. There are existing solutions on the cloud as a subscription-based model but they are not practical as it is very expensive and suited for large enterprises. Various home networks are prone to attacks without the proper configuration of security rules and firewalls. There are various advanced persistent threats being developed by nations everyday. Hence there is a need for a safe and secure system which alerts individuals and businesses, so that necessary steps can be taken. There are various papers with different methodologies and tools being used and we have evaluated them and come up with the best solution which is economical and efficient solution. To detect intrusion, possible techniques existing are signature-based intrusion detection (IDS) system and anomaly-based intrusion detection system (IDS) [1]. Among these existing systems, we have to decide which detection

H. P. A. Kumar · J. S. Gokul · K. N. Sudhakar (✉)
Department of ISE, CMR Institute of Technology, Bengaluru, India
e-mail: sudhakar.kn@cmrit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_11

engine best suites the need. Signature-based IDS works on historical identity of a threat described as signatures reflecting malware attack pattern, network scanning pattern, server penetration attempting pattern and many more. If such a pattern/s is observed in the live traffic, alert is triggered to notify the concern to take counter measures to tackle the situation. Anomaly-based IDS or behaviour-based IDS look into the network traffic abnormalities correlating it to a malicious attack pattern rather referring to payload delivered between the interacting nodes. Anomaly-based IDS tools depend on baselines, i.e. the unusual activity which deviates from existing statistical averages of previous activities referring to network traffic instead of signatures as mentioned in the earlier technique. For example if a network exists which has local working demography like a company and the incoming traffic symbolizes accessing HR records across globe is an indication of intrusion setting an alarm flag. If we try to compare the two techniques, fewer false positives occur with signature-based IDS but fail to detect new threats whose signatures are not know leaving the system vulnerable to attacks. Whereas the occurrence of false positives is pretty high in anomaly-based IDS. If the system is well configured, there are high chances of identifying previously unknown threats. Network-based IDS (NIDS) works on the principle of inspecting all the traffic crossing the networks. If such an activity is found at intermediate devices like switches or routers, a traffic mirror copy is received to NIDS devices for further analysis to threat detection.

2 Motivation

There are various advanced persistent threats being developed by numerous communities every day for varying selfish purposes. It is very difficult to trace the cause of the issue, and in most of cases it goes unnoticed for a long period of time until huge damage is registered. In various corporate companies there are teams dedicated to identifying the threats and mitigating issues. There are an increasing number of applications using user data as well as mobile payment applications. The money from user banks can be easily hacked and transferred only by knowing their phone number. The attacker can deploy a wide range of attacks such as spyware or man in the middle attacks and intercept the network traffic for determining the user details. User details such as WiFi passwords, router passwords, surveillance camera passwords and many more are having very low level of security. There are a number of phishing attacks and scams online which are spreading rapidly and can also be purchased at a cheaper rate on and online marketplaces like darknets. Community with selfish intentions see this as a potential area for criminal operations and cyber attacks. Attackers or hackers with distinct skills are seeing these vulnerabilities as an opportunity to provide hacking as a service. Hence there is a need for a safe and secure system which alerts society and businesses to take necessary steps to safe guard assets from fraudsters. In this regard numerous solutions and varied methodologies are proposed in which most of them rely on propitiatory tools. In this process we have come up with an optimal solution which is economical and efficient.

3 Ease of Use

3.1 Container-Based Solution

The solution is easy to deploy in every home network. This solution provided is a container-based solution, platform independent and persistent working well within the existing system configuration. Since the entire technology stack is bundled into a single container, it is quite easy to update and maintain to current versions by which the software defends itself with potential new vulnerabilities evolving in real time.

3.2 Comparative Case Study in Existing Technology

In the process of exploring existing open-source IDS systems, comparative analysis of different existing tools was done. In the process found the following pros and cons;

Snort Being the foremost open-source IDS system, it lacks in easy-to-use GUI administrative console for which we have to rely on other tools like BASE, Sguil and others.

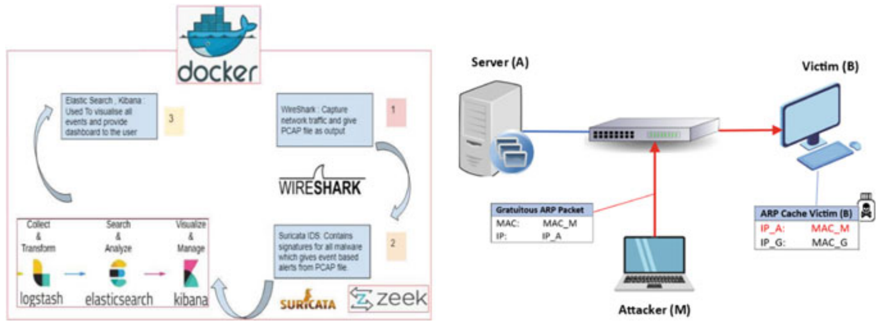
Suricata [2] Compares to others similar tools, the IDS engine is built around modern, multi-threaded, and highly scalable code base. Through PF_RING and AF_PACKET it provides native support of hardware acceleration for varied hardware vendors.

Zeek Provides support for both signature and anomaly-based IDS. Its analysis engine has the potential to convert captured traffic into a series of events. Whereas an event could be practically any request made over network.

In our implementation we have opted Suricata IDS along with Wireshark and the Elastic stack which is summarized in Fig. 1a.

3.3 (a) Tools Used (b) Implementation Architecture

WireShark In the list of network protocol analysers, Wireshark is the most primitive and formidable tool which provides a microscopic view of network traffic. It is widely used by the wider communities like government agencies, enterprises, educational institutions and the public.



(a) Tools used (b) Platform

Fig. 1 Working model and the architecture of implementation platform

Docker [2] Constituting a set of Platform as a Service (PaS) products that use OS-level virtualization to deliver software in packages called containers. It provides a platform to deploy and maintain software without worrying much about regular configuration changes.

Suricata The primary reason for using Suricata is its popularity among the user community. Being open source, it gives greater flexibility of usage to the user community.

Elasticsearch, Logstash, and Kibana (ELK) [4] These three tools blend well in analysing and visualizing network traffic. To know about them, Elasticsearch is a leading analytics engine to perform open searches to generate network search results. Logstash being open source is a pretty good server-side data processing pipeline that retrieves data from multiple sources at the same time and also transforms and passes it on to "stash" like Elasticsearch. Kibana is open source and a good tool for visualizing data in the form of charts and graphs in Elasticsearch [4].

To understand the working, the incoming and outgoing packets are captured from Wireshark [5] and the data is sent to the Suricata IDS, which contains the list of emerging threats definitions like viruses, malware, and worms. Later the data is sent to the ELK stack to analyse and find the various parameters such as country of origin, the severity of the threat, and the number of threats.

Later we visualize the threat patterns on a dashboard for the users to monitor and take necessary action against threats. This ensures that the user can easily mitigate and take necessary actions against the threats. The various logs are monitored, and the end user can know about the threats in the network. The visualized data can be used later to isolate the threat and as a defense towards law enforcement. This helps the user of the system stay safe from the ever-growing cyber threats.

3.4 Flow of Data

Following are the process steps for threat observations;

- The packets from the wireshark capture are taken, and it is sent to the Suricata IDS.
- In the Suricata IDS we have the emerging role threats containing signatures for different viruses and worms. Once the pcap file is sent to Suricata we compare this with the already existing threats.
- This data has to be visualized; hence we used the ELK stack and the data must be sent from the ids to the ELK stack; hence we ensure that the end user is aware of the different threats.

3.5 Detection of Attack on the Data Link Layer

1. ARP Poisoning: The following are the steps for a successful attack taking Fig. 1b as reference:

- M would pretend to be B to A: It will send a gratuitous ARP reply with the pair: IP_B->MAC_M
- M would pretend to be A to B: it will send a gratuitous ARP reply with the pair: IP_A->MAC_M Because of the TTL in hosts ARP caches, an attacker would need to send these packets at intervals lower than the timeout (usually every 30 s is a good choice) as shown in Fig. 1b.

2. Detection of ARP on the Stack Detection of ARP on the Stack by analysing the mac_flood pcap file as shown Fig. 2a. From Fig. 2b, we can see that the host is Kali machine.

From Fig. 2b we can find under the alert signature we can see that Suricata can detect the different signatures.

(a) (b)

Fig. 2 **a** Host identification, **b** alert signatures

From Fig. 3a under the alert category generated we can see that see that the IDS can identify that there is potentially bad traffic. In Fig. 3 we can see the different fields and observations.

Hence in the visualization, we can determine that the host is a Kali machine trying to perform MITM attacks which can steal the data in transit.

3.6 Detection of Attack on the Transport Layer

If there are devices running scans using tools such as NMAP it can be identified as there are excessive SYN packets being sent without completing the three-way handshake.

Also if there is usage of different flags we can identify the irregularities and finally, if a single host is sending traffic such as SYN flag to multiple hosts we can flag it as an anomaly. More specifically, Nmap sends a series of SYN packets to an open port and tries to perform OS fingerprinting, by attempting to identify the Initial Sequence Number (ISN) generation formula in the returned SYN/ACK and comparing the returned values to a file that contains expected returned values for a series of checks analysing the nmap_sequence number prediction.pcap.

Scanning by a single IP sends multiple SYN flags. Here we can see that the conn_state index of ZEEK IDS identifies the state of the connection.

Analysing strange_scanning.pcap: Obviously, we are dealing with a scanning attempt, but what is that destination port 0 (TCP) all about is to be known. A host is not supposed to be listening on port 0. A scanning tool sending packets to port 0 (TCP) of the remote host is trying to identify if the remote host is alive. If the remote host is alive, then, a RST response should be sent to all those packets. As shown in Fig. b we can see that under the history it is Sr and the id.resp_p is 0. conn_state is REJ which says that the connection is rejected.

Figure 4a Messages, Fig. 4b conn_state index of ZEEK IDS, Fig. 4c gives the total count of requests which signifies that there are multiple connection requests sent but was aborted by the originator. Figure 4d shows us that from the IP address 192.168.153.154, there are over 800 requests which indicate a port scanning in the network. Hence from the various response messages we can determine that there is an attacker scanning a network to determine the various ports open.

3.7 Detection of Attack on the IP Layer

The IP layer's functions are related to how packets are transferred from one hop to another. Source and destination IP addresses are used by the IP layer for inter-host communication. IP addresses reside in the IP header of the IP packet. It should be noted that IP packets travel individually as they are directed to their destination. For example, IP packets from the same source that are traveling to the same destination could get there through different routes.

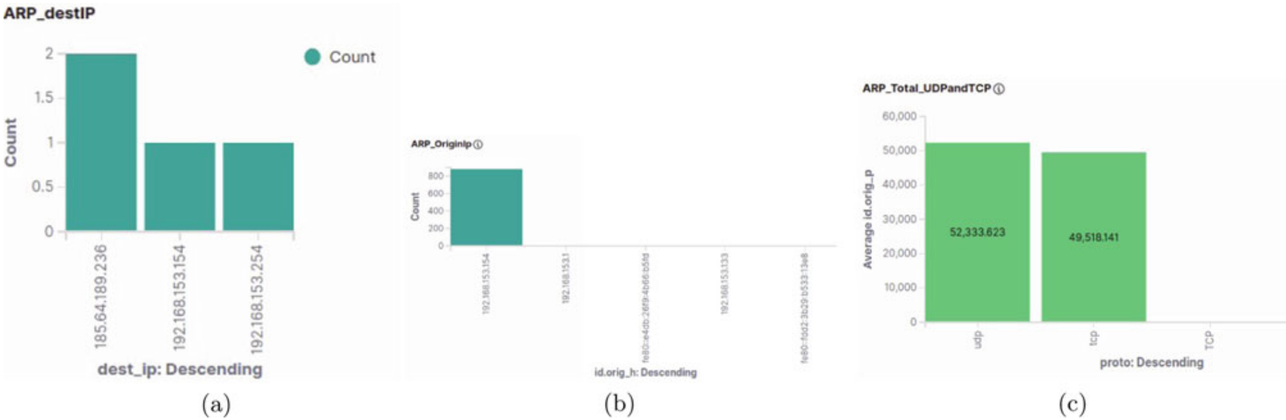


Fig. 3 **a** Depicts the destination IP versus count of the packets, **b** depicts the number of ARP requests sent which is abnormally high, **c** depicts protocol UDP which is higher than TCP which carries the ICMP message protocol

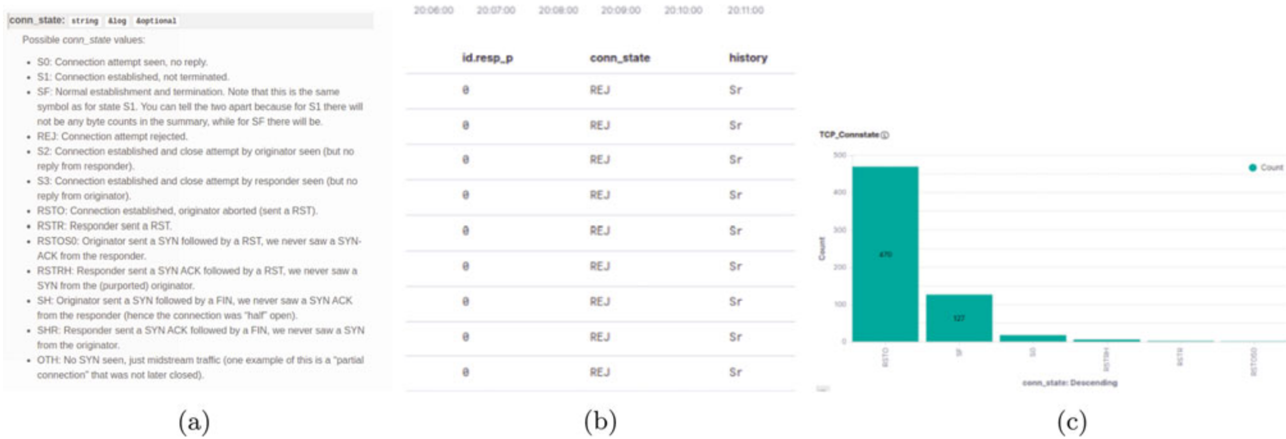
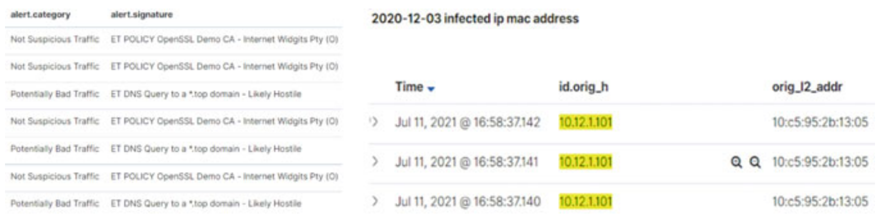


Fig. 4 **a** Messages, **b** conn_state index of ZEEK IDS, **c** gives the total count of requests which signifies that there are multiple connection requests sent but were aborted by the originator, **d** shows us that from the IP address 192.168.153.154, there are over 800 requests which indicate a port scanning in the network



(a) (b)

Fig. 5 **a** Depicts facts to help distinguish normal and suspicious HTTP traffic, **b** represents some facts to help distinguish normal and suspicious HTTPS traffic

During packet delivery, a lot of things can go wrong. The IP layer has no built-in mechanism to identify when a packet gets lost, expired, or dropped. The Transport protocol or the application itself is responsible for identifying and resolving any packet loss.
We analyse the pcap file by passing it to Suricata which has a set of emerging threat rules to identify the malicious activity taking place inside the packet and find the infected IP address.
After getting the infected IP we are next going to focus on finding the hostname, domain name, mac address, and username of the infected IP address.

Hence using these visualizations we can determine that there has been an attack from the IP 10.12.1.101.

3.8 Detection of Attack on the Application Layer

Analysing 2020-12-03-traffic-analysis.pcap, we analyse the pcap file by passing it to Suricata which has a set of emerging threat rules to identify the malicious activity taking place inside the packet and find the infected IP address.

1. HTTP Protocol HTTP traffic consists of a series of requests and responses known as messages. Client will make a request and server will respond. HTTP responses include a 3-digit status code. HTTP messages include a message header and body (Figs. 5, 6 and 7).

2. HTTPS (Hypertext Transfer Protocol Secure) HTTPS is the secure version of HTTP.
The “S” refers to Secure Socket Layer / Transport Layer Security (SSL / TLS). HTTPS also establishes a handshake similar to TCP but more complicated. Below is a brief summary: Both the client and the server need to agree on the protocol version. Both the client and the server need to select cryptography algorithms.

Optionally authenticate to each other. Use public key encryption techniques to establish secure communications.

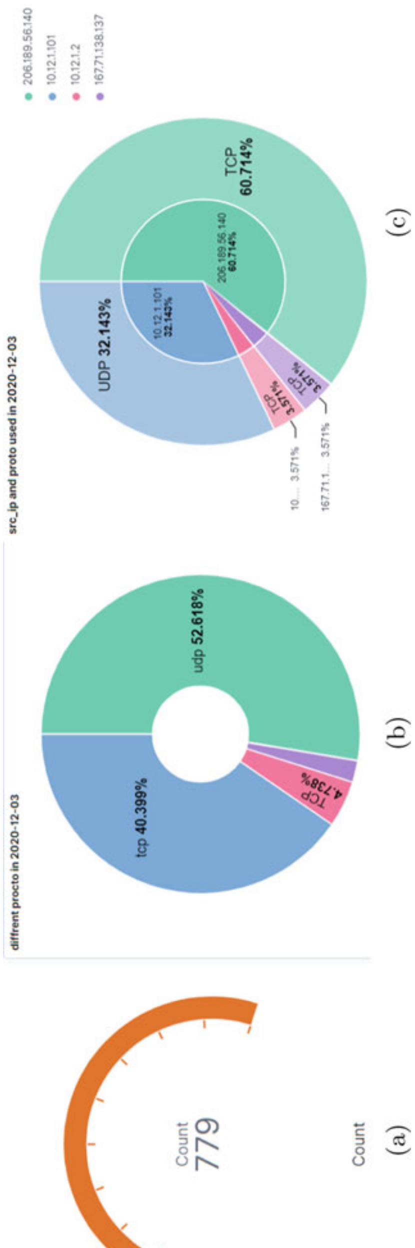


Fig. 6 Depicts the Conn_States

(a)

Normal HTTP Traffic	Suspicious HTTP Traffic
Port 80, TCP Port 8080, TCP (used as alternate) Port 8088, TCP (used as alternate)	Malicious binaries (backdoors), scripts, web shells, etc. will use this port because typically in all corporate environments the port is open.
Plaintext traffic	If the traffic is encrypted then most likely it's being used for malicious traffic. Malicious traffic can be in plaintext as well.
Web server typically in FQDN format.	Server will point to an IP address instead of FQDN format.

(b)

Normal HTTPS Traffic	Suspicious HTTPS Traffic
Port 443, TCP Port 8443, TCP (used as alternate)	Malicious binaries (backdoors), scripts, web shells, etc. will use this port because typically in all corporate environments the port is open.
Encrypted traffic	If the traffic is not encrypted and Secure Sockets Layer packet details are empty within packet details then that will fall under suspicious.
Web server typically in FQDN format.	Server will point to an IP address instead of FQDN format.

(c)

(d)

(e)

Fig. 7 **a** Depicts facts to help distinguish normal and suspicious HTTP traffic, **b** represents some facts to help distinguish normal and suspicious HTTPS traffic, **c** depicts the Conn_State, **d** depicts the Conn_State, **e** depicts the threats in wireshark which is Client Hello which is abnormal

4 Conclusion

The increasing affordability of electronic devices has made them accessible to the general public, enabling the widespread use of applications for various daily needs. However, this convenience comes with risks, as evidenced by instances of malware infecting devices such as smart fridges and posing threats to users' networks. The complexity and cost of current cloud-based solutions and antivirus software make them difficult for average users to implement and maintain.

This project proposes a user-friendly solution that provides a high level of abstraction through easy-to-read dashboards and alerts. By collecting and analysing network data using tools like Wireshark and intrusion detection systems, the solution enables users to identify and respond to potential threats effectively. The deployment process is simple and quick, thanks to the use of a docker container. Additionally, the solution benefits from regularly updated signature detection for identifying malicious traffic. Overall, this project offers an affordable and manageable approach to cybersecurity for the average user, allowing them to understand and protect their network without requiring in-depth technical knowledge. The generated dashboards can also be shared with cybercrime departments for further analysis and utilized by machine learning algorithms to identify attack patterns and predict new threats.

5 Future Scope

Running on high-specification hardware for faster analysis, distributed hardware allows us to run multiple copies of the data in parallel for analysis, and this makes the solution highly fault-tolerant in nature. REST API integration with other messaging services allows real-time messaging to be sent to the end user whenever there is an anomaly in the network, this reduces the time for the analysis and response to the threat, and also various applications can have a common platform to analyse this data. Elastic stack upgradability causes many issues, and this can be countered using Docker. The solution can be improved further by deploying on a Kubernetes cluster and managing from various services such as Terraform to check, maintain and automate the deployment. The usage of machine learning algorithms can be done to help automate and improve the efficiency of the detection algorithms. Various supervised learning methods such as neural networks and decision trees can be used to further increase the chances of reducing the false negatives during the detection phase.

References

1. Khraisat A, Gondal I, Vamplew P et al (2019) Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity 2:20. https://doi.org/10.1186/s42400-019-0038-7
2. Ren J, Guo J, Qian W, Yuan H, Hao X, Jingjing H, Building an effective intrusion detection system by using hybrid data optimization based on machine learning algorithms

3. White JS, Fitsimmons T, Matthews J, Quantitative analysis of intrusion detection systems: snort and Suricata
4. Patgiri R, Varshney U, Akutota T, Kunde R, An investigation on intrusion detection system using machine learning. Jan SU, Ahmed S, Shakhov V, Koo I, Toward a lightweight intrusion detection system for the Internet of Things
5. Kumar A, Maurya HC, Misra R, A research paper on hybrid intrusion detection system

Detection of Autism Using Artificial Intelligence

Soumabha Mitra, Srinath K., V. Gowri Manohari, D. Poornima, and K. Karunya

Abstract Autism Spectrum Disorders are developmental disorders that affect children and get worse with age. People with autism have trouble interacting and communicating with others. They also have problems with restricted and repetitive behaviors and interests. People with autism have different ways of learning to move and pay attention to problems. Early diagnosis can improve quality of life through care and treatment, but in many developed countries, it is too late to diagnose children with autism. Autism is very difficult for trained medical professionals to diagnose because there is no direct medical test. As a result, children require intensive supervision, which takes longer to detect, which in many cases may not be possible. Moreover, given the existing stigma surrounding mental health in many parts of society, many children remain unrecognized until it is too late. This project will give you the most accurate results and reduce the time it takes to observe your children. This experiment used computer vision for child eye tracking. In our model, we used this technique to improve detection accuracy and reduce detection time. Eye tracking helps detect autism. The experiment is divided into several sections to provide both autism detection. In this experiment, we used CNN. This is because it is more accurate than any other available technique. This experiment will help medical professionals reduce the time needed for diagnosis.

Keywords Autism diagnostic observation schedule · Early detection · Eye tracking · Detection · CNN

S. Mitra (✉) · K. Srinath · V. Gowri Manohari · D. Poornima · K. Karunya
Sathyabama Institute of Science and Technology, Chennai, India
e-mail: msoumabha@gmail.com

K. Srinath
e-mail: srinath.ka16@gmail.com

V. Gowri Manohari
e-mail: gowrimanohari.cse@sathyabama.ac.in

D. Poornima
e-mail: poornima.d.cse@sathyabama.ac.in

K. Karunya
e-mail: karunsai.ifs@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_12

1 Introduction

1.1 Autism Spectrum Disorder (ASD)

Communication and conduct are impacted by the developmental disease known as autism. Early childhood is when it is frequently identified, but there is no one test for it, so it might be challenging to find. Instead, to make a diagnosis of autism, clinicians use behavioral observations, developmental evaluations, and accounts from parents and carers. Machine learning algorithms are now being used by researchers to identify autism. These algorithms may be able to spot autistic symptoms that are hard for human observers to spot by analyzing patterns in behavioral data such as language usage and social interactions. Machine learning algorithms have been proven in several studies to be useful in diagnosing autism, but further study is required to evaluate their accuracy and reliability. It is essential to stress that a thorough evaluation by a licensed physician is still necessary and that machine learning algorithms should not be utilized as the only technique of diagnosis.

1.2 Signs of Autism Spectrum Disorder (ASD)

The symptoms described here suggest a possible case of autism spectrum disorder. Individuals with autism may exhibit inappropriate laughing and giggling, and may not be sensitive to pain. They may struggle to make eye contact properly, and may not respond appropriately to sounds. They may not have a wish for cuddling or be able to express their gestures effectively. Additionally, individuals with autism may not interact with others or show appropriate attachment to people, instead displaying attachment to inappropriate objects.

1.3 Autism Spectrum Disorder in Adults

ADS is difficult to detect in adults: Difficulty interpreting social cues and understanding social norms Difficulty initiating or maintaining social interactions Difficulty communicating and expressing needs and wants or obsessive behaviors or interesting sensory processing, e.g. It is important to note that the severity of signs and symptoms varies greatly between individuals with ASD and not all. ASD experiences all these symptoms. Diagnosing ASD in adults requires a comprehensive examination by a licensed physician. The preceding signs and symptoms are not always present in autistic persons, and they may also encounter additional ones not listed above. The indications and symptoms of ASD differ from person to person, while there may be some overlap with other conditions such as attention deficit hyperactivity disorder (ADHD). Additionally, the symptoms may vary across genders. Because their symptoms may be more covert and milder, some persons may appear to be able to handle social settings better than others. As a result, diagnosing ASD might be more difficult.

1.4 Autism Symptoms in Adults at Work

Depending on how severe the illness is, ASD symptoms can vary widely from person to person. ASD symptoms, or ones identical to them, can be noticeable at work. Your co-workers claim that you have robotic speech. You don't like it when cleaning services reduce the items on your desk to dust since each one has a particular spot. You excel at math and computer programming but struggle in other subjects. Speak to your co-workers in the same manner as you would with your relatives and friends. Unintentionally produces noise during meetings, such as repeatedly cleaning one's throat. It might be tough to discern while speaking with your manager if they are pleased or upset with your performance. You may showcase your skills. Additionally, 40% of those with autism have IQ levels that are ordinary or above average.

1.5 Diagnosing ADS

ASD cannot be diagnosed with a single test; instead, reports from parents or other carers together with observations of a person's behavior and developmental progress are used to make the diagnosis. A health care provider may look at a person's behavior, speech, and social interactions as well as their growth and medical background to make an ASD diagnosis. To evaluate a person's social and communication abilities, health professionals may also administer standardized exams and fill out questionnaires.

It is crucial to remember that ASD can be challenging to identify and that an accurate diagnosis necessitates a thorough assessment by a trained medical professional. For those with ASD, early diagnosis and intervention are advantageous. Because of this, it's critical to visit a doctor if you have concerns about a loved one or for yourself.

1.6 Living with Autism

Autism may make daily life challenging for both sufferers and their families. Autism spectrum disorders can cause difficulties with social interaction, communication, and situational adjustment. They can also cause repetitive or obsessive behaviors and interests. Living with autism requires a customized strategy since each person is affected by autism in their own unique way. A person with autism and her family may, however, use a variety of coping mechanisms and tools to manage the condition and have productive lives.

1.7 Treatment for Autism

- Early Intervention: People with autism benefit from early diagnosis and treatment because they may better address developmental delays and achieve better results.
- Professional Therapy: A range of treatments, such as occupational, speech, and behavioral therapy, aid in the skill development and symptom management of persons with autism.
- Supportive education helps persons with autism flourish academically and socially by providing special education programmers and accommodations in conventional schools.
- Community Resources: People with autism and their families have access to a wide variety of organizations and resources. Self-help organizations, summer camps, and leisure activities.
- Individualized Treatment Plan: It's critical to collaborate with your healthcare provider to create a treatment strategy that is suited to your particular requirements and objectives.
- People with autism and their families may achieve their full potential and lead satisfying lives by locating the appropriate services and support.

2 Literature Survey

One of the main approaches used in the literature is to use artificial intelligence algorithms to classify individuals as having ASD based on various characteristics. These characteristics may include behavioral indicators such as social interaction and communication skills, and brain imaging data such as functional magnetic resonance imaging (fMRI) and electroencephalography (EEG).

Many studies have reported good results using artificial intelligence techniques for ASD detection.

[1] A study confirmed that it was possible to classify children with an accuracy of 85% by classifying children with autism and children with normal development using an artificial intelligence algorithm based on eye tracking data lower.

[2] The research focuses on leveraging artificial intelligence (AI) technologies to enhance the screening and diagnosis of autism spectrum disorder (ASD). Traditionally, ASD detection relies on behavioral observations, but this study aims to introduce a more objective, data-driven approach. The literature review explores various studies that have incorporated AI in the assessment process, aiming to discern if additional behavioral data could be utilized to differentiate ASD characteristics.

[3] A. R. Mohamed Shanavas, M. S. Mythili. We used categorization algorithms to perform an ASD investigation. The primary goal of this article was to identify autistic difficulties and autism levels. SVM and fuzzy algorithms were used in this neural network, coupled with WEKA tool, to improve student conduct and social interactions and to investigate human relationships.

[4] D.P. Wallet Al., Me. Kosmicki1, V. Sochat, M. Duda, and D.P. Wallet Al. It detects autism using a simple feature set search approach. The authors employed an artificial intelligence technique to examine the clinical evaluation of ASD in this way. ADOS was administered to a behavioral subgroup of children on the autism spectrum. The ADOS has four modules. In this study, eight different artificial intelligence methods were utilized. On 4540 scoresheets, this entails progressive backward feature detection. We detect ASDrisk using 9 of the 28 behaviors in module 2 and 12 of the 28 behaviors in module 3 with an overall accuracy of 98.27% and 97.66%, respectively.

[5] Di Martino, Yan C.G, Li Q, DenioE, Castellanos F. X, Alaerts K & Milham M. P. (2014). The Autism Brain Imaging Data Exchange: Toward a large-scale evaluation of the intrinsic brain architecture in autism. Molecular psychiatry, 19(6), 659–667.

[6] This research paper explores the use of artificial intelligence (AI) methods in the automatic detection of Autism Spectrum Disorder (ASD) through MRI neuroimaging. The study reviews and discusses various approaches to leveraging AI for this purpose. The authors highlight the potential of AI algorithms in enhancing the accuracy and efficiency of ASD diagnosis using neuroimaging data. The paper was published in Frontiers in Molecular Neuroscience in October 2022.

[7] The study analyzed placental tissue from preterm births, comparing 10 infants with Autism Spectrum Disorder (ASD) to 10 controls. Using AI, they identified 4,870 significant CpG loci related to ASD. Deep Learning accurately predicted ASD with 100% accuracy using 5 specific CpG markers. The research also pinpointed five prenatal pathways linked to ASD.

[8] A research paper, published in Biomedicines in June 2023, surveys the utilization of Artificial Intelligence (AI) in diagnosing Autism Spectrum Disorder (ASD) using advanced neuroimaging techniques like Diffusion Tensor Imaging (DTI) and functional Magnetic Resonance Imaging (fMRI). The authors, Eman Helmy, Ahmed Elnakib, Yaser ElNakieb, Mohamed Khudri, Mostafa Abdelrahim, Jawad Yousaf, Mohammed Ghazal, Sohail Contractor, Gregory Neal Barnes, and Ayman El-Baz, explore the potential of AI as a tool for enhancing ASD diagnosis, providing valuable insights into this evolving field. The paper sheds light on the intersection of AI, neuroimaging, and ASD diagnosis.

Some of the studies have reported good results using these approaches, with classification accuracies ranging from 80 to 87%. These papers discuss the use of various artificial intelligence techniques, such as support vector machines, decision trees, and neural networks. Artificial intelligence algorithms are only as good as the data they are trained on, and more research is needed to determine the accuracy of these approaches.

Overall, the literature suggests that artificial intelligence may become a useful tool for ASD detection and improve the accuracy and efficiency of ASD diagnosis. However, it is important to note that artificial intelligence algorithms are only as good as the data they are trained on, and more research is needed to determine the generalizability and robustness of these approaches.

3 Proposed Methodology

3.1 Steps in Detecting Autism

Data collection: Gathering a dataset of individuals with and without autism is the initial stage. This dataset needs to include details on the diagnosis of autism as well as behavioral, physiological, and/or brain imaging data. Data Preprocessing: Data Preprocessing is the next phase. This involves preparing the data for machine learning by cleaning, standardizing, and changing it. Extraction of important characteristics from data that might reveal the existence of autism is known as feature extraction. The most crucial characteristics should be identified in this stage using methods like feature selection and dimensionality reduction. Model selection: Depending on the data type to address, pick the best machine learning model. Models like decision trees, random forests, support vector machines, and convolution neural networks can be included in this. Training and Validation: Using preprocessed data to train a model and validation data to assess the model's performance. To increase performance, you might need to fine-tune model settings now. Regardless of the training and validation sets, test the model on the test dataset to assess how well it generalizes. The model is ready to be deployed in a real environment once it has been trained and verified. Applied in hospitals or clinics. Monitor: Keep a close eye on the model's performance and retrain it as necessary.

3.2 Convolution Neural Network

Convolutional neural networks (CNNs) are a family of deep learning algorithms that are frequently employed for the analysis of photos and videos, but they may also be applied to the analysis of other forms of data, including voice and text. Many studies have employed CNNs to diagnose autism using brain imaging data such as magnetic resonance imaging (MRI) and functional MRI because they are particularly well suited for image classification tasks (fMRI). In these investigations, in order to find patterns and characteristics connected to autism, researchers employed CNNs to map the anatomical and functional brains of people with and without autism examined the pictures. CNNs are frequently used to categorize fresh pictures as autistic or not after being trained on big databases of brain scans. For instance, he discovered that the network could accurately classify people as autistic or not in one research that utilized CNNs to evaluate structural MRI pictures of the brains of children with and without autism was discovered to be accurate at classifying. In another research, CNNs were employed to analyze brain fMRI images taken while children with and without autism performed a social cognition test to pinpoint the brain areas connected to the disorder.

3.3 CNN Architecture to Detect Autism

The convolutional neural network architecture known as VGG16 was introduced by the Visual Geometry Group (VGG) at the University of Oxford in their 2014 study titled "Very Deep Convolutional Networks for Large-Scale Image Recognition". It is renowned for having a straightforward design and a high degree of accuracy for picture classification applications. VGG16 is trained on the ImageNet data set and has 16 layers, thus the name VGG16. Multiple convolutional and max pooling layers precede fully linked layers in the architecture. It may be adjusted for use in different computer vision applications and is frequently used as a pre-trained model for picture classification tasks.

The model's initial component is made up of several convolutional layers. These layers examine various aspects of the image and spot patterns. A variable filter size is used to analyze the picture in each convolutional layer. This enables the model to identify various patterns across the picture. The second component of the model is made up of several totally interconnected layers. Based on the patterns that the convolutional layers have identified, these layers are utilized to categorize pictures. Finally, output the probabilities for each class using a SoftMax layer. This enables the model to determine the class to which a picture belongs.

3.4 VGG16 to Detect Autism

Numerous applications involving the identification of autism, such as the detection and diagnosis of autism spectrum diseases, have made use of VGG16 (ASD). It has been used to identify facial expressions, posture, hand gestures, and other autism-related behavioral indicators. VGG16 has also been used to detect items in photos, assisting in the discovery of environmental patterns that might be connected to autism. In order to recognize items in the movie, we ultimately utilized VGG16. This aids in monitoring autistic people's conduct.

3.5 Dataset Description

The dataset used in this study was obtained from Kaggle.com and consisted of 2450 training images. The images were divided into two categories: 'Autistic' and 'Non-Autistic', with 1225 images in each category. To prevent any bias in the results, the

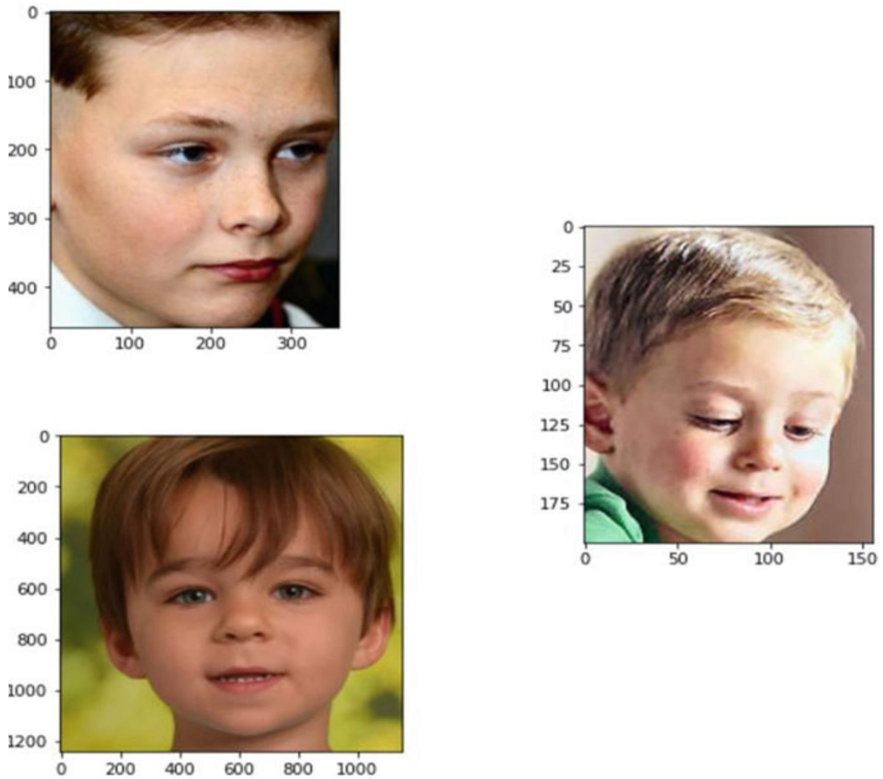


Fig. 1 Image description of dataset

images were randomly shuffled before being used for training. The balanced distribution of images between the two classes ensures that the dataset is representative of both groups (Fig. 1).

3.6 Data Preprocessing

In the given task, we are dealing with a dataset of train images that are of different dimensions. To prepare this dataset for model training, we first redefined the dimensions of the images to make them uniform. This was achieved by resizing the images uniformly before feeding them to our model. To implement this, we wrote a function that reads the images, resizes them, and specifies the labels for each image based on the file name. Specifically, we can label an image as ‘1’ if it belongs to the ‘Autistic’ class and ‘0’ if it belongs to the ‘non-Autistic’ class. Once we have implemented this function, we check the number of images and their labels to ensure that everything

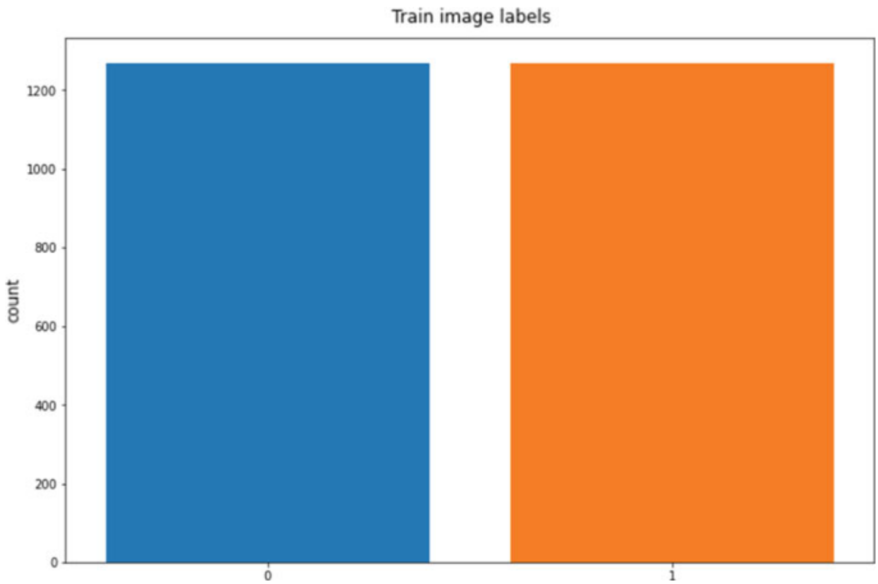


Fig. 2 Visualization to check the number of autistic and non-autistic image

has been done correctly. The resulting dataset can then be used for model training (Fig. 2).

This visualization shows that equal number of Autistic and Non-Autistic data values are present. After resizing the images and labels for our training data, we can convert them into NumPy arrays for ease of use in our machine learning models. A NumPy array is a powerful data structure that allows for efficient manipulation of large datasets, which is particularly important when working with image data.

4 Results and Discussion

4.1 Pre-trained Model for Transfer Learning

For image classification tasks, we utilize a pre-trained Convolutional Neural Network (CNN) model called VGG16, which has been trained on a vast amount of image data. The base model of VGG16 comprises 19 layers, including 13 convolutional and 6 fully connected layers.

4.2 Creation of Custom Model

To connect a pre-trained convolutional neural network (CNN) model to a custom fully connected layer, we can use transfer learning. This involves using the pre-trained weights from the CNN model as the initial weights for our custom fully connected layer (Fig. 3).

This allows us to take advantage of the pre-trained model's ability to extract features from images, while still allowing us to train our own specific classification layer. When it comes to training the model, we need to specify the loss function, optimizer, and metrics. For binary classification, we can use 'binary-cross entropy' as the loss function. The optimizer and learning rate can be selected through hyper-parameter tuning to reduce the loss. To prevent overfitting, we can use an image data generator to perform various image transformations on the available training images. This can include scaling the image pixel array between 0 and 1, as well as performing zoom and flip operations to increase the variety of the training data. By increasing the amount of training data, we can help ensure that the model does not overfit by training on a small dataset. It is important to note that we should also perform scaling on the validation images, but we should not augment the images in the validation dataset. The purpose of the validation dataset is to test the performance of our model on data that it has not seen before. Therefore, it is important that the validation data is representative of the real-world data that the model will encounter.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vgg16 (Functional)           (None, 4, 4, 512)         14714688
_________________________________________________________________
flatten (Flatten)            (None, 8192)              0
_________________________________________________________________
dense (Dense)                (None, 512)               4194816
_________________________________________________________________
dropout (Dropout)            (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 513
=================================================================
Total params: 18,910,017
Trainable params: 4,195,329
Non-trainable params: 14,714,688
_________________________________________________________________
```

Fig. 3 Pre-trained VGG16 model

4.3 Model Training

To build a machine learning model, it is important to first train it on a set of training data. Once the model is trained, it is then evaluated on an unseen set of validation data to measure its performance, ensuring that machine learning models can make accurate predictions on new, unseen data is crucial. During the training phase, it's important to set the number of epochs, which refers to one complete pass through the training dataset. The number of epochs required depends on the complexity of the data and the model architecture. After training, the model's performance is evaluated on validation data to detect overfitting or underfitting. Overfitting occurs when the model performs well on the training data but poorly on the validation data, indicating that it has simply memorized the training data instead of learning from it. Conversely, underfitting occurs when the model performs poorly on both the training and validation data, suggesting that it's too simple to capture the underlying patterns in the data (Fig. 4).

The learning curves of a machine learning model are important visualizations that help us understand how the model is performing during training. Typically, we plot the training loss and validation loss over the number of epochs to observe their behavior. In a well-trained model, the training loss should decrease as the number of epochs increases. This indicates that the model is effectively learning from the training data. However, the validation loss can show some noise around the training loss. This is because the validation set is typically smaller than the training set, which

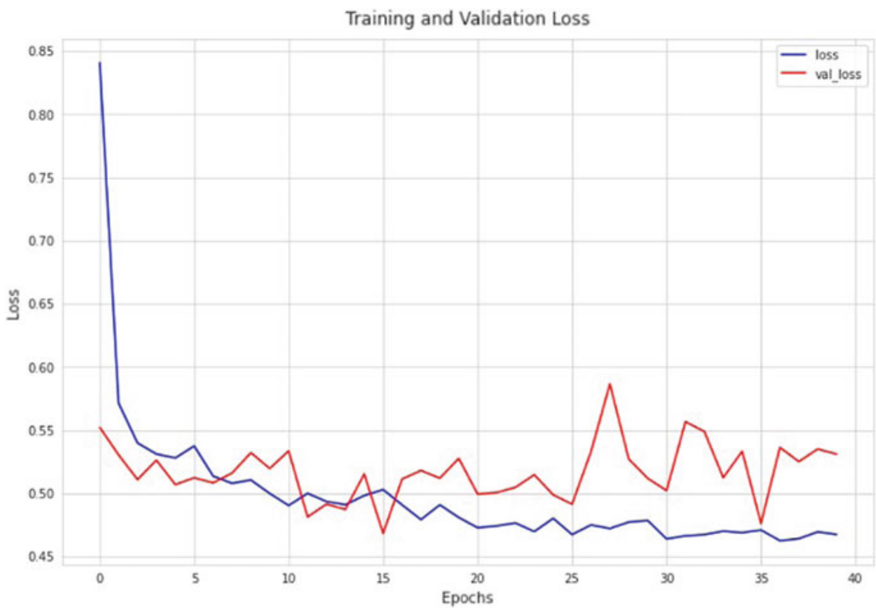


Fig. 4 Training and validation loss graph

can lead to some randomness in the validation loss. In some cases, we might also observe that the validation loss tends to increase after a certain number of epochs, say around 15–17 epochs. This is an indication that the model is overfitting to the training data. Overfitting occurs when the model becomes too specialized to the training data, leading to poor generalization to new, unseen data. Moreover, the validation accuracy can also fluctuate a lot. This is expected, especially when the validation set is small, and it might not necessarily indicate a problem with the model. However, it is important to take note of the general trend of the validation accuracy over the number of epochs.

4.4 Model Validation

The Receiver Operating Characteristic (ROC) curve's Area Under the Curve (AUC) is a common metric for assessing the effectiveness of classification models, especially in datasets with balanced class distribution. AUC measures how well the model can differentiate between positive and negative classes, with higher values indicating better performance. An AUC of 1 denotes a perfect classifier, while an AUC of 0.5 suggests a random classifier. However, datasets in real-world scenarios are often imbalanced, and accuracy alone is insufficient for evaluating model performance. In such cases, precision and recall become crucial metrics as they consider False Positive (Type 1) and False Negative (Type 2) errors. Precision calculates the proportion of true positive predictions among all positive predictions, while recall calculates the proportion of true positive predictions among all actual positive instances. Therefore, to compare model performance, we must evaluate multiple metrics such as precision, recall, and AUC. The classification report summarizes these metrics based on the model's predictive power for both positive and negative classes (Fig. 5).

4.5 Results

Based on the evaluation metrics, the accuracy of our model is 73.33%. Although this may seem like a reasonable score, there is still room for improvement. One factor that could be contributing to the modest accuracy is the size of our training data. As the data set is small, the model may not have had enough examples to fully capture the complexity of the underlying patterns in the data. Therefore, increasing the size of the training data could potentially lead to a more accurate model. Furthermore, we can experiment with different techniques such as data augmentation, regularization, and hyperparameter tuning to further improve the model's performance. Data augmentation is a process of creating new data by applying various transformations, such as rotation, scaling, and flipping, to the existing dataset. Regularization involves imposing constraints on the model to prevent overfitting, which can help in developing a more generalized model. Hyperparameter tuning is another technique that

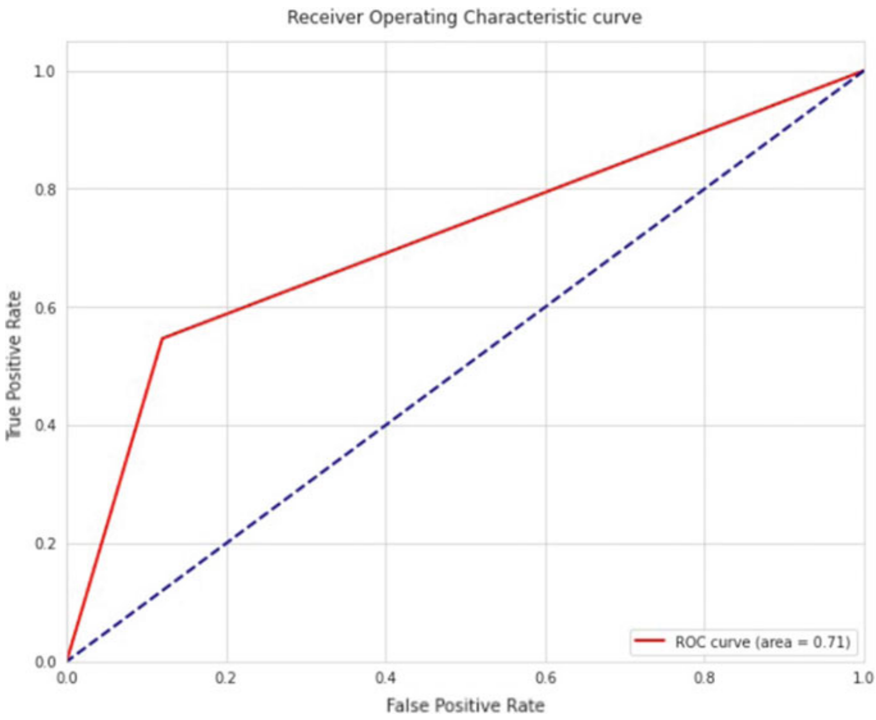


Fig. 5 ROC curve

can be employed to optimize the model's parameters and improve its performance. Moreover, we could explore advanced algorithms or architectures to enhance the model's accuracy. For instance, we could consider using deep learning techniques, like convolutional neural networks or recurrent neural networks that are specifically designed to capture complex patterns in the data (Fig. 6).

5 Improvisation

Providing a large dataset for training a machine learning model can help it to generalize better on new data. This is because the model can learn more diverse and representative patterns from a larger amount of data, and therefore make better predictions on unseen examples. However, simply providing more data is not always sufficient to improve model performance, as the quality and diversity of the data also play a crucial role. To prevent overfitting and evaluate model performance, it is common practice to split the data into training, validation, and testing sets. However, with limited data, the validation set may be too small to provide a reliable estimate of model performance. In such cases, cross-validation can be used to split the data

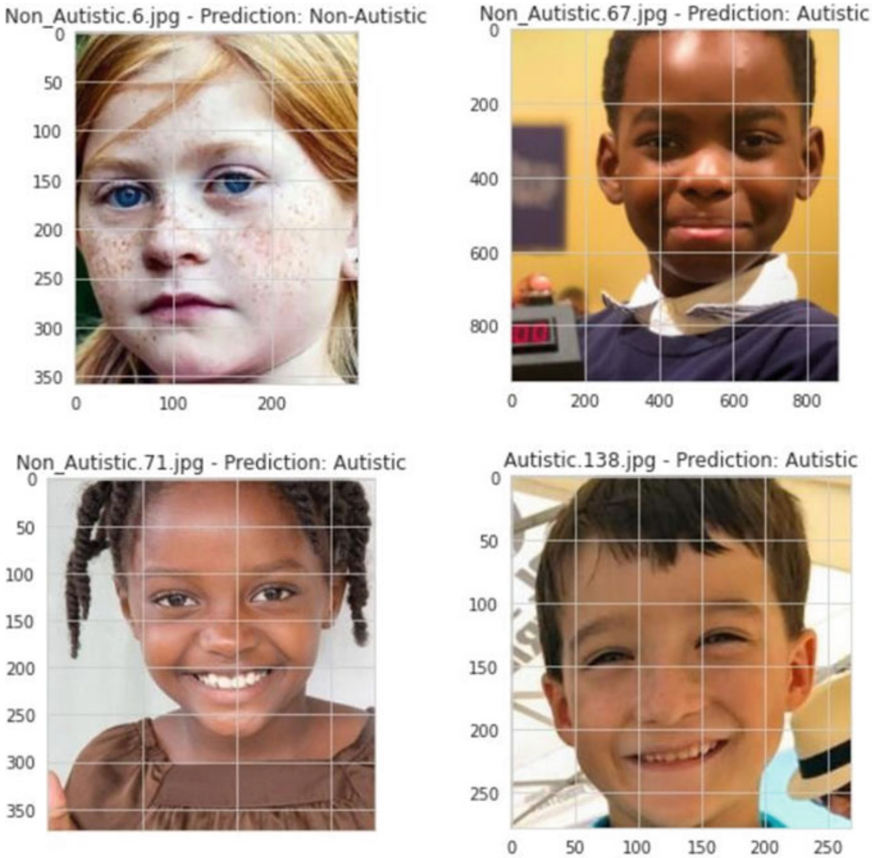


Fig. 6 Image with their prediction and actual value

into multiple folds and perform multiple training and validation cycles, which can help to better estimate the model's performance. In addition to training on a large dataset and using cross-validation, another strategy to improve model performance is fine-tuning. Fine-tuning involves taking a pre-trained model, such as VGG16, and retraining some of its top layers on a new dataset to customize the feature extraction for a specific task. This can help the model to learn task-specific patterns and improve its performance on the new dataset. Fine-tuning can be especially useful when the new dataset is smaller than the original pre-training dataset, as it allows the model to leverage the knowledge learned from the pre-training on a smaller dataset.

References

1. Zhao Z, Tang H, Zhang X, Qu X, Hu X, Lu J (2021) Classification of children with autism and typical development using eye-tracking data from face-to-face conversations: machine learning model development and performance evaluation. J Med Internet Res 23(8):e29328. https://doi.org/10.2196/29328. PMID: 34435957; PMCID: PMC8440949
2. Song DY, Kim SY, Bong G, Kim JM, Yoo HJ (2019) The use of artificial intelligence in screening and diagnosis of autism spectrum disorder: a literature review. Soa Chongsonyon Chongsin Uihak 30(4):145–152. https://doi.org/10.5765/jkacap.190027. PMID: 32595335; PMCID: PMC7298904
3. Mythili MS, Mohamed Shanavas AR (2014) A study on autism spectrum disorders using classification techniques. Int J Soft Comput Eng (IJSCE) 4(5):88. Published By: Blue Eyes Intelligence Engineering & Sciences Publication Pvt. Ltd. ISSN: 2231-2307. Retrieval Number: E2433114514/2014©BEIESP
4. Kosmicki JA, Sochat V, Duda M, Wall DP (2015) Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Transl Psychiatry 5(2):e514. https://doi.org/10.1038/tp.2015.7. PMID: 25710120; PMCID: PMC4445756
5. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, Anderson JS, Assaf M, Bookheimer SY, Dapretto M, Deen B, Delmonte S, Dinstein I, Ertl-Wagner B, Fair DA, Gallagher L, Kennedy DP, Keown CL, Keysers C, Lainhart JE, Lord C, Luna B, Menon V, Minshew NJ, Monk CS, Mueller S, Müller RA, Nebel MB, Nigg JT, O'Hearn K, Pelphrey KA, Peltier SJ, Rudie JD, Sunaert S, Thioux M, Tyszka JM, Uddin LQ, Verhoeven JS, Wenderoth N, Wiggins JL, Mostofsky SH, Milham MP (2014) The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry 19(6):659–67. https://doi.org/10.1038/mp.2013.78. Epub 2013 Jun 18. PMID: 23774715; PMCID: PMC4162310
6. Moridian P, Ghassemi N, Jafari M, Salloum-Asfar S, Sadeghi D, Khodatars M, Shoeibi A, Khosravi A, Ling SH, Subasi A, Alizadehsani R, Gorriz JM, Abdulla SA, Acharya UR (2022) Automatic autism spectrum disorder detection using artificial intelligence methods with MRI neuroimaging: a review. Front Mol Neurosci 15:999605. https://doi.org/10.3389/fnmol.2022.999605. PMID: 36267703; PMCID: PMC9577321
7. Bahado-Singh RO, Vishweswaraiah S, Aydas B, Radhakrishna U (2022) Artificial intelligence and placental DNA methylation: newborn prediction and molecular mechanisms of autism in preterm children. J Matern Fetal Neonatal Med 35(25):8150–8159. https://doi.org/10.1080/14767058.2021.1963704. Epub 2021 Aug 17. PMID: 34404318
8. Helmy E, Elnakib A, ElNakieb Y, Khudri M, Abdelrahim M, Yousaf J, Ghazal M, Contractor S, Barnes GN, El-Baz A (2023) Role of artificial intelligence for autism diagnosis using DTI and fMRI: a survey. Biomedicines 11(7):1858. https://doi.org/10.3390/biomedicines11071858

Analyzing and Predicting Diabetes Chronic Disease Using Machine Learning Techniques

K. Risheedhar Varma, Amarnath Pulagam, V. Gowri Manohari, P. Nandini, and J. Joan Niveda

Abstract The incidence of diabetes has been on a steady rise for several decades, becoming a significant global health concern. This condition impairs the body's ability to utilize blood glucose effectively, often resulting from poor dietary habits, lack of exercise, aging, or other underlying medical issues. Without proper care and attention, diabetes can result in severe health consequences such as renal failure, heart attacks, strokes, eye damage, blindness, kidney failure, and other organ damage. However, with early detection and appropriate intervention, this condition can be managed effectively. To improve the accuracy of predicting diabetes in its early stages, we have utilized a range of machine learning techniques. By leveraging datasets directly collected from patients, machine learning models have shown significant improvements in prediction outcomes. Our research study employs a variety of ensemble algorithms, including Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic Regression (LR), and Gradient Boosting (GB) to predict diabetes in each dataset. By comparing the accuracy of various models, we have identified the most precise machine learning algorithm. Furthermore, we have integrated our model with a web application using Flask Framework to enable easy and efficient prediction of diabetes. Our study emphasizes that machine learning-based classification and a proper pre-processing pipeline for clinical data can effectively predict diabetes and potentially aid in early intervention. In conclusion, our

K. Risheedhar Varma (✉) · A. Pulagam · V. Gowri Manohari · P. Nandini · J. Joan Niveda
Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India
e-mail: risheedharvarma@gmail.com

A. Pulagam
e-mail: Pulagamamarnath7@gmail.com

V. Gowri Manohari
e-mail: gowrimanohari.cse@sathyabama.ac.in

P. Nandini
e-mail: nandini.p.cse@sathyabama.ac.in

J. Joan Niveda
e-mail: joanniveda.cse@sathyabama.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_13

research shows promising results and encourages further development of machine learning models for diabetes prediction.

Keywords K-Nearest Neighbor · (KNN) Logistic Regression (LR) · Support Vector Machine (SVM) · Gradient Boosting (GB) · Flask Framework

1 Introduction

A chronic autoimmune condition called diabetes type 1 alters how the body makes and utilizes insulin. Type 1 diabetes is not brought on by lifestyle decisions, in contrast, to type 2 diabetes, which is linked to lifestyle factors. Instead, it is brought on by the immune system targeting and decimating pancreatic cells that make insulin, which is required for controlling blood sugar levels.

Type 1 diabetes is characterized by weight loss, tiredness, and decreased eyesight, increased thirst, and urination, intense hunger. Type 1 diabetes can cause major problems such as heart disease, nerve damage, kidney damage, and visual loss if it is not managed. Insulin therapy is often used as part of type 1 diabetes treatment. It's crucial to remember that type 1 diabetes cannot be cured but may be controlled with the right care and medication. Working closely with their medical team, remaining informed about their illness, and taking action to control their blood sugar levels are all things that people with type 1 can do.

Type 2 diabetes is like a complex puzzle with many interlocking pieces. Some pieces are determined by our genes, while others are shaped by our lifestyle choices. Unhealthy habits such as overeating, sitting for long periods, and consuming sugary, processed foods can all play a role in the development of this disease. Fortunately, there are ways to solve this puzzle and reverse its effects. By adopting healthy lifestyle changes like exercising regularly, eating nourishing foods, and shedding extra pounds, we can improve our insulin sensitivity and give our pancreas the support it needs to function properly. To ensure normal blood sugar levels are maintained, some individuals may require the use of medications or insulin therapy. While type 1 diabetes is a challenge that cannot be prevented, type 2 diabetes is a puzzle that can be solved with the right tools and mindset. By taking control of our health and making positive changes, we can prevent or delay this disease and live our best lives possible. Although many individuals find controlling their diabetes challenging, living a long and healthy life with diabetes is feasible with the right care and treatment. Complications can be avoided or delayed by maintaining a healthy lifestyle, checking blood sugar levels, and scheduling routine doctor's appointments. Diabetes education is essential because those with the disease must comprehend the significance of controlling their blood sugar levels, taking their medications as directed, and altering their lifestyles. It is crucial to remember that diabetes is a complicated condition that includes several metabolic abnormalities. With the right treatment, people with diabetes may have active, healthy lives. Maintaining vigilance for possible threats and taking preventative measures is critical for limiting their impact.

Several decision support systems are provided by many information mining algorithms to aid healthcare practitioners. The decision support system's accuracy demonstrates its effectiveness. As a result, the objective is to build a decision support system capable of precisely predicting and diagnosing a certain circumstance. To anticipate the onset of diabetes, several machine-learning methods have been developed. We use K-Nearest Neighbor (KNN), Gradient Boosting (GB), Logistic Regression (LR), and Support Vector Machine (SVM) in our work. The K-Nearest Neighbors (KNN) algorithm works by locating the k-closest data elements to a new point and predicting the new point based on the average value of those points. Support Vector Machine (SVM) is a popular machine learning technique employed mainly for solving classification and regression problems. SVM's primary objective is to find the suitable boundary, or "decision boundary," between various classes of data. Gradient Boosting (GB) is a machine learning strategy that joins several ineffective models, to produce a more robust and accurate model. It operates by training several models, each of which aims to fix the mistakes caused by the model before it. The ultimate model, which combines all of these models, can predict outcomes more accurately than any of the separate models could. Logistic regression (LR) is a type of supervised learning method used to solve binary classification problems. By considering one or more predictor variables, LR estimates the probability of the binary outcome variable while imposing certain constraints and the probability estimates between 0 and 1 using a logistic function. Using Flask Framework, we further link the model with a web application. The findings of this work suggest that diabetes may be predicted successfully and efficiently using ML-based classification and a suitable pre-processing pipeline for clinical data.

2 Related Works

VijiyaKumar [1] pioneered the Random Forest technique for developing a model that accurately predicts early diabetes in a patient. The study found that the Random Forest model was able to accurately predict diabetes in patients with an accuracy of 76.3%. This result is significant since early detection of diabetes can lead to early interventions and improved patient outcomes. The study also identified the most important features for predicting diabetes, with body mass index and age being the most influential factors. Nonso Nnamoko [2] presented Diabetes Onset Forecasting: An Ensemble Supervised Learning Method. According to the findings, ensemble models frequently outperform individual constituent classifiers. This is not given because a variety of factors, particularly during fundamental training, may limit their potential to enhance performance. Deepti [3] evaluated the efficiency with which the classification algorithms such as Decision Tree, SVM, and Naive Bayes were executed. Although using the Pima Diabetes Database (PIDD) in their research, the greatest accuracy of their work was only about 76.3%. The obtained dataset was then employed in the planned research attempt, which aimed for an 80% success rate. Dehghani [4] performed a comprehensive evaluation of 21 papers that employed machine learning algorithms to predict diabetes. Upon analysis, it was discovered that the commonly utilized methods for the task were decision trees, support vector machines, and neural networks. The datasets utilized, the characteristics chosen, and the assessment measures vary throughout the investigations. The scientists came to the conclusion that machine learning algorithms have a lot of potential for predicting diabetes, but the results need to be confirmed using bigger and more diverse datasets. Deeraj Shetty [5] proposed creating an Intelligent Diabetes Disease Prediction System that uses a diabetes patient database to analyze diabetic sickness. Using data mining, this program anticipates diabetes disease. They recommend using Bayesian and KNN (K-Nearest Neighbor) algorithms to a database of diabetic patients and analyzing them by taking numerous aspects of diabetes into consideration to predict the onset of the chronic disease. Glmnet, RF, XGBoost, and LightGBM were among the regression models examined by Kopitar [6] to forecast the occurrence of type 2 diabetes mellitus. The objective of this research was to evaluate and contrast the efficacy or effectiveness of novel machine learning algorithms with traditional regression techniques. The core aim of this research was to enable early forecasting of impaired fast glucose and fasting plasma glucose levels (FPGL) using innovative and distinct methods. Zhang [7] presented a detailed overview of the application of machine learning algorithms for diabetes prediction in this review study. They addressed the uses of several machine learning approaches, such as support vector machines, decision trees, and random forests, in diabetes prediction. They also examined much research that employed machine learning algorithms to predict diabetes and evaluated their strengths and weaknesses. Using a dataset of 500 individuals, Almedia [8] performed a comparison research to examine the efficacy of many machine learning approaches for predicting diabetes. The gradient boosting algorithm was determined to be the most accurate, with an AUC of 0.91. According to Ayush Anand's [9] research, the CART prediction model may be implemented with 75% accuracy to a highly classified dataset. Blood pressure has been established as a significant risk factor for diabetes. In this study, Wang [10] proposed a paradigm that utilizes machine learning to predict diabetes. The framework employs a feature selection strategy to minimize the dimensionality of the data and increase prediction accuracy. They tested the framework on a dataset of 5000 diabetic patients and attained an accuracy of 82.5%. The authors also compared their framework's performance to that of other cutting-edge machine learning algorithms and discovered that their framework surpassed others in terms of accuracy and computing efficiency.

3 Dataset

The origins of this remarkable dataset can be traced back to the esteemed National Institute of Diabetes and Digestive and Kidney Diseases. The aim of this dataset is to facilitate accurate diagnostic predictions of diabetes in patients, based on a series of diagnostic measurements that are encompassed within the dataset. To ensure the highest quality of data, rigorous selection criteria were applied to a much larger

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

Fig. 1 Dataset sample

database from which this dataset was derived. This dataset exclusively features female patients of Pima Indian descent who have reached the age of 21 years or above. Within this dataset, a multitude of medical predictor variables have been included, alongside a single target variable, known as Outcome. The predictor variables cover a broad spectrum of health metrics, ranging from the number of pregnancies a patient has had, to their age, insulin levels, BMI, and beyond (Figs. 1 and 2).

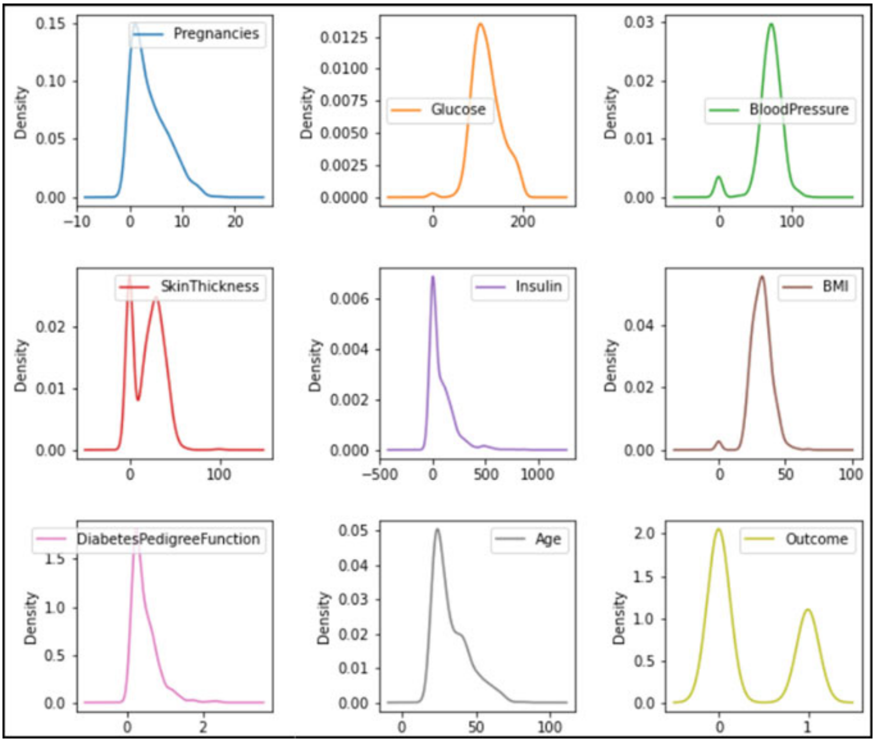


Fig. 2 Univariate density plots

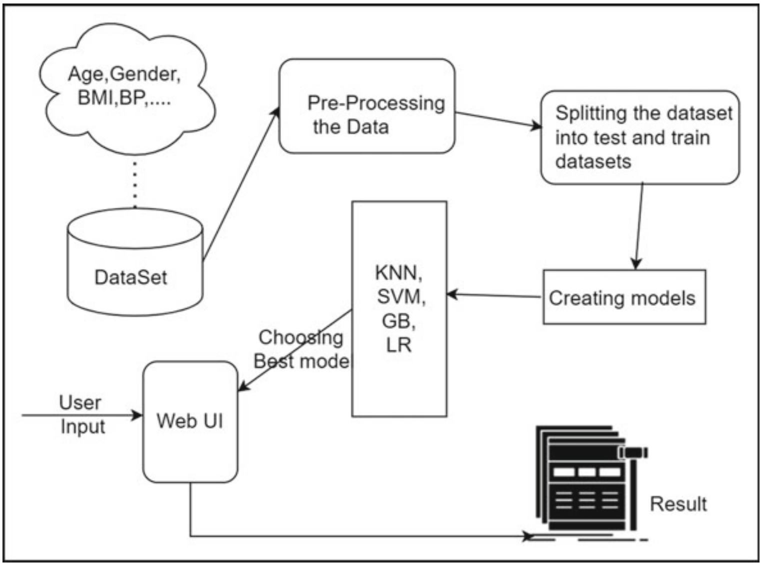


Fig. 3 Workflow

4 Methodology

See (Fig. 3).

4.1 Pre-processing the Data

Data pre-processing is like the secret ingredient that gives your machine learning models the extra boost they need to succeed. It involves the careful cleaning, transforming, and organizing of raw data to ensure that it's ready for analysis. Preparing this dataset for analysis requires a range of techniques, such as handling missing values, tackling outliers, scaling features, encoding categorical variables, selecting features, and splitting the dataset. These steps are essential for building robust and accurate machine learning models that can predict the onset of diabetes with precision. In machine learning, splitting a dataset into training and testing sets is an important step. The dataset is divided into two parts, where the test size parameter determines the proportion of the data used for testing, typically set to 20%, leaving 80% for training. The model is then trained on the training set, and its parameters are adjusted to minimize errors. The performance of the model is evaluated using the testing set, with the aim of developing a model that can make accurate predictions on new data.

4.2 Algorithms Used

As part of our research, we have investigated a range of machine learning algorithms, consisting of KNN (k-Nearest Neighbors), SVM (Support Vector Machines), GB (Gradient Boosting), and LR (Logistic Regression).

K-Nearest Neighbors (KNN): calculates the similarity between a new data point and existing data points to assign the new point to a category or predict a value. In classification tasks, the algorithm assigns the new point to the class that is most common among its K-Nearest Neighbors in the feature space. In regression tasks, it takes the average of the values of the K-Nearest Neighbors to predict the value of the new data point. KNN is a basic supervised learning algorithm that considers the similarity between the new and existing data points.

Logistic Regression (LR): Logistic Regression is a powerful tool in the world of machine learning that excels in classifying data. Its ability to model the complex relationship between input variables and output probabilities is unmatched. By using a logistic function to transform the input variables into a linear combination, Logistic Regression can predict which class the output variable belongs to with impressive accuracy. This algorithm's versatility is demonstrated by its capability to tackle both binary and multiclass classification problems, providing interpretable and straightforward results that are easily understood by users.

Support Vector Machine (SVM): SVM is a proficient algorithm in the realm of machine learning that effectively segregates data into distinct classes by identifying the most appropriate hyperplane with the maximum margin. The margin corresponds to the space between the hyperplane and the nearest data points from each class. To handle both linear and non-linear data, SVM transforms it into a higher-dimensional space, thereby enhancing its efficacy.

Gradient Boosting (GB): Gradient Boosting (GB) is a machine learning strategy that combines multiple weak prediction models, often decision trees, to generate a final, more accurate prediction. Its primary objective is to sequentially integrate models that aim to address the errors of their predecessors. GB accomplishes this by minimizing the loss function during each iteration, typically using the mean squared error for regression issues and the cross-entropy loss for classification tasks. One of the remarkable aspects of GB is its exceptional precision, along with its ability to withstand noise and outliers, thus making it a widely popular approach for data scientists and machine learning enthusiasts.

5 Results and Discussion

5.1 Model Performance

The performance of machine learning algorithms is influenced by the characteristics and size of the dataset, with each algorithm having its strengths and limitations. Proper evaluation of classification models is essential in machine learning, and various metrics are available for this purpose. These metrics include accuracy, precision, recall, F1-score, and ROC-AUC score. Accuracy measures the ratio of correct predictions to the total number of predictions made by the model, considering both true positives and true negatives. Precision, on the other hand, measures the proportion of true positives among all positive predictions, providing insight into the model's ability to identify positive instances accurately. Recall, also known as sensitivity, assesses the proportion of true positives correctly identified by the model. The F1-score provides a single score that balances both precision and recall by calculating the harmonic mean of the two. Conversely, the ROC-AUC score evaluates the model's ability to differentiate between positive and negative classes, determined by plotting the true positive rate against the false positive rate at different threshold levels. By examining these metrics, we can determine the effectiveness of the classification model and guide the algorithm's selection for a particular dataset, ultimately improving the overall performance of the model (Fig. 4).

In this study, we presented a machine learning model for diabetes prediction using a dataset containing various demographic, clinical, and lifestyle factors of patients. The SVM model achieved an overall accuracy of 0.76 on the test dataset, indicating its potential for accurate diabetes prediction. Here, the goal is to achieve high accuracy,

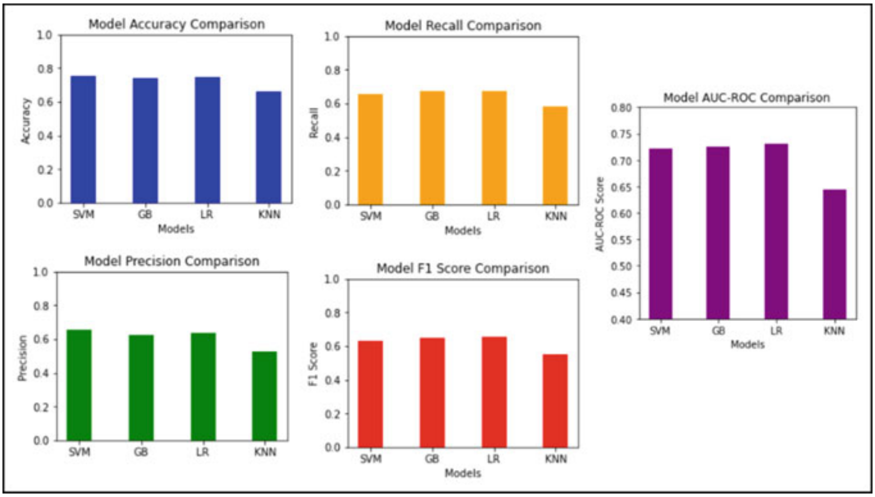


Fig. 4 Models performances

so SVM is the best choice as it has the highest accuracy in the table. On the other hand, if the goal is to minimize false positives, then precision may be the most important metric, and, in this case, also SVM has higher precision scores than the other models. If the goal is to detect all positive instances with high recall, then Logistic Regression may be the best choice as it has the highest recall score.

5.2 *Web-UI Implementation*

The final step of the project involves making a web application that can accept attribute values like BMI, blood pressure, glucose levels, and family history of diabetes and test the trained model to predict the output. First, you'll need to create a user interface for the web application, allowing users to input their attribute values. Once the user clicks the "test" button, the application will need to send these values to the trained model for prediction. We used SVM for training the model which is accurate and precise. To deploy the trained model in the web application, you can use a Flask's Development server for testing. This will allow you to run the application on a local server, making it accessible for testing. Finally, it's important to test the web application thoroughly before making it available to users. This includes testing the user interface, the connection to the trained model, and the security measures. Once everything is working properly, you can launch the application and make it available to users (Figs. 5 and 6).

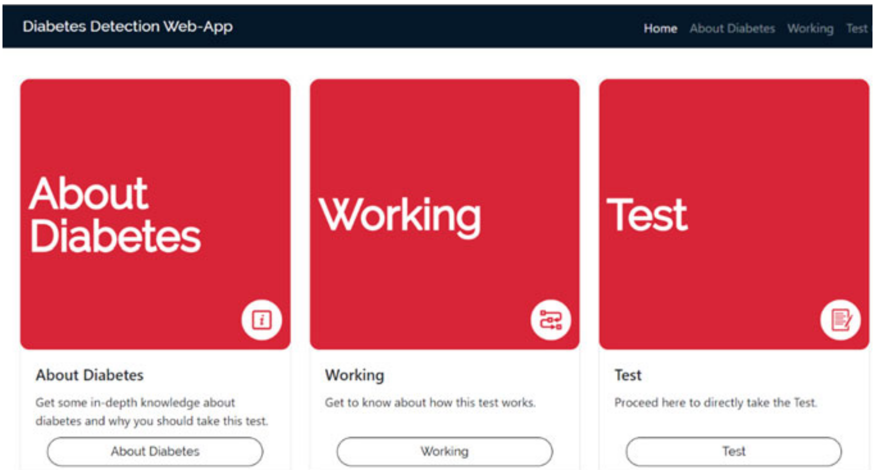


Fig. 5 Home page of the UI

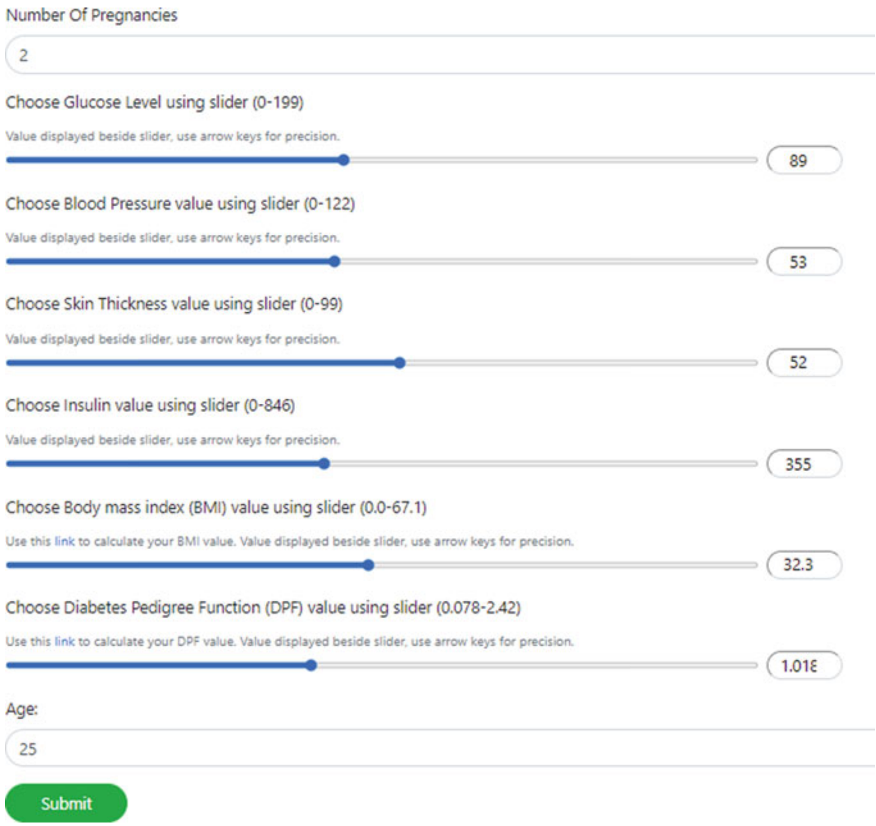


Fig. 6 User input for prediction

6 Conclusion

In conclusion, we discovered that SVM is the most effective machine learning method out of the rest. We assessed the accuracy of the predictions before applying the specific method to the project, and the results were excellent. By using machine learning models for diabetes prediction, it is possible to identify persons at risk for diabetes and aid them in adopting preventative actions. It is crucial to keep in mind that these models should be verified using sizable and varied datasets and that the outcomes should be evaluated in light of other elements like a patient's medical history and physical exam. In the future, there is still space for progress. Incorporating extra information, such as genetic and lifestyle characteristics, into the models is one possible direction for future study. Incorporating deep learning strategies, such as neural networks, may also help the models perform better. Overall, using machine learning models to predict diabetes has the potential to significantly help with early diagnosis and disease prevention. The accuracy and efficiency of these models must be continually improved via research and development in this area.

References

1. VijiyaKumar K, Lavanya B, Nirmala I, Caroline SS (2019) Random forest algorithm for the prediction of diabetes. In: 2019 IEEE international conference on system, computation, automation and networking (ICSCAN). Pondicherry, India
2. Nnamoko N, Hussain A, England D (2018) Predicting diabetes onset: an ensemble supervised learning approach. IEEE Congr Evol Comput (CEC)
3. Sisodia D, Sisodia DS (2018) Prediction of diabetes using classification algorithms. Int Conf Comput Intell Data Sci (ICCIDS2018)
4. Dehghani M, Saghaei A, Amini M (2020) Prediction of diabetes using machine learning algorithms: a systematic review. J Diabetes Metab Disord 19(1):35–46
5. Shetty D, Rit K, Shaikh S, Patil N (2016) Diabetes disease prediction using data mining. Int Conf Innov Inf Embed Commun Syst (ICIIECS)
6. Kopitar L, Kocbek P, Cilar L, Sheikh A, Stiglic G (2020) Early detection of type 2 diabetes mellitus using machine learning-based prediction models. Sci Rep 10:1–12
7. Zhang Y, Liu X, Wang L, Xiong J (2019) A comprehensive review of machine learning for diabetes prediction. J Healthc Eng 2019:1–15
8. Almeida TA, de Sousa RM, de Albuquerque VHC (2019) A comparative study of machine learning techniques for diabetes prediction. In 2019 IEEE international conference on systems, man and cybernetics (SMC) pp 1196–1201
9. Anand A, Shakti D (2015) Prediction of diabetes based on personal lifestyle indicators. In: 1st International conference on next generation computing technologies, 978-1-4673-6809-4
10. Wang Q, Zhou Y, Chen X, Chen H (2015) A machine learning-based framework for diabetes prediction. IEEE Trans Industr Inf 11(6):1418–1425

Challenges of Multimedia Education in the Lives of Human Beings

V. Parvathy and Deepanjali Mishra

Abstract The role played by multimedia in today's world is manifold and cannot be ignored. Multimedia in education can be defined as the use of text, audio, video, graphics and animation while imparting lessons in a classroom. The use of multimedia in classrooms can help the students in a big way if teachers are able to combine their pedagogical skill with modern technology. Multimedia aids the students in grasping information quickly by offering them a visual presentation of the subject rather than creating an image by themselves in their minds. Multimedia offers a better interactive platform than lecturing in the classrooms. The use of multimedia in classrooms depends on many factors like availability of technological tools, well trained teachers, the student size and the teacher's mettle in handling multimedia for making improvement in the academic performance of his or her students. Even though multimedia can make classroom teaching live by enhancing student participation and enthusiasm, it also faces several challenges. Technology-aided education especially through power point presentations is very common in developed and developing countries but receives cold response in underdeveloped countries owing to financial constraints. Apathy of authorities, lack of proper training for teachers, a huge class size and insufficient awareness of modern ways of teaching methodologies are some among the major reasons for the lack of multimedia-enabled education in many schools and colleges in various parts of the world. The effective implementation of multimedia in education can make educationists design curriculum in accordance with world standards along with students getting greatly benefitted by exploring new vistas of knowledge and academics at par with global best practices. The introduction of ICT in teaching through smart classrooms with the use of LCD and internet connection has the potential to attract girls from low-income families to come to school on a regular basis and can aid the respective countries in empowering their female population through education, skill development and generation of employment. Computer science is a field that is progressing at a fast pace and it is one among the most important reasons why ICT is being introduced in education in educational institutions in different parts of the world. One can see that with developments in technology, maximum number of jobs are generated. Multimedia-aided education

V. Parvathy · D. Mishra (✉)
School of Humanities, KIIT University, Bhubaneswar, India
e-mail: Deepanjalimishra2008@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_14

helps the learners to be more competent, think more widely, helps them in skill development and perform tasks according to set standards. Hence, it is of paramount importance for the governments, NGOs and other social and political organizations of every nation to empower the youth by imparting them technology-based education though multimedia in order to make them capable of performing various jobs demanding multimedia skills. These jobs include not only the ones in information technology but also in other fields of study too with computers and networking already making an unquestionable place for themselves everywhere. This research paper focuses on the need of multimedia-aided education in the present times and lays bare the practical challenges faced in educating students with the help of multimedia in the present day scenario.

Keywords Multimedia · Teaching · Development · Youth · Social media · Challenges

1 Introduction

The age in which we live can be better defined as the age of information and technology and we can see it making an imprint in every field of study, jobs and business. Technology now plays an inevitable role in education mainly through multimedia. In the field of education, the methodology of teaching has undergone a sea change with respect to the teacher teaching the students, the way the students perceive and develop the concepts in their minds and the interaction in classrooms between teacher and students and the interpersonal communication among students. It is beyond doubt that multimedia plays a huge role in the dissemination of education making breakthroughs in the quality of the content generated with the help of information technology that can give one access to any piece of information at one's fingertips. Although there are several positive effects of multimedia learning, one must not forget the disadvantages caused by it. Multimedia teaching is often welcomed as interactive and lively but it can also have a very bad impact on the brains of the young. As a result, students face psychological problems like anxiety, addiction, isolation from their peer groups and even attention deficit disorders. With the onset of the social media platforms, visual media, YouTube and OTT platforms like Netflix, Disney Hotstar, amazon prime video etc., the use of technology and networking has become a part of our daily especially those of youth and students. It can be pointed out that technology has obviously made people connect with their loved ones, share their views and to get experience on different subjects without time lag even though they are on a different part of the globe. But it's high time that we discuss, chalk out plans and find solutions for the same connectivity harming students, diverting them from studies and misguiding them even to be a part of anti-social and criminal activities due to overuse or abuse of multimedia. We must realize that multimedia is a powerful tool capable of making great changes in the way people think, talk and act and that its importance is only going to increase in the years to come. What should be recommended is a careful, time bound and judicious use of multimedia with the supervision of parents which can be of great use to students in improving their academic record and learning how to use internet and networking keeping in mind their own safety and security.

2 Advantages of Multimedia

The use of multimedia in education has its own advantages and disadvantages. Some of the advantages of multimedia education are listed below.

Effective learning strategy—Educating children through multimedia is a very effective strategy of making students understand the concepts clearly. It helps a long way in making students retain the studied concepts in their brains and even reproduce them with more clarity and precision [1]. Teachers having good knowledge of multimedia become more capable of teaching students by making use of advanced methods like educational websites and engaging students with blog writing. This has the potential to make students grasp the latest developments taking place around them and remaining updated and upgraded to latest technologies.

Traditional Pedagogy versus technology—Multimedia has the versatility to engage a much larger audience when compared to classroom teaching. One of the most important advantages of multimedia education is that it can be accessed by students at the same time from any part of the blog [2]. For example, video conferencing through google meet, google duo, zoom etc., can help students connect with their teachers online and learn lessons without any time lag unlike traditional method of teaching.

Interpersonal Communication—The use of multimedia in teaching enhances interpersonal communication between the teacher and students as well as among students. The story telling method of teaching by using computer graphics, sound and animation has proved to keep up students focused for a much longer time ensuring their better involvement in their learning processes [3]. It has been proved that story telling method not only engages students better but also helps them in learning difficult concepts in a much more simpler and easy way.

Easy access—Multimedia helps in saving time and reducing costs when compared to traditional methods of teaching. People irrespective of their age can learn through multimedia as it is not time dependent or place dependent. One can learn through multimedia according to their own convenience by sitting in any part of the world. It helps working professionals to get educated and promotes distance education through advanced technology and networking capabilities.

User-friendly aspect of Multimedia—Multimedia helps in disseminating high quality information to its users. The various aspects of multimedia like text, audio, video, graphics, animation etc., can be provided through websites and even blogs making the user or student to indulge in the learning process more precisely with minute level detailing [4]. The user gets an opportunity to connect their minds with the most recent and upgraded information with the whole internet network as a canvas for them to give wings to their innate thoughts and ideas.

Stimulated Learning—It is noteworthy that Multimedia appeals to the four senses of man and, hence, reduces the time required to learn even the difficult concepts. Multimedia makes the whole learning process less boring, more interactive and improves the attention and involvement of the students [5]. This is evident in today's teaching strategies making use of three dimensional and even four dimensional technologies in virtual labs and classrooms [6].

Efficient and Quick learning—Various researches have proved that multimedia enhances brain's ability to find an association between verbal and visual representations of the educational content. This in turn makes one improve his high level thinking skills, problem-solving skills and ability to work in groups. Their quick grasping and understanding of the content helps them think along the similar terms of understanding with respect to other complicated situations as well [7]. This can be invariably seen in the learning of coding in computer applications as well as computer programming which has increased manifold among students studying different subject streams in schools, colleges and universities. Coding, cloud computing, data science etc., are learned by people irrespective of age from different walks of life amply showing the influence of multimedia in current society.

Logical problem-solving—Moving pictures, graphics and animation are more capable of making an imprint in the brain when compared to studying with textbooks alone [8]. With the introduction of computer science as a subject in the schools and with computer programming becoming very relevant for every job, students now use programming to logically solve problems not only in the field of computer science but also related fields like information practices and mathematical computations and modeling. Hence, multimedia has paved the way for more studies in the applied levels of very complicated subjects too. This is evident in students taking the help of multimedia not only for information retention but also for information processing and converting the information processed for finding logical solutions to problems in related areas. Since multimedia resents content in the visual mode, it becomes easy for students to process the content effectively as a majority of the thought processes in human brain take place according to visual impulses received [9].

Availability of vast amount of resources—With the developments in the technological world, now majority of the population have access to smartphones, tablets, laptops, notebook computers and even the technological gadgets for entertainment [10]. They are now free to search for any information according to their comfort at the fingertips. Online information is widely searched for and has become the very common source for researchers, teachers and students. This increases the confidence of students and ensures their active participation in the class.

Better information exploration—It is not a surprise to say that children are the best beneficiaries of information explosion. With this, students are able to search for any information about which they have a doubt. This is done regardless of subject, place, time and identity [9]. For example, in a history class, students can search for monuments of historical importance, heritage sites; in a science class, they can search for planets, stars, solar system; in a Mathematics class, they can search for various mathematical models which make the subject easy, in a biology class, they can search for rare varieties of plants, endangered species of animals etc. All these information are available as graphical content and animation. Hence, anything and everything they need to know are now at their behest because of the continuous process of development occurring in the field of multimedia and animation. This has been of great help not only to students but also the teaching community at large who have gone for a blend of traditional lecturing with multimedia teaching for an information upgrade and applicability of their subjects in daily life.

Moreover, multimedia can be used for promoting genuine social causes and also helps in making the learner become more creative and express his or her emotions freely.

3 Disadvantages of Multimedia

However, there are many disadvantages for using multimedia for classroom teaching. For example, lack of technological devices in the case of unavailability of both software and hardware. Multimedia teaching in classrooms demands teachers and students well versed with modern technology. A good amount of time is necessary for planning, designing, developing as well as evaluating activities making use of multimedia. The need to store data and information while using multimedia is another challenge for both students and teachers alike.

Addiction to multimedia—One among the major problems associated with multimedia is the overuse of multimedia by the learner. Students tend to spend much more time on internet than necessary. According to the latest findings, learners spending much time on multimedia become less active and become vulnerable to numerous psychological disorders. Because of excessive use of internet, they lose sleep and suffer from poor concentration in classrooms [11].

Multitasking and multimedia—Multitasking has become a necessity in today's busy world and people do it mainly with the help of multimedia and internet. But this can be very fatal at times when students watch videos on internet while learning and other productive activities. Multimedia addiction is an important menace to address as it distracts students from reality, finally ending up with virtual reality and cyberspace as their abode instead of giving importance to human relationships, family, friends and people [6]. Addiction to multimedia especially with regard to online gaming has taken the lives of many students including children and teenagers.

Lack of Interaction—Multimedia overuse can lead to lack of productive interactions between students and teachers. Many learners end up in not knowing how to approach learning through sound, text, graphics and animation content. This is not the case with classroom teaching where teacher communicates directly or physically with students. Doubt clearing sessions between teachers and learners have found to be more effective when made physical and this also improves the mental well-being of students as they need care from the teachers as mentors during their growing up years. The interaction between learners and multimedia content is comparatively very less when compared to classroom learning.

Time Consumption—We can list out various positive aspects and benefits of multimedia. At the same time, we are supposed to think about its drawbacks too. Teachers tend to spend a good amount of time on internet on their laptops and mobile phones preparing the multimedia content for teaching. Students too get hooked to tabs, laptops and smartphones at a time when they are supposed to cultivate healthy habits like socializing with their peer group, reading books and exercising [12]. With the coming of the age of internet and multimedia, students show a tendency to isolate themselves from others, become introverts getting badly influenced by internet content like feeling desperate to match with the standards of people who pose for photos and write content regularly for marketing and projecting themselves as the best and successful.

Marketing through Multimedia—With the advancement in technology, multimedia has become a great platform for marketing. Kids and teenagers often suffer from various psychological issues by viewing too much of the marketing content or ads that come while searching for information related to education. But marketing with the help of multimedia is now considered an important business strategy and hence cannot be neglected. One of the major reasons for multimedia development is need to marketing and once internet and social media platforms cease to be marketing friendly, it can lead to the collapse of multimedia.

Dependency on resources—Multimedia demands a huge amount of resources ranging from a consistent power supply to a good data connectivity. Using multimedia takes up lot of storage space on one's computer and is costly due to high power consumption. Multimedia use demands experts from various fields in order to produce a remarkable output. This is also a factor that distracts learners as it makes them lose interest in their studies and day today activities eventually ending up in the virtual space by creating videos, reels and other content for entertainment purposes to market themselves and attract people. This makes them lazy underactive learners. Such a situation can be averted by proper supervision of parents by limiting the amount of resources given to children for multimedia use.

4 Multimedia and Mental Health

Human beings are social animals and social life is a very important part of human life. Every man is depended on other men for his day-to-day activities, performing jobs and for survival. In today's world of rapid progress in technology, making associations and befriending people even from different parts of the globe is made possible. It has been proved that, socialization and strong bond between people has got a very important role to play with regard to a person's happiness and overall well-being. Socialization has been playing a pivotal role in reducing anxiety and avoiding depression to a very large extent. It also improves one's self-esteem, self-respect and can increase a person's life span. Multimedia especially social media plays a huge role in beating loneliness and has proven to be necessary for both mental as well as physical health.

In the present scenario, youngsters using social media platforms like Facebook, Twitter, Instagram, LinkedIn, Tumblr, YouTube etc., have become very common. Even though social media plays a huge role in making people connected to each other, it is high time that we stop turning a blind eye to its side effects and long-term consequences. Psychologists point out that real and personal contact between people is much more effective in making people happy and this can help in the better secretion of happiness hormone like serotonin and other hormones essential in maintaining mental stability like dopamine, oxytocin, endorphin etc. [13]. Even though it remains a fact that multimedia makes people enter into positive relationships, it is also true that social media platforms also get misused for fake relationships with users making fake profiles for luring the youth especially girls into unsafe havens. Overtime engagement with social media can even lead to inhibition dealing with people and situations in real life, loneliness, extreme despair, dejection and even suicidal tendencies at large.

Cyber bullying is perhaps considered as the biggest menace in today's world of networking and connectivity with social media platforms becoming the breeding grounds for hate videos, religious, political and even personal, which can lead to permanent wounds in the minds of people especially children in their years of intense psychological and sociological development. With the youth spending excessive time on multimedia, they tend to severe all connections with the real world and get hooked to the virtual space which they consider as more safe and entertaining than the real world leading to too much confinement to self and no longer becoming a part of the society. The usage of multimedia as a veil to cover up one's real self and to project one as very happy, contented and successful in life than others is a very detrimental tendency while the reality can be too depressing and annoying to the user. Implementation of stringent laws to prevent overuse and misuse of multimedia and spreading awareness among children, teenagers and youth regarding the lurking dangers in using social media platforms is mandatory if we need to promote a healthy balance between multimedia use and social life.

5 Tackling the Harmful Effects of Multimedia

Reducing time—Reducing the time spent on social media is a necessary step to curb multimedia misuse. A good sleep routine not only keeps us refreshed and energetic the whole day but can also reduce the symptoms of depression and is a great way of beating work pressure, stress as well as elevate one's mood. With this simple remedy, one can resume one's focus on daily requirements, concentrate on studies as well as bringing more efficiency and productivity in their jobs. This is widely known as digital detox with many people now giving up their smart phone use for feature phones. It is advised to reduce the time spent on social media and internet browsing with the help of apps, switch off the phones at regular intervals in a day, completely avoid its use before going to bed at least three hours before bed time, to turn off one's notifications, limit the time of checking new messages and responding for them, uninstall social media apps which he or she finds less important.

Change of Focus—Be specific about what you want to search on social media rather than merely scrolling the pages through posts of your friends or relatives. This can not only prevent the user from wasting his time but also makes him aware of the dos and don'ts of multimedia [14]. This step can significantly help in using multimedia for killing time. Other than seeing the posts of your new friends on social media, one can try to go through the posts of one's dear ones, enquire their well-being and feel like a family instead of wasting time to check, post comments or click likes for the posts of strangers over the fear of losing them and their friendship. This can be a solution for getting bored and browsing inadvertently on the websites. A clear motivation can make the multimedia experience more satisfying and enjoyable.

Finding Alternatives—Many times, we ponder over substitutes that can replace as well as reduce our multimedia use. Some of them can be becoming more active in the social circles, by being religious, reading books of one's choice, opting for healthy entertainment options like listening to music, playing musical instruments, picking up a new hobby, exercising regularly, going for a walk among a few. All these measures can be greatly beneficial to curb social media overuse and act effectively in de-addiction procedures [15].

Countering the fear and inhibition of not getting noticed is an important step in de-addiction. Counselling sessions by psychologists to improve the temperament of people addicted to multimedia use can help a long way in reducing the problems of addiction. Users should be advised to concentrate on what they have instead of brooding over what they lack. A healthy social life makes one understand the reality of everyday life and can make them stick on to a healthy lifestyle keeping aside their unhealthy routine of watching multimedia all the time when they find themselves isolated and dissatisfied. They must be made to know the difference between what is right and what is fake. They must be taught that each and everyone in the society are unique in one way or the other and everyone has got their own share of problems be it psychological, physical, financial or social [16].

Spending more time with offline friends—People, in general, always like to be in others' company in order to reduce boredom, be happy and healthy. This is one of the prominent reasons why most people look up to social media as a suitable option as it helps them create a relationship with so many people at a time. In this way, social media acts as a great facilitator in building life-long relationships. But if a person is deeply interested in offline friendships, that can be encouraged more than online friendships. This is accepted, more especially, in the wake of cyber bullies and other criminals with fake profiles on social media. Relationships are obviously an important feel good factor for all people irrespective of their age. This real-life relationship and socialization makes a much more better effect on the brain and thereby on an individual's mental well-being. Moreover, offline friendship is a great mood lifter and assures better bonding between peers and hence more meaningful. Offline relationships can be developed by setting aside a good amount of time with one's family and planning regular meetings at cinemas, over a cup of tea at restaurants and even while working out in gymnasiums. One can also try to catch up with old friends in schools and colleges as there are many number of alumni associations and cultural clubs coming up especially in cities. Whatever may be the situation, it is advised to go for real-life relationships rather than virtual ones with strangers owing to lack of safety and security, a major cause of concern for cyber police as well.

The need to express gratitude—It is said that showing courtesy or gratitude to others even for small favors can be a great way to feel happy and contented. This can be practiced to get rid of anxiety, stress, peer pressure, job-related stress, dissatisfaction and lack of satisfaction with the present situation. This also makes us feel more confident, energetic and also improves our productivity and efficiency in doing our work. Keeping all our memories fresh by taking a note of them especially the most cherished and joyful ones helps us keep up the spirit of our life without giving to boredom or brooding. These memories make us feel very positive and hopeful and keep us active within our social circles. Taking enough time for self-reflection is another way of coping with problems caused by multimedia. Giving expression to one's thoughts, feelings and emotions can be better possible offline and in a direct way, so that both the speaker as well as the listener feels the conversation to be more engaging. In the networking world, we mostly depend and await the likes, comments and shares of our friends to define ourselves and this is a very dangerous tendency especially among kids and teenagers who sit and wait for getting others compliments for their posts often wasting hours instead of doing something creative and productive in their daily lives. Parents must see that their kids never fall a prey to psychological disorders like depression and anxiety because of scrutiny on social media platforms. Parents should see that their children do not take to social media for getting appreciation and gaining others' attention. This can not only hamper their concentration on studies but can also make them a prey for long-time psychological disorders.

Saying 'No' to unhealthy comparisons and competitions—Instead of focusing on others' haves and have nots, children and teenagers, in particular, and adults, in general, should concentrate on their own positive and negative qualities. Practicing mindfulness can be a great way of reducing problems with regard to multimedia overuse. There have been numerous cases in which people are duped by online fraudsters as in the case of unauthorized and unapproved online games that mint money, cheating the youth who get deeply disappointed because of common everyday realities of life. Such people tend to go for online gambling and make easy money. Parents must see that their children take failure and success in a positive way and impart them the lesson that there is always a chance for improvement and betterment in their lives. The government also has a major role to play in removing addictive and regressive apps from social media in order to protect a country's youth and make them dutiful citizens of the country where they work for the progress of themselves as well as their country.

Taking the leadership and volunteering –Just like an individual socializes, makes social contacts and friend circles, he or she can get more benefitted out of socializing through volunteering. Taking part in various competitions conducted by social welfare organizations, the adventure clubs, sports clubs, community service organizations, festival celebrations etc., can help a long way in making people feel very less lonely and well connected. Volunteering and offering help to others through NGOs and even spending time with pets can make one live a very happy and contented life without giving in to networking on multimedia. Volunteering from a young age makes children develop good leadership qualities, remain focused on studies and remain enthusiastic and live their lives happily and successfully.

6 Conclusion

It is a well-known fact that nowadays, human life is largely dependent on machines. Gone are the times when people used to do things manually and this has led to the work replaced by machines. In the current scenario, individuals getting benefitted by multimedia and internet services should also ponder about the disadvantages and harmful effects of multimedia. Sitting at computer for a long time leads to a sedentary lifestyle causing various lifestyle diseases. Adults as well as children need a fine tuning in their lives with a good amount of time to be spent on inculcating good habits and hobbies instead of relying on multimedia for entertainment. We have seen many online games claiming young lives and this tendency to find solace and happiness in life must be nipped in the bud. Schools, colleges, universities and other authorities can play a huge role in avoiding multimedia overuse and misuse by monitoring the daily activities of students in their institutions. Parents ought to see that facilities like multimedia should be used in a productive way for learning purpose and not for getting distracted from everyday realities of life. The nation as such has a collective responsibility in preventing the young from getting misguided in life due to easy access to multimedia these days instead of using it positively for the betterment in their lives.

References

1. Clark RC, Mayer RE (2016) E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning. Wiley
2. Evans C, Gibbons NJ (2007) The interactivity effect in multimedia learning. Comput Educ 49(4):1147–1160
3. Rieber LP (2005) Multimedia learning in games, simulations, and microworlds. In: The Cambridge handbook of multimedia learning, pp 549–567
4. Astleitner H, Wiesner C (2004) An integrated model of multimedia learning and motivation. J Educ Multimedia Hypermedia 13(1):3–21
5. Schnotz W, Lowe R (2003) External and internal representations in multimedia learning. Introduction. Learn Instr 13(2):117–123

6. Mayer RE, Moreno R (2002) Animation as an aid to multimedia learning. Educ Psychol Rev 14:87–99
7. Plass JL, Jones LC (2005) Multimedia learning in second language acquisition
8. Leacock TL, Nesbit JC (2007) A framework for evaluating the quality of multimedia learning resources. J Educ Technol Soc 10(2):44–59
9. Mayer RE (2002) Multimedia learning. In: Psychology of learning and motivation, vol 41, pp 85–139. Academic Press
10. Brunken R, Plass JL, Leutner D (2003) Direct measurement of cognitive load in multimedia learning. Educ Psychol 38(1):53–61
11. Mayer RE, Heiser J, Lonn S (2001) Cognitive constraints on multimedia learning: when presenting more material results in less understanding. J Educ Psychol 93(1):187
12. Boyle T (1997) Design for multimedia learning. Prentice-Hall, Inc.
13. Pea RD, Gomez LM (1992) Distributed multimedia learning environments: why and how? Interact Learn Environ 2(2):73–109
14. Mishra D (2021) Implications of E-learning in education: an analysis. In: Next generation of internet of things LNNT. Springer
15. Mayer RE (2003) The promise of multimedia learning: using the same instructional design methods across different media. Learn Instr 13(2):125–139
16. Mayer RE, Moreno R (2002) Aids to computer-based multimedia learning. Learn Instr 12(1):107–119

Effective Detection of Liver Disease Using Machine Learning Algorithms

Amol Dhumane, Sudeep Pawar, Rajas Aswale, Tejas Sawant, and Shishir Singh

Abstract The human body consists of many important organs one of which is the Liver also known as the largest integral organ, present inside the human body it has certain functions which help in maintaining the ecosystem inside. The liver is known for performing close to 500 functions that are vital to keeping the human body working efficiently. Many people die due to liver diseases nowadays and it is impossible to keep track of everyone with the number of doctors present and the ever-growing population. It is necessary to detect liver disorders or it can lead to many lethal problems such as coma or even death in certain cases. Many researchers have performed the spotting of liver disorders through various ML models. In this paper, we use and compare various classification techniques coupled with the types of liver diseases. Random Forest Classification, Support Vector Machine Classifier, and Logistic Regression are the techniques we will be working on within this particular project on the basis of various features cut out from the dataset. Comparative analysis is taken into consideration to measure their accuracy results for knowing the better algorithm for further analysis and use with successful detection and prediction of liver disorders.

Keywords Liver disorders · Machine learning · Classification · Random forest · Support vector machine · Logistic regression · Comparative analysis

1 Introduction

The liver is one of the most critical and important organs of the internal system of our bodies; it is present on the lower right-hand side of the body above the waist. It plays a huge role in the removal of toxic substances, protein synthesis, production of bile, albumin production, and production of bilirubin in the right amount and releases enzymes such as aminotransferase and globulin, various chemicals as well which help in the digestion of food efficiently. It has a special ability to regenerate itself

A. Dhumane (✉) · S. Pawar · R. Aswale · T. Sawant · S. Singh
Pimpri Chinchwad College of Engineering, Pune, India
e-mail: amol.dhumane@pccoepune.org

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_15

which is a very rare quality, if for example there is only 30% of liver tissues remaining then generally it takes 7–12 days to completely regrow the liver without any loss of functionality. In order to stay completely fit, we need to take good care of our livers or else the malfunctioning of it would lead to many diseases such as liver cancer, cirrhosis, liver failure, hepatitis, fatty liver disease, ascites, and hemochromatosis. There are many reasons for liver disorders take place, some of them are obesity/ fleshiness, drawing tattoos, excessive use of alcohol, genetically inherited, etc. We can predict that a person may have liver disorder based on the signs and symptoms listed for a liver disease which are abdominal pain, loss of appetite, jaundice, itchy skin, dark urine color, fatigue, etc. Liver disorder disturbs the patient's life to huge extent and according to the statistics, liver disease caused lakh deaths in India which is 3.17% of the total deaths, and is ranked 8th in terms of life expectancy list of diseases which are the major cause of death. Early detection and treatment as soon as possible are the only way to get rid of this problem. Doctors and practitioners decide that a person is affected based on CT scans, ultrasounds, liver biopsies, MRIs, and blood tests. It is a difficult task to perform so it takes a lot of time. So, in order to assist doctors, we use machine learning algorithms which play a vital role in disease identification. Building a model through ML helps extract important data. Due to the non-linear characteristics and complex structure of the datasets, it is difficult for the researchers to predict accuracy at a satisfactory level, and the presence of outliers and missing values makes it difficult to make predictions.

In this particular paper are going to use various classification and regression models such as Logistic Regression (LR), Random Forest (RF), and Support vector machine (SVM) for the prediction and also try to increase the accuracy of the models which we are going to implement. The remaining part of the paper will give you a clearer picture of the entire procedure that we will follow in order to achieve our goal and get the desired output. The paper further is segmented into various sections or we can say step-by-step procedural conduct of how we go about things from the study of past related work on liver diseases to the proposed methodology which includes our work from data description, preprocessing, mining, etc.

2 Literature Survey

This section describes research that was conducted in a variety of sectors.

The author of the study [1] segmentation provides distinguishing knowledge about the seriousness of liver disorders based on the size and shape of the lesion. With the contrasting of liver lesion segmentation performances, the authors of this research examined the two theories. They created three distinct segmentation modules with different windowing models. On this basis, they contrasted the trainable window models and the preset window's capabilities in segmenting liver tumors. They also looked for room to improve their results by increasing the number of contrast windows. They utilized the (Lits) dataset that was published in ISBI 2017.

In this paper [2], the mathematical model known as artificial neural networks (ANNs) is an inspiration derived from the biology of the human nervous system. Medical decision-making employs ANN. The goal of this ANN model is to reliably identify unsafe groups of people that may suffer disease succession by anticipating the early-stage warning in patients with HBV-HCC. They registered 2117 patients between 2008 and 2016, therefore dividing them into training datasets ($n = 1480$) and validation datasets ($n = 637$) at random. In the training group, there were 434 (29.5%) cases of illness progression within a year, 174 (40.1%) fatalities within that year, and 310 (71.4%) fatalities within three years of the disease's progression. Of the 784 patients who had a 1-year Hepatology International 13 succession-free survival, only 173 (22.1%) of them were no more in a span of 3 years. 181 (28.1%) of the validation group's participants experienced sickness progression after a year, 93 (51.4%) of them passed away, and 134 (74.0%) did so after the disease had advanced for three years. Of the 327 patients who had a one-year succession-free survival, only 82 (25.1%) lost it in a span of 3 years.

Paper [3], purpose of this paper is the early identification of Patients who are at risk for colorectal liver metastases which can aid in personalizing care and enhancing oncological outcomes. The purpose of the survey was to determine whether a radionics model on the basis of machine learning could predict the development of metachronous metastases in patients with colorectal cancer (CRC) by analyzing the tested characteristics of the patients. In 41 patients' colons and 50 patients' recta, the primary tumor was discovered. 67 patients showed no signs of metastases during a follow-up that lasted at least 24 months following the initial diagnosis. In total, there were 24 patients (26.4%) who acquired metachronous metastases; 10 of these metastases appeared between 7 and 12 months after the primary CRC diagnosis, and the other 10 between 13 and 24 months. There were no such differential values that separated any aspects between the patients in the training and validation sets.

In this paper [4], one of the most essential therapies for the end (last)-stage of liver disorder is liver transplantation, but there is a significantly greater need for livers than there are donors for donating available. The Model for End (last)-stage Liver Disease (MELD) score is a frequently used method for ranking victims, but prior research has shown that the MELD score may be less accurate at predicting postoperative outcomes. Physiological measurement values are used to generate new features and clinically relevant features that we then pick using random forest (RF). Additionally, we suggest a novel imputation technique to address the issue of missing data, and the outcomes demonstrate that it performs better than the alternatives. Patients' blood test results from one to nine days before surgery are used to build the predictive model.

Paper [5], the enormously notable independent predictor of recurrence for hepatocellular carcinoma (HCC) is microvascular invasion (VMI), however, determining this before surgery is difficult. In this work, they look into how multi-parametric MRI radionics is often used for forecasting mVI conditions prior to surgery. For 99 patients with HCC, they retrospectively gathered pre-operative multi-parametric liver MRI scans. Surgery and a pathology-confirmed diagnosis of mVI were given to this patient. They created machine learning classifiers using radionics features they derived from manually segmented HCC regions to forecast mVI status. They assessed how well such classifiers performed when constructed on five separate and combined MRI sequences. They looked at the outcomes of combining features collected from the peritumoral marginal zone and the tumor region, as well as the effects of each separately.

Paper [6], with a C statistic of at least 0.90, the Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo), created at a conferral center in the United States and internally validating, greatly exceeds the competition. A welcome advancement in the field is the application of ML to predict PSC results. The application of ML in PSC in this study, which examines a cholestatic condition, is especially appropriate given the requirement to combine diverse, frequently shifting factors into a single score.

Through the suggested study effort in the paper [7], an advancing algorithm known as LIVERFAST is used to assess the diagnostic effectiveness of the machine learning (ML) algorithm for measuring liver damage using the most recent AI technologies.

The distinction of NASH from straightforward steatosis and the detection of advanced hepatic fibrosis are crucial difficulties for NAFLD patients. A specialized machine learning (ML) algorithm named LIVERFAST was made using a dataset with 2862 distinct medical assessments of biomarkers for this potential study's analysis of the staging of 3 different liver lesions to diagnose fatty liver. Of these assessments, 36% were used for training the algorithm, and 64% made up the validation dataset.

In their study [8], the purpose was to effectively detect Disease Free Survival (DFS) of HCC patients after LR using an ML algorithm based on readily available information. The clinical and laboratory data were chosen in accordance with earlier studies that identified indicators for predicting HCC outcomes. Bioinformatics was used to find pertinent model-relevant variables. Although not all variables in the multivariate analysis were independently predictive, their integration in the RF ML increased the detection level of the end model.

Using this approach, they could create and test the RF model that effectively detected initial DFS after liver restationing in the case of HCC using common clinical and laboratory information. Future treatment allocation may vary as a result of ML modeling, with low-risk patients receiving LR and high-risk patients being listed for LT.

Paper [9], for the purpose of screening patients for liver disease, the author of this research made a compactly attached Deep Neural Network (Dense DNN) based on the 13 most frequent LFT (liver function test) indicators and demographic data. 76,914 samples were used in the dataset to test the algorithm. The dense DNN Area Under the Curve is 0.8919, DNN is 0.8867, Random Forest's is 0.8790, and Logistic Regression is 0.7974. Deep learning models perform best with traditional approaches. When it comes to deep learning techniques, Dense DNN performs better than DNN.

Paper [10], this study's goal was to use machine learning to create prediction models for HCC linked to Chronic Hepatitis C (CHC). To find the important factors that can accurately predict the occurrence of HCC, raw data from 4423 CHC patients were examined. In order to predict the existence of HCC, this survey built HCC classification models using a variety of machine learning techniques. Statistics revealed that HCC existence was related to age, alpha-fetoprotein (AFP), alkaline phosphatase (ALP), albumin, and total bilirubin characteristics. Several ML techniques/models were used to build several HCC classification models. The suggested models for classifying HCC offer a good circumference under the receiver operating characteristic curve (AUROC) with a high degree based on diagnostic precision.

3 Proposed Work Model

Herein there will be a step-by-step process to work on the dataset. These steps are data collection, data pre-processing, splitting of the dataset into training and testing accuracy analysis, and modeling.

3.1 Data Collection

Collection of the dataset for liver diseases like liver cancer, ARLD (alcohol-related diseases), and inherited liver diseases from Kaggle (Indian liver patient records). The dataset used by us contains the following features Age of the patient, Gender of the patient, Total Bilirubin, Direct Bilirubin, Alkaline Phosphatase, Alanine Aminotransferase, Aspartate Aminotransferase, Total Proteins, Albumin, Albumin, and Globulin Ratio.

3.2 Data Pre-processing

Data pre-processing is one of the key factors of machine learning models and if there are any missing values, outliers, and irrelevant features, it may affect model accuracy. The performance of the model mostly depends on the quality of data. We use the Min-Max and Standard Scaling techniques to transform the data, and they are imported from sklearn.pre-processing library. Then for handling the missing values, we make use of Simple Imputer which is a scikit class that then uses different strategies like mean, mode, and median. Our motive is to increase the prediction accuracy of the model. We can deal with missing values by various methods like replacing those values with means of that column etc.

Splitting into training and testing sets—in this step, the dataset will be divided into parts in a ratio of 75–25 with the help of train_test_split function which is imported from sklearn.model_selection library, wherein we take the first 75% of the dataset for the training dataset and the remaining 25% for testing.

3.3 Modeling

To predict whether the patient has liver disease or not, we will be using three techniques which will consist of two classification models and one regression model, which are further defined in Fig. 1.

Data set acquired from Kaggle: https://www.kaggle.com/datasets/uciml/indian-liver-patient-records

3.3.1 Support Vector Machine

Support vector machine (SVM) is one of the most efficient supervised learning models used in machine learning. It can be used for both regression and classification problems but is mainly used for classification models and we will implement the same. We collect the dataset and define the x and y variables wherein y is the output i.e., our main objective of whether a patient has liver disease or not and x consists of the input parameters. Then we transformed the data using the standard scalar method from sklearn.preprocessing library. We imported train test split methods from sklearn.model_selection, Then we defined (x_train,x_test,y_train,y_test) and we passed hyperparameters (test_size, random_state etc.) following which we import SVC from sklearn.svm furthermore creates an object fitting it into the model. After tuning parameters using hyperparameter tuning and grid searchCV, we got the best parameters as: 'C': 1, 'degree': 2, 'gamma': 10, 'kernel': 'rbf' using these we got.

Accuracy Score: 73.50% (Graph 1).

3.3.2 Random Forest

Random forest is a machine learning method that can be applied to both regression and classification problems. It constructs several decision trees and combines those trees to get a more well-built and precise prediction. The random forest can handle the overfitting problems of decision trees when building a model from training samples. We don't transform the data as herein it is in a hierarchical manner which precedes the extra tree classifier technique. We follow the same steps that we have followed in SVM just importing the Random Forest Classifier from sklearn.ensemble, followed by fitting into the model and then generation of accuracy. By using hyperparameter tuning and grid searchCV we got hyperparameters values as follows- 'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 150.

Accuracy Score: 74.25% (Graph 2).

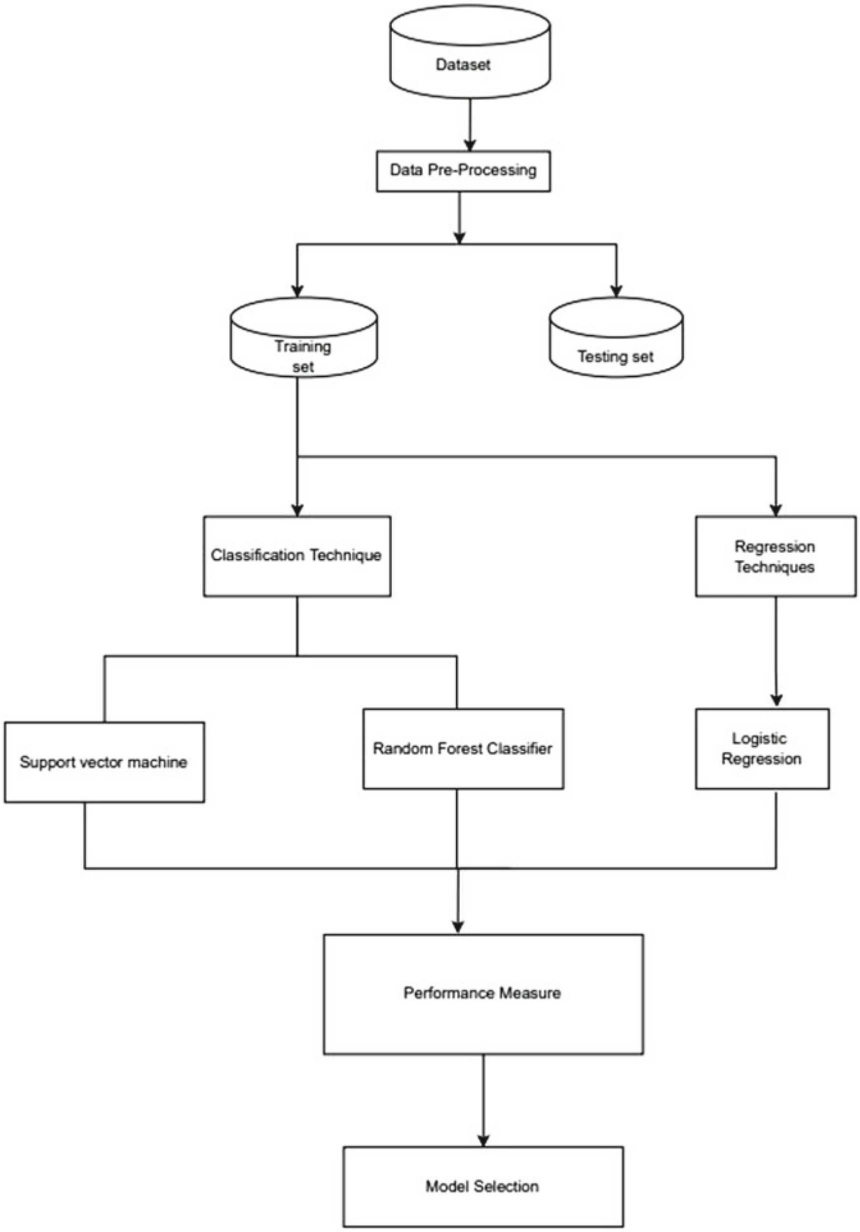
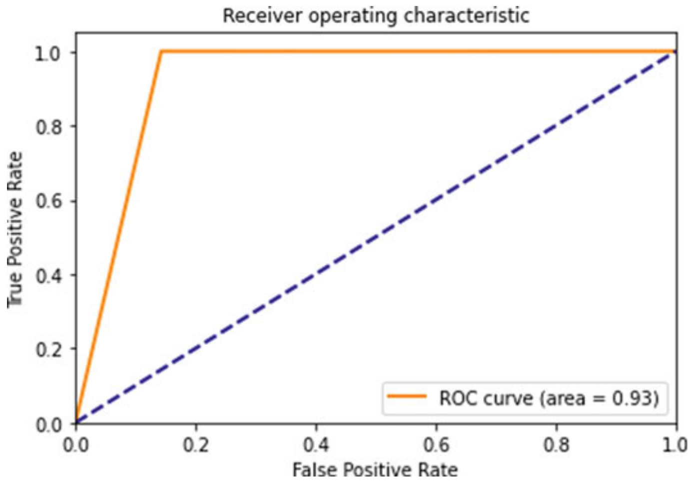
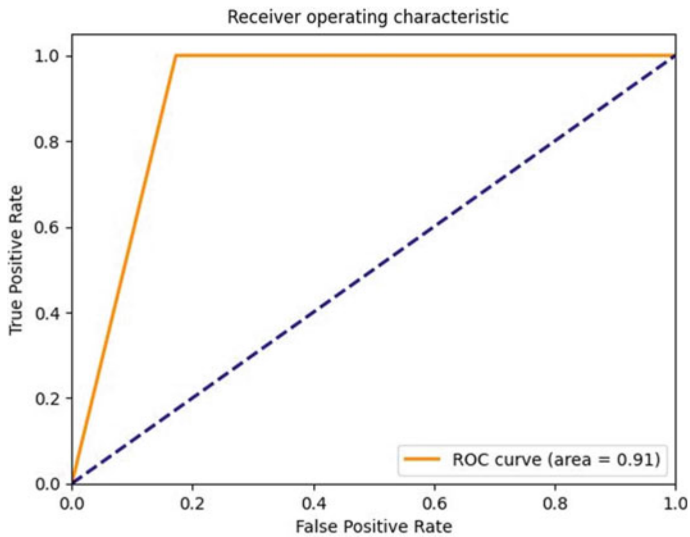


Fig. 1 Workflow diagram



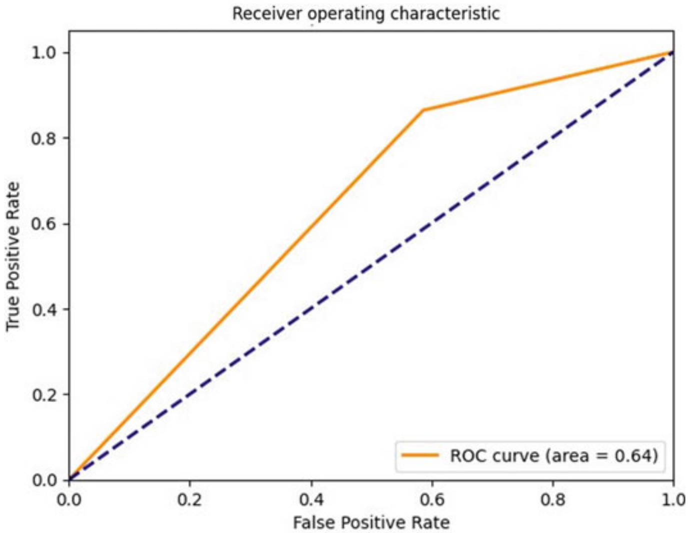
Graph 1 ROC curve based on support vector machine



Graph 2 ROC curve based on random forest

3.3.3 Logistic Regression

Logistic regression is one of the regression techniques used in machine learning that models the probability of an event. It is used to classify two-class problems. Logistic regression predicts the output of a categorical dependent variable. Therefore, the outcome must be a categorical or discrete value. In this case, the output will be



Graph 3 ROC curve based on Logistic Regression

whether the patient is having liver disease or not. During implementation, we import logistic regression from sklearn.linear_model, and then create the object fitting it into the model and get the prediction results.

Accuracy Score: 75.21% (Graph 3).

4 Conclusion

We applied machine learning-approved algorithms on a dataset consisting of liver-prone patients. The early detection of liver disorder is necessary to help the patient know about the seriousness of the level by which his/her liver has been affected. The main issue that brings to light upon this is that the doctors and practitioners aren't able to detect the failure early which causes a lot of problems later on when the disease becomes close to incurable. So, by using the above machine learning techniques and with the help of various libraries, we can achieve the main objective of our project which is the early-stage detection of liver diseases, and provide a helping hand to the healthcare sector by performing the necessary. The survey paper is based on several researchers' findings. The Logistic Regression is the best algorithm that can be used and applied to get the best possible accuracy according to the outcome of our implementation with 75.21%, followed by the Random Forest model with 74.36% and SVM with an accuracy score of 73.5% (Fig. 2; Table 1).

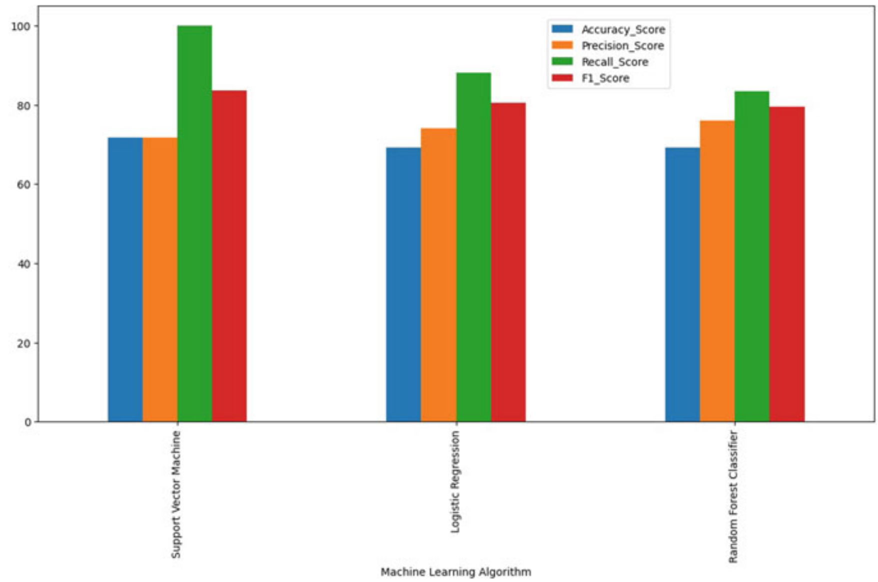


Fig. 2 Graphical Study of the results based on the table above

Table 1 Comparison between various machine learning algorithms

Algorithm	Accuracy score	Precision score	Recall score	F1 score
Support vector machine	73.50	73.50	100	84.7
Random forest	75.21	77.67	93	84.7
Logistic regression	74.36	78	90.7	83.9

References

1. Kwon J, Choi K (2020) Trainable multi-contrast windowing for liver CT segmentation. In: 2020 IEEE international conference on big data and smart computing (BigComp). IEEE, pp 169–172
2. Liu X, Hou Y, Wang X, Yu L, Wang X, Jiang L, Yang Z (2020) Machine learning-based development and validation of a scoring system for progression-free survival in liver cancer. Hepatol Int 14:567–576
3. Taghavi M, Trebeschi S, Simões R, Meek DB, Beckers RC, Lambregts DM, Verhoef C, Houwers JB, van der Heide UA, Beets-Tan RG, Maas M (2021) Machine learning-based analysis of CT radiomics model for prediction of colorectal metachronous liver metastases. Abdom Radiol 46:249–256
4. Liu CL, Soong RS, Lee WC, Jiang GW, Lin YC (2020) Predicting short-term survival after liver transplantation using machine learning. Sci Rep 10(1):56–54
5. Nebbia G, Zhang Q, Arefan D, Zhao X, Wu S (2020) Pre-operative microvascular invasion prediction using multi-parametric liver MRI radiomics. J Digit Imaging 33:1376–1386
6. Shung DL, Assis DN (2020) Machine learning in a complex disease: PREsTo improves the prognostication of primary sclerosing cholangitis. Hepatology 71(1):8–10

7. Aravind A, Bahirvani AG, Quiambao R, Gonzalo T (2020) Machine learning technology for evaluation of liver fibrosis, inflammation activity and steatosis (LIVERFAStTM). J Intell Learn Syst Appl 12(2):31–49
8. Yao Z, Li J, Guan Z, Ye Y, Chen Y (2020) Liver disease screening based on densely connected deep neural networks. Neural Netw 123:299–304
9. Schoenberg MB, Bucher JN, Koch D, Börner N, Hesse S, De Toni EN, Seidensticker M, Angele MK, Klein C, Bazhin AV, Werner J (2020) A novel machine learning algorithm to predict disease free survival after resection of hepatocellular carcinoma. Ann Transl Med 8(7)
10. Hashem S, ElHefnawi M, Habashy S, El-Adawy M, Esmat G, Elakel W, Abdelazziz AO, Nabeel MM, Abdelmaksoud AH, Elbaz TM, Shousha HI (2020) Machine learning prediction models for diagnosing hepatocellular carcinoma with HCV-related chronic liver disease. Comput Methods Programs Biomed 196:105551

Digital Financial Asset Management Using Blockchain for Succession Planning of Wealth

Tejas Patil, Shubham Patel, Siddhanth Shah, Ravi Thacker, Sudhanshu Gonge, Rahul Joshi, Ketan Kotecha, Renuka Agrawal, and Manikrao Dhore

Abstract This paper surveys and implements the application of blockchain and related technology in digital financial asset management and succession planning of wealth. Digital financial assets, such as cryptocurrencies, NFTs, and digital real estate, have become increasingly popular in recent years, but their management and distribution after the death of the owner can be a difficult and complex process. Having conducted systematic research into the area of blockchain and its various use cases, we determined how it affects wealth creation, planning, management, succession planning, and portfolio performance. Blockchain, with its decentralized and immutable nature, provides a solution to this problem by enabling the creation of smart contracts and digital wills that can be programmed to automatically transfer

T. Patil (✉) · S. Patel · S. Shah · R. Thacker · S. Gonge · R. Joshi · K. Kotecha · R. Agrawal
Symbiosis Institute of Technology, Pune, Maharashtra 412115, India
e-mail: tejas.patil.btech2020@sitpune.edu.in

S. Patel
e-mail: shubham.patel.btech2020@sitpune.edu.in

S. Shah
e-mail: siddhanth.shah.btech2020@sitpune.edu.in

R. Thacker
e-mail: ravi.thacker.btech2020@sitpune.edu.in

S. Gonge
e-mail: sudhanshu.gonge@sitpune.edu.in

R. Joshi
e-mail: rahulj@sitpune.edu.in

K. Kotecha
e-mail: head@scaai.siu.edu.in

R. Agrawal
e-mail: renuka.agrawal@sitpune.edu.in

M. Dhore
Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India
e-mail: manikrao.dhore@vit.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_16

ownership of digital assets to designated beneficiaries upon the death of the owner. This paper examines the current challenges in tracking and transferring ownership of digital assets after the death of the owner, and how this technology could be used to impact and improvise the succession planning of wealth for digital assets. The paper also explores the potential benefits and limitations of using blockchain for this purpose, as well as potential future developments in this field.

Keywords Blockchain · Wealth planning · Succession planning · Digital financial assets · Cryptocurrencies · NFTs · Digital monetizable assets

1 Introduction

Blockchain technology is drawing a lot of recognition as a potential disruptor in a number of sectors, including finance, banking, IoT, supply chain management, and management of digital assets. It is a distributed and highly decentralized ledger which can securely and openly record transactions [1].

Keeping track of all the digital assets and ensuring that they are passed to the suitable beneficiary in the event of the owner's passing are two of the most significant issues in managing digital assets. This can be particularly challenging for digital assets that are not maintained in a centralized location and may need clear ownership records, such as cryptocurrencies, NFTs, and digital real estate [2–5]. This problem can be solved by using blockchain technology, which offers a safe and transparent means to track and transfer digital assets ownership. Ethereum, EOS, TRON, Corda, and Hyperledger are among the top blockchains for managing digital financial estate [6–8]. Smart contracts can be used to build investment funds that automatically distribute profits to investors or digital wills that automatically transfer ownership of digital assets to designated beneficiaries upon the death of the owner.

Digital estate planning is essential in the context of modern wealth planning. This study aims to show how technology powered by blockchain could be utilized to more effectively distribute as well as manage digital assets after the owner's passing, hence reducing the risk of losing a fortune. Blockchain technology makes it possible to track who owns what digital assets in a safe and transparent manner, and it has the potential to fundamentally change how we manage and move money [8]. Scalability is a significant concern since certain blockchain networks' current infrastructure might not be able to manage the huge amount of transactions necessary for general adoption [2, 3, 9]. Additionally, the enforceability of smart contracts and digital wills may be in doubt due to the fact that the legal and regulatory environment for digital assets is still developing. Security and privacy concerns are a further major factor. Even while transactions and ownership records are highly secure because to blockchain technology, there are still worries about the possibility of hacking and other types of cybercrime [3]. Despite these difficulties, it is obvious that blockchain technology has the power to significantly alter the way we manage, store, and transfer digital assets [1, 10, 11]. The ownership of digital assets may be recorded and transferred in a secure and transparent manner thanks to blockchain technology, which has the potential to increase the effectiveness and efficiency of digital financial asset management and succession planning.

The goal of this study is to give a thorough analysis of the cutting edge at the moment and also in prospective time ahead advancements in this area. This study also highlights the necessity of creating digital estate planning in order to safeguard and maintain digital assets [9]. Additionally, it covers the many methods and instruments that can be applied to carry out digital estate planning.

2 Literature Review

Xi and Pattabiraman (2022) in [6] elaborate the Ethereum blockchain makes use of smart contracts, and the programming language Solidity's low-level operations have the potential to be misused, leading to security flaws. The Solidity community has identified certain insecure code patterns and advises developers to avoid using them, however, these patterns are still prevalent in actual smart contracts written in the world. Lee (2022) et al. in [7] focus on the drawbacks of conventional digital asset management systems and the ways in which blockchain technology might be applied to address them. It highlights that the EU General Data Protection Rules (GDPR) has brought data protection to the forefront, leading to conflicts between GDPR principles and blockchain. Many smart contracts deployed for digital asset management are rarely invoked, leading to a lack of creativity in smart contract designs. In order to overcome these problems, the study suggests SPChain, a more sophisticated and private digital asset management framework that complies with GDPR and is powered by blockchain. Hyperledger Fabric could be utilized with the proposed SPChain to handle digital art. Jani (2020) in [8] discuss the concept of smart contracts and their implementation in such as Hyperledger Fabric, Corda, Ethereum, EOS, Rootstock (RSK), and Stellar, among other blockchain systems. Different smart contract platforms were compared, and smart contract applications were categorized with illustrative examples. Camposano (2021) et al. in [10] explore the use of digital twins in the AEC/FM industry in Finland, conducting semi-structured interviews with C-level executives and project managers. Waheed and He (2020) in [1] explore how the Users of the Internet of Things (IoT) are increasingly concerned about their security and privacy due to the increasing number of cyber threats. Ross (2020) et al. in [11] explored financial fundamental in addition to discussing the potential advantages and disadvantages of employing NFTs in GLAM institutions; the authors give an overview of important NFTs and selling platforms. Understanding the technology plays a very crucial role in developing the solutions for managing the same. Waheed (2020) et al. in [3] expound upon the growth of the Bitcoin network has led to increased centralization, with a minute number of nodes controlling a large part of the network's connectivity. It could be understood from the literature survey that there is a need for a platform to manage the digital financial assets and handle the succession planning for the same.

3 Methodology

The methodology for a blockchain project focused on Digital Financial Asset Management and Succession Planning of Wealth using Blockchain had been as follows.

Figure 1 deals with the working principle, which works while transmitting the digital assets from one cryptocurrency account to another or vice-versa. Party 1 is the owner of a particular digital financial asset, and party 2 is the beneficiary on the receiving end. Both the parties are being mentioned in the smart contract initially and a lock in period for its execution is being set. A transaction is being created which shall be triggered in the event of the demise of the owner or if the owner is unable to extend the lock in period. Post the lock in period, the smart contract is being executed and the transaction is completed. The entries are respectively then marked in the immutable ledger. The following are the steps which are followed while payment using smart contract methodology for transferring assets from the owner to the beneficiary.

Defining the problem: First identify the specific problem which is being aimed to be solved.

Research and analysis: Determine whether which blockchain is fit for the project.

Designing the solution: Design a blockchain-based for solution for the problem statement chosen appropriately.

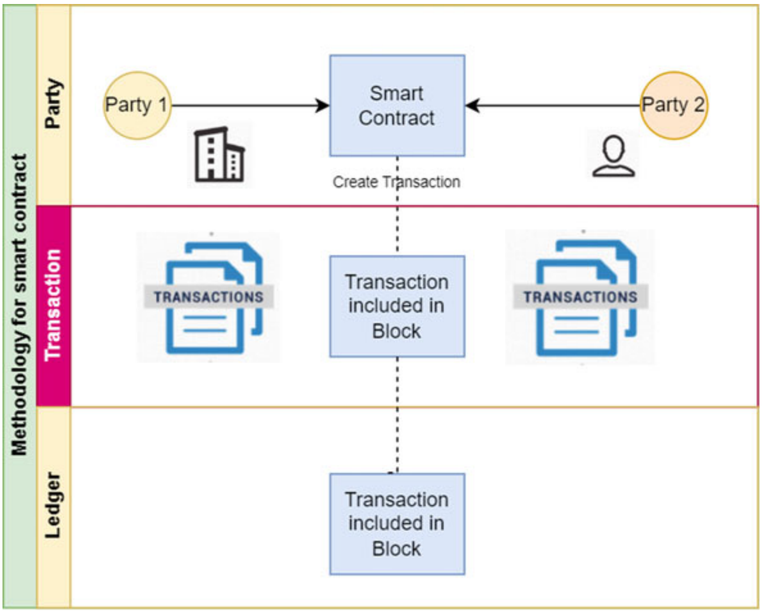


Fig. 1 Basic working principle of blockchain-based transaction

Building the prototype: Write code, create smart contracts, and configure the blockchain network.

Testing and validation: Unit testing, integration testing, and user acceptance testing.

Maintenance and monitoring: Monitor the blockchain network for any issues and address any bugs or vulnerabilities that have been discovered.

Determine Asset Types: Determine the types of digital assets to be managed. This may include cryptocurrencies, NFTs, digital real estate, and other monetizable digital assets.

Smart Contract Development: Designed to execute the transfer only after the owner's death or at a specific time as mentioned in their will.

Identify Beneficiaries: The owner can assign beneficiaries by providing their wallet addresses in the smart contract.

Integration with Web3.js: This will enable beneficiaries to view their inheritance and receive digital assets through the decentralized application (DApp).

Deployment: Configure the blockchain network, and create, update and distribute the digital assets to the users.

Gas Value Calculation: The "price of gas" is the quantity of ether a user is willfully ready to pay for each unit of the gas. The user controls the gas cap. Multiply the gas price by the gas limit to determine the transaction's gas value. E.g., the gas value would be 0.00042 ether if the gas price was 20 gwei and the gas limit was 21,000.

4 Analysis of the Collected Data

In this study, there has been a focus only on a small number of widely used digital financial assets and there has been conducted an in-depth investigation on a specific set of transactions. The asset type, time stamp, transaction amount, and gas value are among the information we have gathered on the trades between owners and recipients. This data was then pooled and evaluated in order to find trends and patterns in the transactions, including common asset types, typical transaction amounts, and variations in gas prices. Through this process, we were able to better understand the dynamics and difficulties of moving digital financial assets and identify potential areas for process improvement.

Table 1 shows different types of digital assets which have been taken into consideration in our study. The approximate gas amount consumed for the transaction between the owner and the beneficiary has been calculated. When a succession planning is done for these digital financial assets, the address of the owner and the beneficiary's wallet is noted in the smart contract. Also, the private keys and the assets which come along with the owner's digital will be transferred to the address of the beneficiary. The approximate gas values are thereby calculated to make these transactions more viable on scale such as on a commercial level.

Figure 2 visualizes the data from Table 1 in form of a pie chart. It plots from the selected sample population from which the primary data is being collected, which

Table 1 Digital assets based on gas value required for transaction

T-ID	Asset type	Asset name	Owner address	Beneficiary address	Gas used (gwei)
1	Cryptocurrency	Bitcoin	0 × 1234	0 × 5678	21,000
2	NFT	CryptoKitty #123	0 × 1234…	0 × 5678	200,000
3	Digital real estate	Virtual Land in Decentraland	0 × 1234	0 × 5678	80,000

forms of digital financial asset are popular and their usage as compared to other digital financial assets. It can be concluded from our sample data that as of the date of publication of this paper, NFTs are requiring more gas amount and are more popular among users who attach value to these personal assets followed by digital real estate and cryptocurrency.

As of January 2023, the average gas price on the Ethereum network is around 80 Gwei (or 0.00000008 ETH) per unit of gas. To calculate the gas value for each transaction in the database, we can use the following formula: Gas value = Gas limit * Gas price.

For example, if we take the first transaction in the database (Transaction ID = 1), the gas value would be: Gas value = 21,000 * 0.00000008 = 0.00168 ETH.

The smart contract enforces any unique terms or conditions that may apply. The smart contract could demand that a specific length of time has passed before the transfer can take place. For instance, if there is a lock in period, the smart contract keeps track of all the transactional information, including the name of the owner, the

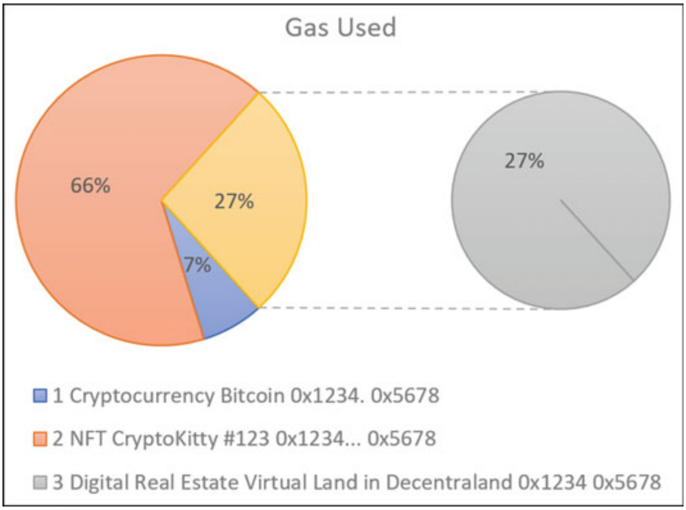


Fig. 2 Pie chart of gas value for different assets

name of the beneficiary, the type of asset, the transferred amount, the transaction's timestamp, and the cost of the gas. The smart contract can be set up to automatically transfer the assets to the selected beneficiary in the event of the owner's death. The smart contract can also be set up to renew or expire in accordance with predetermined criteria, such as a particular date or a predetermined quantity of transactions.

Figure 3 is the interface of an online portal which shall be used to create a smart contract on blockchain to transfer the assets from the owner to the beneficiary. The owner has to give the details like his email address, the receivers email address, the amount, and the secret key along with the lock in period which has to be specified in the number of days. A wallet shall later on prompt up and the user will have to enter their private keys which link their assets to them. All this data shall be stored in a smart contract which shall be executed after the expiration of the lock in period.

The owner can also add multiple beneficiaries and allocate different assets to different beneficiaries. Once this process is done, the assets are being locked and a lock in period is being specified. The address of the owner and beneficiary is also being locked. The interface is designed in a way such that the owner feels in control of his assets.

Figure 4 is a flow diagram explaining the working of the full stack decentralized application. For the front end, a JavaScript based framework like React JS is being used along with different web3js libraries. The user has an option to log into the web interface, and then the program checks for the availability of a smart contract initiated by the owner. If the smart contract is available, then a connection is established between the addresses of the owner and the beneficiary using a smart contract which is based on a popular blockchain, the lock in period is being checked. If it has been expired, the owner is given an option to renew/extend the locking period, and if the owner doesn't chose to respond owing to his/her demise or any other reason, the

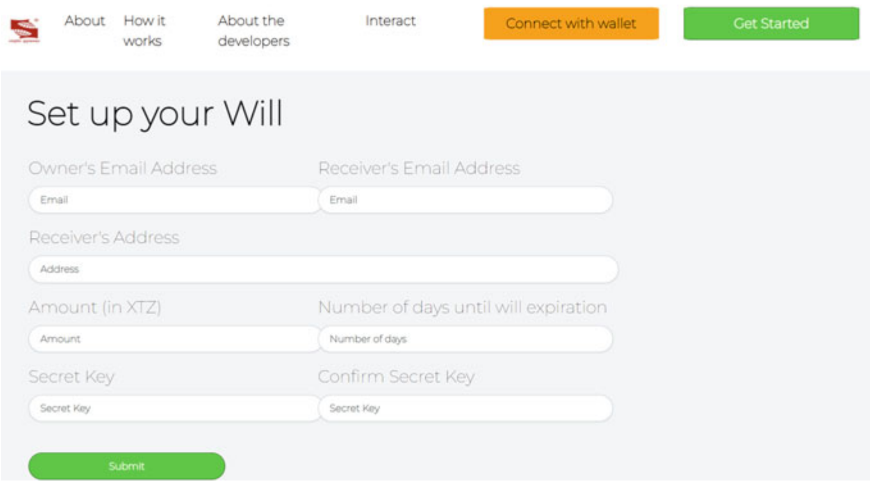


Fig. 3 Interface of online digital will be based on Web 3.0 technologies

transaction is being triggered in the smart contract and the private key along with the assets are being transferred from the owner to the beneficiary.

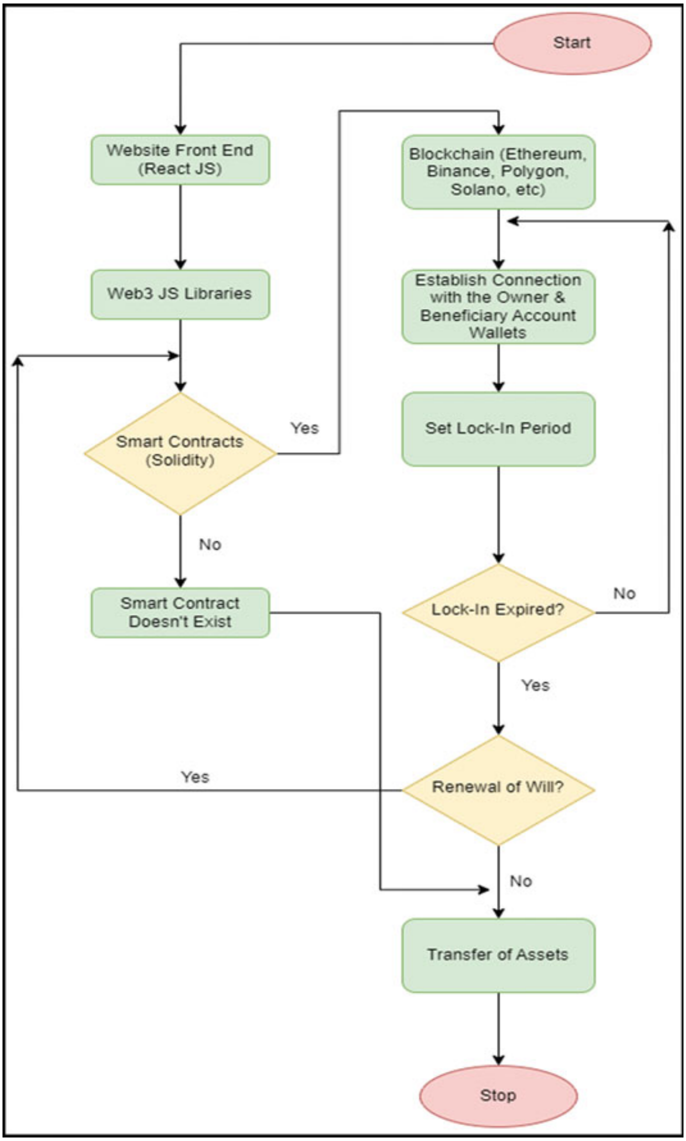


Fig. 4 Working of the full stack decentralized application for creating a digital will

5 Conclusion

Through this research, a blockchain-based system has been created for managing digital financial assets and wealth succession was presented. For the front-end web development and implementation of smart contracts, the project used web3.js, React.js, and Solidity. Gas values for transactions on the Ethereum blockchain were computed using a sample database. For the purpose of illustrating the project's implementation, a block diagram was also made.

The research contributed toward enhancing management and transfer of digital financial assets using blockchain technology which is achieved by building a website with a user-friendly interface using the React framework, which can communicate with smart contracts on the blockchain using the Web3.js library. Solidity is used to write the rules for controlling and transferring digital assets, and the blockchain provides the decentralized infrastructure that ensures the security and immutability of the transactions.

Digital financial assets like cryptocurrencies, NFTs, and digital real estate are becoming more prevalent, but the lack of a proper inheritance plan for these assets can cause significant problems. Blockchain technology provides a viable solution to this problem, as it allows for secure, transparent, and decentralized transactions. The proposed project utilizes web3.js, React.js, and Solidity to create a digital financial asset management system that is secure, transparent, and easily accessible for both owners and beneficiaries. The sample database and gas calculations provide insights into the implementation of the project on the Ethereum blockchain.

In conclusion, the suggested concept offers a workable answer to the issue of digital financial asset inheritance. However, more research is needed to optimize the gas values and transaction costs on different blockchain platforms. This initiative has the potential to significantly affect the management and inheritance of digital financial assets with further development.

References

1. Waheed N, He X, Ikram M, Usman M, Hashmi SS, Usman M (2020) Security and privacy in IoT using machine learning and blockchain: threats and countermeasures. ACM Comput. Surv **53**(6) Article 122 (November 2021), 37. https://doi.org/10.1145/3417987
2. Chu S (2021) Evaluating the financial impact of data breaches involving account credentials, 18 Sept 18. Available at SSRN https://ssrn.com/abstract=3926497 or https://doi.org/10.2139/ssrn.3926497
3. Waheed N, He X, Ikram M, Usman M, Hashmi SS, Usman M. Blockchain: threats and countermeasures. arXiv:2002.03488, https://doi.org/10.48550/arXiv.2002.03488
4. Di Lucca GA, Tortorella M (2021) Towards a framework to compare blockchain platforms for smart contracts. In: 2021 International conference on computational science and computational intelligence (CSCI), Las Vegas, NV, USA, pp 1931–1937. https://doi.org/10.1109/CSCI54926.2021.00359.
5. Ni Y, Sun B, Wang Y (2021) Blockchain-based BIM digital project management mechanism research. IEEE Access 9:161342–161351. https://doi.org/10.1109/ACCESS.2021.3130270

6. Xi R, Pattabiraman K (2022) When they go low: automated replacement of low-level functions in ethereum smart contracts. In 2022 ieee international conference on software analysis, evolution and reengineering (SANER), Honolulu, HI, USA, pp 995–1005. https://doi.org/10.1109/SANER53432.2022.00117
7. Lee WS, John A, Hsu HC, Hsiung PA (2022) SPChain: a smart and private blockchain-enabled framework for combining GDPR-compliant digital assets management with AI models. IEEE Access **10**, 130424–130443. https://doi.org/10.1109/ACCESS.2022.3227969
8. Jani S (2020). smart contracts: building blocks for digital transformation. https://doi.org/10.13140/RG.2.2.33316.83847.
9. Zheng P, Xu Q, Zheng Z, Zhou Z, Yan Y, Zhang H (2022) Meepo: multiple execution environments per organization in sharded consortium blockchain. IEEE J Sel Areas Commun 40(12):3562–3574. https://doi.org/10.1109/JSAC.2022.3213326
10. Camposano JC, Smolander K, Ruippo T (2021) Seven Metaphors to understand digital twins of built assets. IEEE Access 9:27167–27181. https://doi.org/10.1109/ACCESS.2021.3058009
11. Ross D, Cretu E, Lemieux V (2021) NFTs: tulip mania or digital renaissance? In: 2021 IEEE international conference on big data (Big Data), Orlando, FL, USA, pp 2262–2272. https://doi.org/10.1109/BigData52589.2021.9671707

The Secured System for Continuous Improvement in Educational Institutes Using Big Data Analytics

N. Kavya, S. Manasa, M R. Shrihari, T N. Manjunath, and M R. Mahesh

Abstract Data is not just a four-letter word in a present real-world environment. If we consider Big Data, it takes a new part in technology. Big Data in education using the Hadoop platform helps for evaluation of the performance of the student, visualization, and enhancing the performance of the Institution, student, and teacher. The Educational Institute faces several problems in enhancing student performance and bringing effectiveness' to student performance. The Hadoop Platform in Big Data gives an effective improvement in performance in an educational institution.

Keywords Data · Big data · Hadoop · Educational institute

1 Introduction

Nowadays, data is getting generated from different versatile sources in enormous volume. The data generated from the sources such as the internet, mobile phones, institutions, etc., is not in a structured format. The traditional computing system cannot process and analyze huge data, so we require separate tools and techniques. Here, we use Big Data analytics for the processing and analysis of data. Big Data means a large dataset that cannot be processed by traditional computing techniques, and Big Data is the set of technologies created to store, analyze, and manage bulk data.

Hadoop is an open-source framework that allows the storage and process of large datasets in a parallel and distributed manner. HDFS (Hadoop Distributed File System) is used for storing large and also for processing data having different formats such as

N. Kavya (✉) · S. Manasa · M. R. Shrihari
SJCIT, Chickballapur, India
e-mail: kavya6121976@gmail.com

T. N. Manjunath
BMSIT, Bengaluru, India

M. R. Mahesh
NCET, Bengaluru, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_17

structured, semi-structured, and unstructured data. Map Reduce techniques collect huge data and convert it into smaller results. So this technique increases the processing speed of data and helps for parallel processing of data. Since Big Data is capable of analysis of any format of data, the education domain is gaining more attention. For any organization, decision-making is more important to improve its results. In education, Big Data analytics can be used to analyze the data stored in HDFS, and better decisions are made that enhance the performance of the organization. Since Hadoop provides various tools such as a spark, hive, and HBase if we properly utilize them, we can build a website to meet all our needs which we discussed above.

2 Problem Statement

Nowadays, Big Data analytics is widely used in the education sector as mentioned by Vollenbroek in [1] for retaining a large amount of data for analysis, evaluation, visualization, decision-making, etc. of student performance in a secure manner is challenging. So the proposed system is to build a web application that is build based upon the Big Data Hadoop framework also uses map reduce for efficient performance and security is given using the Kerberos. It is the main intension of the proposing title and the proposed system works on increasing the performance of Educational Institution; it may be the student or the whole Educational Institute.

3 Related Work

Big Data in Educational Institute is playing a very crucial role since the data is huge Educational Institute. Because it contains data of students, teachers, Institute, non-staff, and also other additional data, maintaining the data in an effective and secure manner is challenging thing. The usage of the Hadoop mentioned by Alghamdi and Alghamdi in [2] helps the Educational Institute to work on the data in an effective manner and Map Reduce helps in performing the tasks very fast.

3.1 Enhancing the Performance in Educational Institute

Either the data of the student or the data of the Institute or the teacher, enhancing the performance is the important thing mentioned by Alghamdi and Alghamdi in [2] analysis of the result of the student and than Enhancing there performance or it may be data of the Institute to Enhance the admission or maybe the placement so overall it is Enhancing the Performance of the Entire Educational Institute It may be any field.

Enhancing the performance not only includes students but it also includes the teacher's performance based on student feedback or sentiment analysis. The teacher's performance should also be enhanced by providing the required things to the student as mentioned by Wang and Yang in [3].

The performance of the whole Institute matters a lot since all are looking for good continuous performance it should be increasing constantly.

3.2 Designing the Educational Institute

The Educational Institute should be designed in such a way that it should be modernized mentioned by Rafique et al. in [4], and the present technology is well implemented and also the effectiveness and the good decision-making analytics should be confident, so designing the platform in an effective manner is important and should help the student and also the parents for effective analysis students should learn and innovate the things to increase the knowledge the Big Data platform should help them and also as mentioned by Chinsomboon and Piriyasurawong in [5] the Educational Institute to analyze student results, Admissions, etc. and taking the Effective decisions on that and teachers upon the student feedback analysis interact with students.

3.3 Predictive Analysis

Predicting the student's performance may be on the student's educational background or maybe on what they feel upon feedback their strength and the weakness thing mentioned by Munshi and Alhindi in [6] should be effectively predicted and the teacher should be able to reach out to the students and provide the required things. It may be a notes, video, lectures, etc. based upon the student requirements.

Nowadays, as we can see how huge the Educational Institute data is growing considering the students, and the whole Educational Institute so the storing and accessing of such data should be in an effective manner and it should be easy to access and predict future happening things may result from Admissions, etc. should be predicted in an efficient manner.

In order to fill in any gaps in the academic curriculum that students may be experiencing, lecturers and teachers might use predictive analytics. Also, the academic modules can be altered to fit each student's learning needs.

3.4 Secured System

As a result, the term "Hadoop Security" refers to the procedure that secures the Hadoop data storage unit by delivering an impenetrable wall of protection against any cyber threat and provides authentication, authorization, auditing, and other functions.

As we know how important is the security since the institution contains some sensitive data and it is huge, so maintaining as mentioned by Shrihari et al. in [7] such data in a secure manner is challenging implementing accessing the data should be done in a secure manner.

As security came into the picture, many protocols came into the picture. Kerberos authentication protocol obtained by Yang et al. in [8] is one of the protocols which is a somewhat strongest authentication protocols for the Big Data Hadoop Map Reduce technique. As the analysis, prediction security is also important as the data is booming the world, and so providing security is also important.

4 Methodology

The method is to implement a web application based on Educational Institution data using Big Data Hadoop framework which includes Map Reduce by using various algorithms as.

- Collaborative Filtering
- Map Reduce
- Kerberos
- VADER
- Sqoop.

4.1 Collaborative Filtering

- Collaborative Filtering is a technique which is used for identifying the relationship between the pieces of data.
- This technique is frequently used in recommender systems, which identifies the similarities between the user data and items.
- These are the special algorithms designed to predict a user's interest and suggest different products in many ways to retain that interest till the end.
- This system removes data that is useless and redundant as mentioned by Rafique et al. in [9].
- This system filters out all the information intelligently, before showing it to the front users.

- This means that as shown in the below table, if Users A and B both like Product A, and User B also likes Product B, then Product B could be recommended to User A by the system.
- This model keeps track of what course contents and users like and their characteristics. In this system, we are using collaborative filtering to filter the data. Based on the student performance and feedback, the respective courses are recommended.
- Suppose if the student is very much interested in mathematics and scores distinction marks in mathematics, but he is weak at programming, then the courses related to programming are recommended to that student so they can improve their knowledge at programming, and less time is given to mathematics (Fig. 1; Table 1).

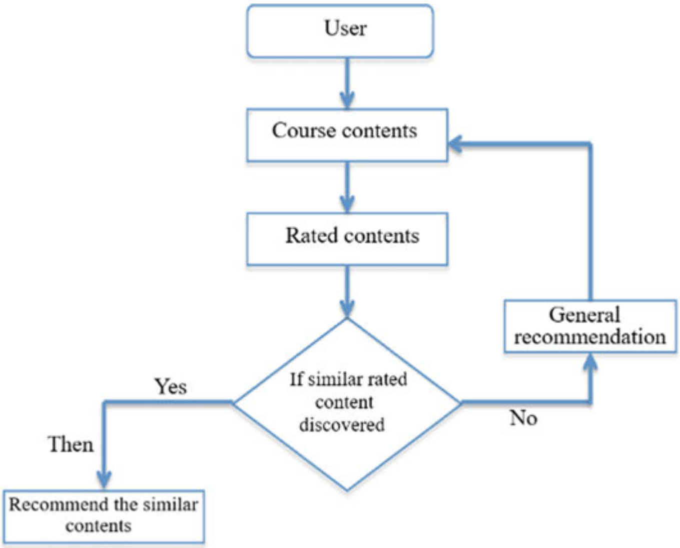


Fig. 1 Collaborative filtering

Table 1 Collaborative filtering

	Product 1	Product 2	Product 3	Product 4
User 1		1	1	
User 2	1			
User 3			1	1
User 4				1

4.2 Map Reduce

- Map Reduce is a Big Data analysis model that processes data sets using a parallel algorithm on computer clusters.
- Map Reduce is a software framework and programming model used for processing huge amounts of data.
- Map Reduce program works in two phases, namely, Map and Reduce. Map tasks deal with splitting and mapping of data while Reduce tasks shuffle and reduce the data.
- Map Reduce is parallel in nature, thus it is very useful for performing large-scale data analysis using multiple machines in the cluster.
- In this system, we have a huge amount of data. In order to process this large amount of data, we use these map reduction techniques.
- Map Reduce divides the data input data into pieces and distributes them among other computers. The input value is broken up into key-value pairs.
- In the system, student's data is stored as value pair.
- In this system, we are using Map Reduce for data processing.
- Since we are dealing with Big Data and Hadoop, data processing becomes easy using Map Reduce.
- At institutions, a large amount of data is generated every day and stored as data blocks on the cluster's numerous 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 the Big Data processing as mentioned by Siledar et al. in [10] (Fig. 2).

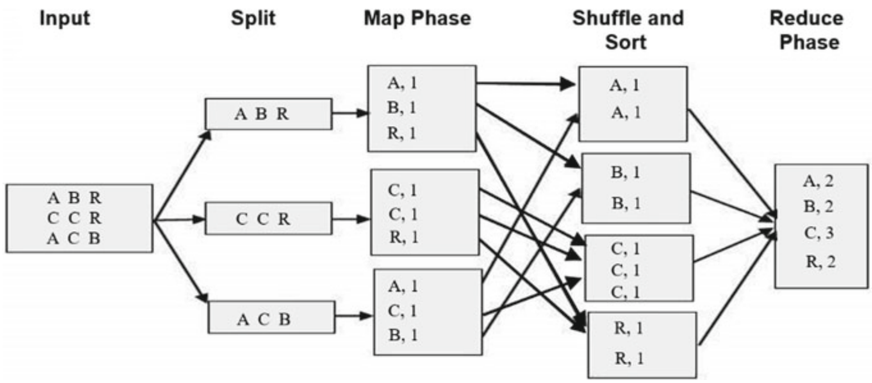


Fig. 2 Map reduce

4.3 Kerberos

- Kerberos is a computer network security protocol that authenticates service requests between two or more trusted hosts around the untrusted network, the key internet.
- Kerberos comprises 3 components Key Distribution Centre (KDC), Client User, and Server with the desired service to access.
- In this system, the Kerberos is used to protect the institutional data for untrusted users.
- KDC consists of three main components

(a) An authentication server (AS): The AS performs initial authentication when a user wants to access a service.
(b) A ticket-granting server (TGS) connects a user with the service server (SS).
(c) Kerberos database: This database stores the IDs and passwords of verified users

- Since security is the basic requirement of any application, in this system, we are using Kerberos for security.
- Kerberos keeps the password private. It will provide services only to authenticated users.
- Any user cannot guess the password and it is hard to reuse stolen tickets.

Kerberos algorithm is used to make sure that it is offering the high level of security, it must be properly set and maintained, just like security protocols. Users who want to access platforms for Big Data education can do so securely using Kerberos mentioned by Shrihari et al. in [11] (Fig. 3).

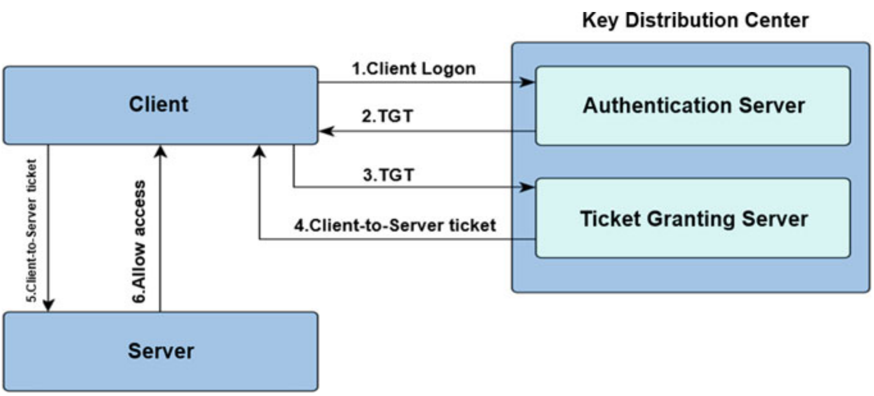


Fig. 3 Kerberos

4.4 Vader

- VADER (Valence Aware Dictionary for Sentiment Reasoning) is an NLTK module that provides sentiment scores based on the words used.
- VADER is built upon widely-known text analysis libraries.
- VADER uses a combination of both Polarity and Intensity. It also analyzes the texts that contain acronyms, emojis, and symbols.
- To know about the text, we add the scores of the word in the sentence and normalize it.
- VADER goes beyond the normal text analysis by focusing on punctuation and CAPITALIZATION of the words.
- In this system, VADER is used for the analysis of students' performance in every subject.
- Vader is an NLP algorithm used for sentimental analysis.
- Based on the feedback received from the student, the sentimental analysis is done. According to their interest, the courses are recommended to them.
- Vader can access the sentiment of any given text without the necessity of previous training.

Vader is a beautiful approach to sentiment analysis where positives and negatives and various other things are identified based on the words as mentioned by Dev et al. in [12]. Sentiment analysis is the approach to finding things based on emotions and nowadays analyzing things become important it may be in the Education field or in any field. It is one of the basic and important parts to know that sentiment analysis becomes very useful to monitor the students and knowing the fact about the student and how the students going to respond to the respective questions and so on.

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 (Fig. 4).

4.5 Sqoop

- Sqoop is a utility for transferring large amounts of data between Hadoop and outside data stores like relational databases (MS SQL Server, MySQL). Data must first be imported from various sources into Hadoop clusters in order to analyze it using Hadoop.

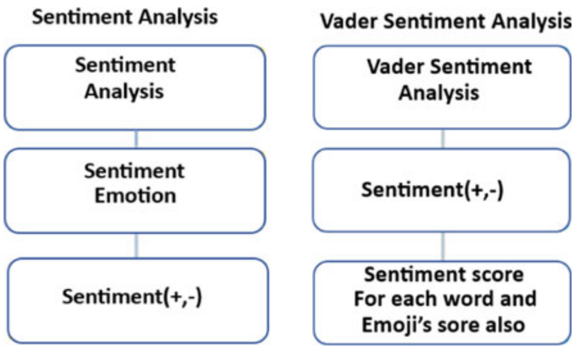


Fig. 4 VADER

- Data must first be put into Hadoop nodes from a variety of sources in order to be processed using Hadoop. However, it came out that loading data from various heterogeneous sources was a very difficult procedure.

Sqoop is useful in the Big Data realm because of a number of characteristics, including:

1. Concurrent import and export. Sqoop imports and exports data using the YARN architecture. On top of parallelism, this offers failure tolerance.
2. Import the outcomes of a SQL query. The output of a SQL query can be imported into HDFS using Sqoop.
3. A comprehensive list of RDBMS connectors. Sqoop offers interfaces for a variety of RDBMSs, including the servers for MySQL and Microsoft SQL
4. Sqoop uses the Kerberos computer network authentication protocol, allowing nodes to safely authenticate users while communicating over an insecure network.
5. Delivers both a full and incremental load. Sqoop can load a database entirely or in sections with just one query.

Sqoop is useful in the world of Big Data because of a number of features:

1. In the Hadoop cluster, Sqoop is used. HDFS receives data imports from RDBMS or NoSQL databases.
2. The data is loaded into HDFS after being divided into various formats using mappers.
3. Ensures that data is exported back into the RDBMS while maintaining its database schema.

In this system, we are using Sqoop for the migration of data from SQL to Hadoop.

5 Results and Discussions

As we know traditional systems for storing the data such as RDBMS, MySQL are lagging in the speed of loading the data and processing, it also provides backup and security.

Whereas Hadoop will load the data at high speed and as a backup since Hadoop stores the data in clusters and also in the education field where we are discussing whether the teacher wants to upload the file or wants to upload the teaching video speed is a very important thing so following example how the speed varies between RDBMS and HADOOP.

Table 2 shows how the data load time varies between RDBMS and HADOOP and how speed Hadoop is in loading the data. In a traditional system like RDMS, loading a large amount of data is somewhat difficult and takes more amount of time to load as well as to process the data. Since Hadoop in Big Data is mainly designed for large data, it stores the data in clusters so it is to access as well as to process the data as the same the below bar graph shows how the data load time differs in data load time. Fundamentally speaking, Hadoop is an open-source infrastructure software framework that enables the distributed storing and processing of enormous amounts of data, or Big Data. A master–slave architecture is used by the cluster system. As a result, huge data can be handled and stored in parallel using this design. Structured (tables), unstructured (logs, email body, blog text), and semi-structured data types can all be evaluated (media file metadata, XML, HTML) (Fig. 5).

As we now know, blocks are what HDFS uses to store data. The smallest piece of data that the file system can store is a block. These blocks are created by processing and segmenting files, and they are then disseminated throughout the cluster while simultaneously being replicated for security.

It supports the fact that Hadoop executes data analysis across the cluster's nodes in parallel. Hadoop's capacity to distribute shuffle and sort operations across a number of cluster nodes shows that it performs better than a typical RDBMS setup while doing data analytics queries.

Let us consider another example for querying the data that is processing the data and considering select count (student id) from students group. By student id, these query execution time is shown in Table 3.

Table 2 Data load time between RDBMS and Hadoop

	RDBMS	Hadoop
Data size in GB	RDBMS data load time	Hadoop data load time
0.5	2.78	1.39
1	3	0.83
2	5.48	1.53
4	10.9	3.16
6	15.64	4.62

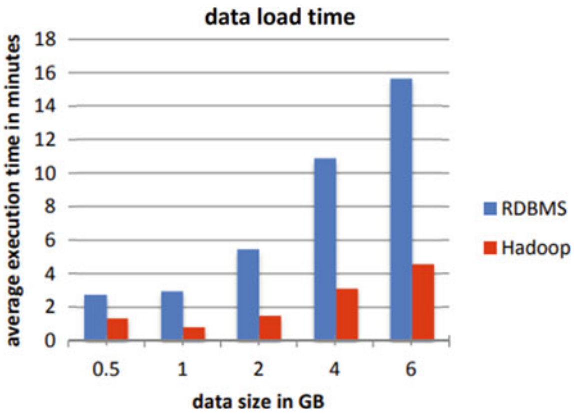


Fig. 5 Data Load time between Hadoop and RDBMS

Table 3 Average analysis time on a Hadoop cluster in minutes for hive table versus RDBMS

	RDBMS	Hadoop	RDBMS	Hadoop
Data size in GB	RDBMS select count group by query time	Hadoop select count group by query time	RDBMS select count like timestamp query time	Hadoop select count like time stamp query time
0.5	1.66	2.2	0.08	2.02
1	4.58	3.04	0.18	2.27
2	8.28	4.03	3.15	3
4	11.65	5.64	5.84	3.59
6	14.83	7.6	7.32	5.96

We can see the query time difference between Hadoop and traditional RDBMS. Hadoop is an open-source framework and it will provide various tools for processing, storing, and analyzing the data such as, Apache Spark, Hive, etc. If the amount of data is in gigabytes, the response time of RDBMS might be very quick. But if the data goes over this specific threshold, RDBMS is not very useful. On the other hand, Hadoop can process this enormous amount of data, but it is useless when we require a quick reaction to any data since Hadoop is used for analyzing, making reports, and processing any other enormous amount of data (Fig. 6).

The above graph shows how the query time differs between the Hadoop and RDBMS.

Hadoop provides enormous scalability in processing power and storage at a relatively low cost that is similar to an RDBMS. Hadoop follows horizontal scalability because it only requires adding additional machines to its cluster in order to improve performance, unlike RDBMS where data growth necessitates increasing machine configuration. Hadoop is made to handle write-once and read-many data. It is ineffective for random reading and writing of a small number of records, unlike RDBMS.

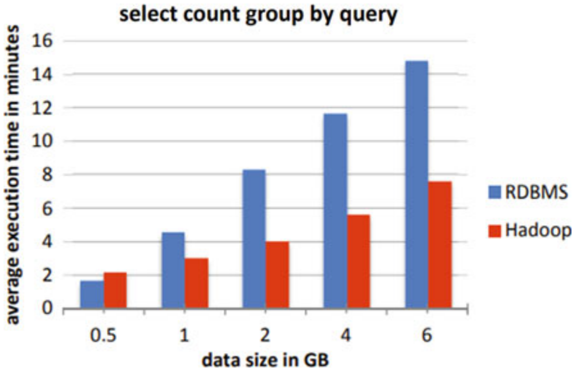


Fig. 6 Query time difference between RDBMS and Hadoop

Whereas RDBMS is excellent for structured data, Hadoop can cope with unstructured, semi-structured, or structured data. ACID features are seen in conventional databases and RDBMS. Hadoop doesn't provide any of these. Before ingesting data, RDBMS requires a schema. Schema is not necessary because Hadoop can handle any data type.

6 Conclusion

The main agenda is to build the web application using Big Data Hadoop framework which also contains the map reduce technique Enhancing the performance of the students, increasing the student results, admissions, etc. the information related to the Educational Institute is done in an effective manner and also the future prediction related to the Educational Institute should be accurate and also providing security since Educational Institute contains the sensitive data it should be secured like Student, Teachers information, overall the web application will look like it is convenient to the student to interact with the teachers and parents to interact with there children and Educational Institute. Since Hadoop Framework is used, it is open source and used for storing large amounts of data. It is easy to process the data and also accessing a large amount of data since it is stored in clusters and also various algorithms such as collaborative filtering, Vader, and Kerberos for security. Sqoop is used for migrating the database from RDBMS to Hadoop as overall if we come to a query time or loading time Hadoop is better than RDBMS.

References

1. Vollenbroek W, de Vries S, Efthymios Constantinides University of Twente, Enschede, The Netherlands NHL Hogeschool, Leeuwarden, The Netherlands (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, 20–21 Nov 2014, Glasgow, Scotland
2. 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
3. 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), Kuala Lumpur, Malaysia, pp 227–230. https://doi.org/10.1109/CISCE50729.2020.00051
4. Rafique A et al (2021) Integrating learning analytics and collaborative learning for improving student's academic performance. IEEE Access 9:167812–167826. https://doi.org/10.1109/ACCESS.2021.3135309
5. Chinsomboon M, Piriyasurawong P (2022) Big data analytics for pre-teacher preparation of higher education in Thailand. In: 2022 Research, invention, and innovation congress: innovative electricals and electronics (RI2C), Bangkok, Thailand, pp 341–346. https://doi.org/10.1109/RI2C56397.2022.9910310
6. Munshi AA, Alhindi A (2021) Big data platform for educational analytics. IEEE Access 9:52883–52890. https://doi.org/10.1109/ACCESS.2021.3070737
7. 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), Nov 2021. https://doi.org/10.1007/978-981-16-1338-8_13
8. Yang Y, Li H, Cheng X, Yang X, Huo Y (2020) A high-security signature algorithm based on kerberos for REST-style cloud storage service. In: 2020 11th IEEE annual ubiquitous computing, electronics & mobile communication conference (UEMCON), New York, NY, USA, pp 0176-0182. https://doi.org/10.1109/UEMCON51285.2020.9298140
9. Rafique A, Khan MS, Jamal MH, Tasadduq M, Rustam F, Lee E, Patrick Bernard Washington, Ashraf I. Integrating learning analytics and collaborative learning for improving student's academic performance. Ernesto Lee (elee@broward.edu) and Imran Ashraf (ashrafimran@live.com)
10. Siledar SK, Deogaonkar B, Panpatte N Pagare J (2021) Map reduce overview and functionality. In: 2021 6th International conference on communication and electronics systems (ICCES), Coimbatre, India, pp 1560–1566. https://doi.org/10.1109/ICCES51350.2021.9489056
11. Shrihari MR, Archana RA, Manjunath TN, Hegadi RS (2018) A review on different methods to protect big data sets, issue 12, p 4
12. Dev C, Ganguly A, Borkakoty H (2021) Assamese VADER: a sentiment analysis approach using modified VADER. In: 2021 International conference on intelligent technologies (CONIT), Hubli, India, 2021, pp 1–5. https://doi.org/10.1109/CONIT51480.2021.9498455

A Brief Review of Text Summarization Approaches

Surbhi Jangid and Sanjay Kumar Jain

Abstract The need for Text Summarization is growing significantly because a huge amount of data is present on the internet. Manual text summarization takes effort and time, therefore, automating text summarization is desirable. Much research work has already been carried out in this area and is still progressing. This paper briefly reviews different methods, datasets, and evaluation metrics for text summarization. The main focus is on neural networks and pre-trained transformer language models.

Keywords Natural language processing · Transformers · Text summarization

1 Introduction

There is a huge amount of text data available online; sometimes we need the data in summary form so that we get useful information from the data in less time and effort. This we can achieve by text summarization, therefore, there is an increase in research demand in the area of automatic text summarization. Summarization of text is the process of generating a short and coherent summary while looking after the important information and general meaning. Text summarization is used in applications like question-answering, information extraction, information retrieval, email summarization, book summarization, sentiment summarization, etc. Figure 1 exhibits text summarization system. Developers still want a text summarization that can produce summaries which (1) include all the important information. (2) Not include repeated data. (3) Is readable and cohesive to the user. In 1958, Luhn's [1] mention in his paper about text summarization. After that many research work happens. The objective of the paper present is to survey the different methods of text summarization and mainly focus on recent pre-trained-based transformer models. First, we introduce the common background of text summarization in (Sect. 2) which includes the (1) Classification of Text Summarization, (2) Summarization Methods, and (3)

S. Jangid (✉) · S. K. Jain
Department of Computer Engineering, NIT Kurukshetra, Thanesar, India
e-mail: surbhijangid786@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_18

Fig. 1 Text summarization system

Language Model. In (Sect. 3), we examine text summarization process with encoder-decoder architecture. (Sect. 4) presents mechanism to improve text summarization system. (Sect. 5) presents an overview of an existing dataset for text summarization. (Sect. 6) presents different methods to evaluate text summarization and compare the performance of model through evaluation metric. At last, in (Sect. 7) conclusion of paper.

2 Background

2.1 *Classification of Text Summarization*

Table 1 exhibits a classification of text summarization systems. El-Kassas and Wafaa in their survey mention these classifications [2]. We mainly focused on criterion based on the approach used. Text summarization has three main approaches as: (1) extractive, (2) abstractive, and (3) hybrid. Extractive approach is the simplest way to generate summary. Here we just need to select the important words, or the sentence from the input text, and combine it to form a summary. Most of the surveys focused on extractive summarization as Gupta, Vishal, and Gurpreet Singh Lehal [3], and Nazari and Mahdavi [4]. Abstractive approach uses the method where it first generates the intermediate form from the input text and then from intermediate, it uses sentences and words which are different from input text to generate the summary. Lin, Hui, and Vincent Ng. done a survey on abstractive text summarization [5]. Hybrid approach combines both extractive and abstractive approaches. Thus, it combines the advantages and disadvantages of both approaches thus overall summarization is improved. Its disadvantages are that it generates less quality abstractive summary if the extractive phase is not done properly because it depends on the extractive phase.

2.2 *Summarization Methods*

We categorized text summarization techniques as: (1) Statistical, (2) Machine learning, and (3) Deep learning methods. In statistical method, we find the important

Table 1 Classification of text summarization system

S. No.	Criteria (based on)	Types of text summarization system	
1	Input document size	1	Single document
		2	Multi document
2	Summarization algorithm used	1	Supervised
		2	Unsupervised
3	Language of generated summary	1	Monolingual
		2	Multilingual
		3	Cross-lingual
4	Nature of the output summary	1	Query
		2	Generic
5	Summarization domain is used for	1	General
		2	Domain specific
6	Summary generated type	1	Sentence level
		2	Full summary
		3	Headlines
		4	Highlights
7	Approach used	1	Extractive
		2	Abstractive
		3	Hybrid

words and sentences from the input data based on the statistical features and combine them to form a summary.

Many NLP statistical techniques are there to determine physical features like title word, term frequency, sentence position and length, TF-IDF, cue words, etc. Machine learning methods are used to train the model for certain tasks, like regression or classification when input data is provided. A few of the important machine learning methods are as Naive Bayesian Classifier, Hidden Markov Model, Methods based on Graph, etc. Deep Learning Techniques in text summarization [6] start growing with the use of neural network along with NLP, there is an opportunity to improve summarization in both extractive and abstractive. In deep neural networks, we can analyze complex problems also. Here, the input to one layer serves as the output to the next layer, so the depth is determined by the number of layers.

2.3 Language Model

A language model is trained on a large corpus of text data to learn the probability distribution of words and their relationships to each other in the language. This enables the model to understand how particular terms are used as well as how the language is written generally. Qiu et al. [7] in his paper categorized these language models as first- and second-generation models. These are pre-trained models; the first-generation model is word embedding in which words are representation as vectors for example word2vec, GloVe, etc.; the second-generation model is contextual encoders which represents the context of a word with respect to the sentence for example BERT, GPT, etc.

3 Text Summarization Process with Encoder-Decoder Architecture

Here, we cover encoder-decoder architecture and how we might enhance the quality of summaries by using different neural networks as encoders and decoders. Abstractive text summarization using neural networks and seq-2-seq models shows good results in generating summaries. Here, the purpose of the model is to generate a fixed length vector representation of the input sequence to fixed length output sequence. Here, the architecture consists neural network encoder layer and a neural network decoder layer. The encoder input sequence is usually transformed into a numerical representation, such as a sequence of embedding, and then passed through the multiple layers of neural network such as GRU or LSTM, to capture the sequence information. The decoder generates the output sequence one token at a time. During each decoding step, the decoder network uses the previous output token and the fixed-length vector generated by the encoder as inputs. The choice of encoder-decoder architecture offers an option to create one's encoder and decoder using regular RNN [8] /LSTM/GRU or bidirectional RNN/LSTM/GRU [9], Transformer, and BART. Our main focus is on transformer models.

Transformers in 2017, Google unveiled The Transformer, a breakthrough for tasks using sequence learning. "The Illustrated Transformer," by Alammar [10] discusses the transformer in straightforward terms. The classic study "Attention is all you need" by Vaswani, Ashish, et al. discusses the transformer in text summary. [11] The transformer is a model that increases speed by paying attention on different parts of the input sequence while computing its output. Unlike traditional models which process input sequentially, transformer architecture processes all the input tokens in parallel, which makes it more efficient for long sequences. In the transformer architecture, the input sequence is first embedded into a continuous vector space using an embedding layer. Then, multiple layers of self-attention and feedforward neural networks are applied to the embedded sequence to transform it into a higher-level representation.

BERT/GPT (Bidirectional Encoder Representation from Transformer/generative pre-trained Transformer) BERT is a pre-trained language model developed by Google in 2018 to perform a variety of NLP tasks, similarly GPT is also a pre-trained language model which Open AI introduced in 2018. These models are trained on a massive amount of textual data present on the internet, using a self-supervised learning approach. Unlike GPT, where training is unidirectional, BERT may undertake bidirectional training. Both systems rely on transformers; BERT is a stack of encoder transformer while GPT is a stack of decoder transformer.

BERT takes the transformer architecture a step further by introducing a new pre-training objective called Masked Language Modelling (MLM) and Next sentence prediction (NSP). In MLM, a certain percentage of the input tokens are randomly masked out, and the model is trained to predict the missing tokens based on the surrounding context. This allows BERT to capture bidirectional context. In the NSP training, the model is given two sentences as input, and it learns to predict whether the second sentence is the next sentence in the document or whether it is a random sentence from another document. This allows BERT to capture not only the meaning of individual words but also the relationship between sentences.

BART (Bidirectional and Autoregressive Transformers) is a pre-trained language model developed by Facebook AI Research (FAIR) in 2019. It is a variant of the transformer architecture that combines both auto-regressive and bidirectional training objectives to improve the quality of text generation. BART large model uses 12 layers of each encoder and decoder compared to the base model's 6 levels. BART is pre-trained on a combination of denoising autoencoding (DAE) and Masked Language Modelling (MLM) objectives. BART has been shown to achieve state-of-the-art results as its ability to capture both local and global dependencies in the input sequence, as well as its ability to handle noise and variations in the input data.

4 Mechanism to Improve Text Summarization System

There are many issues when summaries are generated in text summarization process. To reduce these issues and to improve the summaries quality, we add mechanisms to the basic neural encoder-decoder architecture as attention mechanism earlier is used for neural machine translation later it is applied on text summarization. The encoder intermediate states are used to create context vectors when attention mechanism is applied. Context vectors are used to find the most relevant information when the system is generating the summary, a copying mechanism which can be used when specific input sequence components need to be copied in the output sequence. For example, some facts need to be the same as in input then we need copying mechanism. Coverage mechanism when output sequence generates with reputation of sentences then with the help of coverage mechanism we remove or reduce the repetition problem. Pointer Generator Mechanism the issue of out-of-vocabulary (OOV) terms is resolved by this technique. This mechanism is able to generate new words using a generator while also having the ability to replicate words or facts using a pointer. Distraction mechanism proposed by Chen et al. [12] is to distract the model in such a way that it traverses throughout the documents instead of focusing on a particular area so that it grasps the overall meaning of summarizing a document.

5 Dataset

Dernoncourt et al. [13] mention an overview of datasets in their paper that is used for text summarization as shown in Table 2. Some of the most commonly used datasets are Document Understanding Conference (DUC) datasets that mostly contain news articles. This dataset is based on English language and presented from DUC (2001–2007). LCSTS (Large Scale Chinese Short Text Summarization) dataset is based on Chinese language mostly blogs and contains 2,400,591 number of documents. CNN-corpus dataset is an English language based dataset that is based on news domain. CNN dataset is not supportive of multi-document summarization. Gigaword dataset is also based on a news domain that contains 4,111,240 number of documents. The CNN / Daily Mail dataset contains 312,084 number of documents. It is a widely used dataset that is not supportive to multi-document summarization.

6 Evaluation of Summary Generated

Evaluation metrics in text summarization are used to assess the quality of a summarization system by comparing the output of the system to a reference summary. Text summarization evaluation can be done manually or automatically. In manual evaluation, summaries are evaluated by humans who consider readability, non-redundancy, grammar, referential clarity, topic coverage, conciseness, and focus among other quality indicators. We need an automatic evaluation of summaries because doing so manually requires reading both the original documents and the summaries, thus it's a time taking task. Some commonly used evaluation metric in text summarization are: ROUGE (Recall-Oriented Understudy for Gisting Evaluation): The most used assessment metric in text summarization systems is the ROUGE metric that measures the overlap between the system-generated summary and one or more reference summaries in terms of n-gram co-occurrence. ROUGE scores are calculated for various n-gram sizes (1–4) and can be reported as F_1 scores, precision, or recall. Another used evaluation metric is BLEU (Bilingual Evaluation Understudy). BLEU scores are calculated based on the n-gram precision of the system-generated summary, and can be reported as a single score or as multiple scores for different n-gram sizes. METEOR (Metric for Evaluation of Translation with Explicit Ordering) metric is based on the idea of harmonic mean of precision and recall, with a focus on capturing semantic similarity between the system-generated summary and the reference summary.

Table 2 shows the ROUGE score on text summarization to compare various models on the basis of their encoder- decoder architecture. The table contains Columns 2 and 3 as name of the encoder and decoder used for text summarization. Column 4 presents the mechanism used to handle specific problems, and Column 6 presents the dataset used. Mostly we select research papers which used CNN/Daily Mail dataset so that we can easily compare the models which used the same dataset. Table 2

Table 2 Comparison of different encoder-decoder models used in text summarization on the basis of ROUGE score

S. No.	Name of the encoder model	Name of the decoder model	Mechanism	Dataset	Evaluation rouge score	Year	References
1	RNN	LSTM	Attention	DUC2004, GIGAWORD	28.97	2016	[14]
2	Bi-RNN	RNN	Attention, Copying	LCSTS	35.00	2016	[15]
3	Bi-GRU	GRU	Attention, Pointer-Generator	CNN/Daily Mail, DUC2004, GIGAWORD	35.46	2016	[16]
4	Bi-GRU	LSTM	Attention, Distraction	CNN/Daily Mail, LCSTS	27.1	2016	[12]
5	Bi-LSTM	LSTM	Attention, Pointer-Generator, Coverage	CNN/Daily Mail	39.53	2017	[17]
6	Bi-LSTM	LSTM	Attention, Pointer-Generator, Coverage	CNN/Daily Mail	41.22	2018	[18]
7	GRU	GRU	Attention, Pointer-Generator, Coverage	CNN/Daily Mail NEWSROOM	37.60	2018	[19]
8	Trans-former	Trans-former	Attention	GIGAWORD, DUC2004, LCSTS	38.72	2019	[20]
9	Transformer + BERT	Trans-former	Attention	CNN/Daily Mail	43.85	2020	[21]
10	Transformer + BERT	Trans-former	Attention, Copying	CNN/Daily Mail, NEWS ROOM	44.79	2020	[22]

second last column presents the year of the reference paper we selected which is from 2016 to 2020 so that we can find the evaluation of all recently used models.

The highest score achieved is 44.79 by using the transformer-based encoder-decoder model. The ROUGE score is 41.22 while using the encoder as Bi-LSTM and the decoder as LSTM. The ROUGE score is 38.72 when the transformer is used as an encoder-decoder, therefore, it can be concluded that employing BERT in addition to the transformer and adding some mechanism lead in an overall improvement of the ROUGE score.

Additionally, table draws the conclusion that whereas earlier encoder-decoder models were mostly based on unidirectional RNN/LSTM, with the advancement of research, bidirectional RNN/LSTM was utilized, which improves ROUGE score. Transformer-based architecture is used, according to recent research articles from 2019 and 2020, to describe abstractive text summarization systems. The transformer is highly effective at training because there are no RNN layers, hence backpropagation across time is not required. Hence, the transformer performs better than any other seq-2-seq model for abstractive text summarization. To get better outcomes with text summarization, pre-trained models were effectively used. Seq2seq language creation has seen great success using BART pre-training techniques.

7 Conclusion

We have seen that due to large amount of data, text summarization plays a very important role in saving user's time. In this paper, we reviewed from history to current state models that are used in text summarization process. The current famous summarization is abstractive summarization after the introduction of neural network as it has the ability to improve generate summaries. Abstractive text summarization gained importance after the transformer was published. As a result of this, powerful pre-trained language models like BART, GPT, BERT, etc. are developed.

However, there are number of drawbacks to the current summarization system; those are the future challenges in the field of text summarization as mentioned, few of them are the need for different and large datasets, need for more evaluation metric, flexible stopping criteria during the summary's generation process, and personalized summary generation, etc.

References

1. Luhn HP (1958) The automatic creation of literature abstracts. IBM J Res Dev 2(2):159–165
2. El-Kassas WS et al (2021) Automatic text summarization: a comprehensive survey. Expert Syst Appl 165:113679
3. Gupta V, Lehal GS (2010) A survey of text summarization extractive techniques. J Emerg Technol Web Intell 2(3):258–268

4. Nazari N, Mahdavi MA (2019) A survey on automatic text summarization. J AI Data Mining 7(1):121–135
5. Lin H, Ng V (2019) Abstractive summarization: a survey of the state of the art. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, no 01
6. Rush AM, Chopra S, Weston J (2015) A neural attention model for abstractive sentence summarization. arXiv preprint arXiv:1509.00685
7. Qiu X et al (2020) Pre-trained models for natural language processing: a survey. Sci China Technol Sci 63(10):1872–1897
8. Lopyrev K (2015) Generating news headlines with recurrent neural networks. arXiv preprint arXiv:1512.01712
9. Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. IEEE Trans Signal Process 45(11):2673–2681
10. Jay Allamar's Homepage https://jalammar.github.io/illustrated-transformer. 02 Apr 2023
11. Vaswani A et al (2017) Attention is all you need. Adv Neural Inf Process Syst 30
12. Chen Q et al (2016) Distraction-based neural networks for document summarization. arXiv preprint arXiv:1610.08462
13. Dernoncourt F, Ghassemi M, Chang W (2018) A repository of corpora for summarization. In: Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018)
14. Chopra S, Auli M, Rush AM (2016) Abstractive sentence summarization with attentive recurrent neural networks. In: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies
15. Gu J et al (2016) Incorporating copying mechanism in sequence-to-sequence learning. arXiv preprint arXiv:1603.06393
16. Nallapati R et al (2016) Abstractive text summarization using sequence-to-sequence rnns and beyond. arXiv preprint arXiv:1602.06023
17. See A, Liu PJ, Manning CD (2017) Get to the point: Summarization with pointer- generator networks. arXiv preprint arXiv:1704.04368
18. Gehrmann S, Deng Y, Rush A (2018) Bottom-up abstractive summarization. arXiv preprint arXiv:1808.10792
19. Shi T et al (2021) Neural abstractive text summarization with sequence-to-sequence models. ACM Trans Data Sci 2(1):1–37
20. Duan X et al (2019) Contrastive attention mechanism for abstractive sentence summarization. arXiv preprint arXiv:1910.13114
21. Liu Y, Lapata (2019) Text summarization with pretrained encoders. arXiv preprint arXiv:1908.08345
22. Saito I et al (2020) Length-controllable abstractive summarization by guiding with summary prototype. arXiv preprint arXiv:2001.07331

Transfer Learning-Based Nutrient Deficiency Prediction Model for Sugarcane Crop

R. Madhumathi, S. Raghavendar, B. Jegan, M. Naveenganesh, and T. Arumuganathan

Abstract Sugarcane is a crop of economic importance in many countries including India. This sugarcane crop effectively fixes solar energy and produces around 80 tons per hectare each year. An effective prediction model is necessary to increase total crop productivity because soil nutrient testing methods including colorimetry, optical spectroscopy, soil test kits, and NPK sensors are not practical in real-time applications. In this paper, a model has been created utilizing a pre-trained Deep Learning (DL) model to identify nutritional deficiencies in the sugarcane crop such as Densenet121, VGG16, VGG19, ResNet50, and ResNet152V2. The collected sugarcane images have been used to train, test and validate the model. Findings show that DenseNet121 and ResNet152v2 model gives the greatest accuracy of 93% for 15 epochs and batch size of 16.

Keywords Agriculture · Deep learning models · Transfer learning · Nutrient deficiency prediction

R. Madhumathi (✉) · S. Raghavendar · B. Jegan · M. Naveenganesh
Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India
e-mail: madhumathi.r@srec.ac.in

S. Raghavendar
e-mail: raghavendar.2001154@srec.ac.in

B. Jegan
e-mail: jegan.2001101@srec.ac.in

M. Naveenganesh
e-mail: naveenganesh.2001129@srec.ac.in

T. Arumuganathan
ICAR-Sugarcane Breeding Institute, Coimbatore, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_19

1 Introduction

Agriculture is an important sector in the country's development as we rely on it to produce food for our population and its sustainability. This sector is constantly evolving and changing, due to the ever-growing demand for food. It is anticipated that by the year 2050, there will likely be a significant rise of the demand for agricultural goods due to rise in population. To support this demand, new approaches need to be explored in producing the crops and food products that are more nutritious and less harmful to the environment. One such potential application is to produce crops using less land. Some examples of this type of cultivation include organic farming and precision farming. Hydroponics and Aeroponics are other examples of growing crops without soil. These newer methods are more environmentally friendly because they do not require large areas of land. Also, in the current trend, AI-based automated agricultural methods are getting popular and many such companies like Amazon started investing in agriculture. Agriculture is being innovated in many ways to produce more crops with lesser resources, and deep learning is one of the emerging field in agriculture [1, 2].

Nutrient is considered as an important factor that contribute to increased sugarcane production. To quantify the nutrients present in the field, farmers use various techniques in agriculture. Soil testing is one of the most commonly used techniques in agriculture to determine nutrient levels. By analyzing soil samples, farmers can determine the pH level, organic matter content, and nutrient deficiencies within their fields [3, 4]. Another technique that farmers use to find nutrients in agriculture is plant tissue analysis. Additionally, farmers can also use precision agriculture technologies such as GPS mapping and drones to identify areas of the field that may need more or less of certain nutrients and to apply fertilizer accordingly, ensuring that the crops receive optimal nutrition and irrigation [5]. Through these techniques, farmers can develop specific nutrient management plans to ensure that their crops receive the necessary nutrients of required quantity at proper time. Overall, the use of various techniques in agriculture to find essential nutrients demonstrates the importance of precise and accurate nutrient management for optimal crop growth and sustainability.

2 Deep Learning Models Used in Agriculture

In this work, a system for estimating the nutritional content of Valencia orange leaf using machine learning method was developed [6]. In this study, variety of pre-trained Convolution Neural Network (CNN) architectures with various optimization strategies to detect olive disease [7]. The tests revealed that the MobileNet architecture and RMSprop algorithms outperformed alternative combination of the pre-trained models.

In this paper, five ImageNet topologies are used to generate a collection of pre-trained models for AgriNet and trained dataset collection of 1,60,000 images [8].

The maximum classification accuracy (94%) and F1-score (92%) were attained by AgriNet-VGG19. In this paper, a deep learning framework is created using data from remote sensing to classify water supply systems at the provincial scale [9]. A U-Net topology with a Resnet-34 infrastructure is used and provides an average accuracy of 85%, the model is effective at segmenting images based on spatial data.

In this investigation, the YOLO-V5 model worked effectively with an accuracy of 92.34% to identify period greenery in images captured by a drone [10]. In this study, transfer learning models are used to identify signs of nutritional shortage for sugar beet crops [11]. The best results were obtained using EfficientNetB4's, with Top-1 accuracies of 98.65% and 98.52% on both datasets. This article suggests a supervised neural model to estimate SOM content using optical and near-infrared spectroscopy, and DenseNet approach for reducing the quantity of data pre-treatment [12].

This study proposes an intelligent classification system for pH and soil nutrients using the weighted voting ensemble approach [13]. About the partitioning of pH and soil nutrients, this model's accuracy is 0.9281 and 0.9497 respectively. In this study, transfer learning model is employed to identify nutritional stress in the plant [14]. To increase classification accuracy, three DL architectures were used and a dataset is created with 880 photos of magnesium and calcium deficiency in tomato plants gathered from the greenhouse. The maximum accuracy of 99.14% was achieved by the ensembled model of VGG16 with SVM classifier.

In this paper, a deep learning model with an accuracy of 98.8% is proposed to optimize the expenditure cost and increase the profit and productivity of the cotton [15]. In this study, a system that automatically identifies and categorizes biotic stresses on paddy crop is developed [16]. For the study, three of the most prevalent and harmful biotic stresses-leaf blast, brown spot, and hispa are examined. According to the test findings, the ResNet-50 model performs better than the other pre-trained CNN models, obtaining an accuracy of 92.61%. This study describes a comprehensive technique for efficiently segmenting the damaged vegetation in RGB images [17]. The model is trained using a variety of methods, including infrared-dark channel ratio, infrared-dark channel subtraction, and others.

In [18], weed identification of a bell pepper field was undertaken using pre-trained deep learning architectures and produced better accuracy among the trained models. In this paper, an ensembled model is proposed to classify hot pepper pest and disease images were used to train 8 pre-trained DL models [19]. In this research, a model is proposed to predict the type of disease that will initially affect tomato and grape leaves [20]. The VGG model proved successful in identifying Multi-Crops leaf disease with good accuracy.

3 Transfer Learning Model for Nutrient Deficiency Prediction

In this work, a nutrient deficiency prediction model for sugarcane crop is proposed using the pre-trained deep learning models such as DenseNet121 and ResNet152v2.

3.1 Data Collection

The images have been collected from the sugarcane farm of ICAR–Sugarcane Breeding Institute which is located at Veerakeralam, Coimbatore. A collection of around 250 images has been captured to train the model.

3.2 Framework for Evaluation of Nutrient Deficiency

After the collection of images, it undergoes a series of steps which is shown in Fig. 1. The image dataset has been improved using preprocessing methods like cropping, enhancing, and resizing, which are then followed by a data augmentation process. Then the dataset can be amplified by making small adjustments to the original data or by creating new data using machine learning models.

In this system, a deep learning model has been trained on a large and diverse dataset with Densenet121 and ResNet152v2 frameworks which can be fine-tuned to achieve good accuracy. The transfer learning methodology has been used to pre-train the model, which serves as a good starting point since it has already learned low-level features such as edges and corners that can be useful for nutrient deficiency prediction.

3.3 Algorithm for Nutrient Deficiency Prediction

Step 1: Collect sugarcane crop dataset with nutrient content labels (N, P, K).
Step 2: Convert RGB to grayscale images.
Step 3: Normalize grayscale images.
Step 4: Apply data augmentation techniques such as rotation, scaling, and flipping.
Step 5: Split dataset for training, testing, and validation.
Step 6: Use pre-trained CNN model such as Densenet121, Resnet152v2 and train the model.
Step 7: Use softmax activation function to obtain class probabilities.
Step 8: Define a loss function such as categorical cross-entropy.

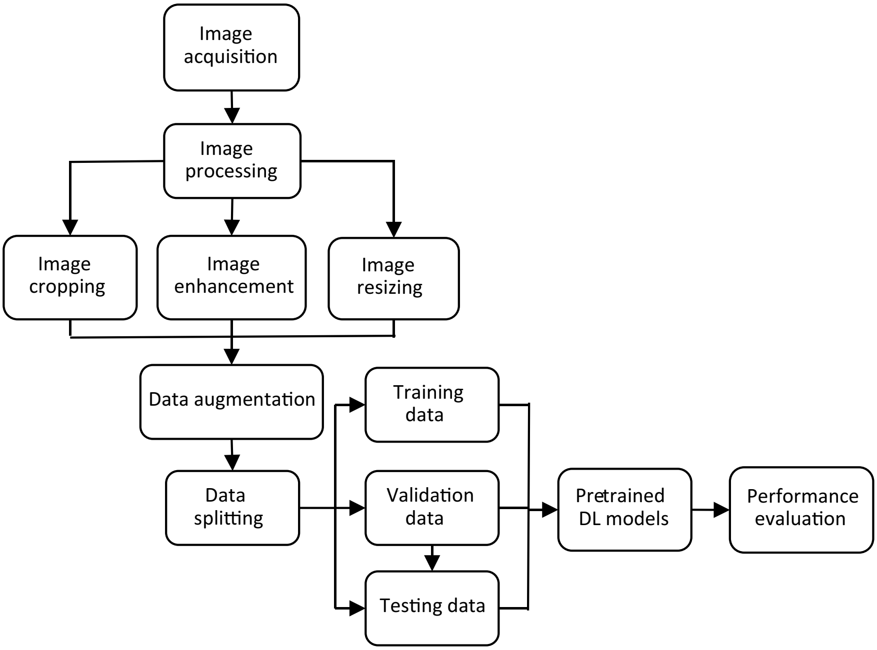


Fig. 1 Flow diagram of the nutrient deficiency model

Step 8.1: Measure the difference between the predicted and actual nutrient content labels.

Step 9: Use optimization algorithm such as Adam to minimize the loss function and update the model parameters.
Step 10: Train the model on the training set for a specified number of epochs, adjusting the learning rate and other hyperparameters as needed.
Step 11: Test and evaluate the model based on performance metrics such as confusion matrix, accuracy, etc.

4 Results

A collection of 250 sugarcane leaf images has been captured from the ICAR—Sugarcane Breeding Institute, Coimbatore. Based on the signs of nutrient deficiencies, the images have been divided into three categories namely nitrogen, phosphorus, and potassium. The leaves with low levels of nitrogen will be a pale yellowish-green color, those with low levels of phosphorus will be reddish, reddish-violet, or violet in color, and those with low levels of potassium will be a bluish-green color which is shown in Table 1. For the model's train, test, and validation phases, these images

Table 1 Leaf Nutrient Deficiency Based on Color Symptom

Nutrient type	Color deficiency symptom
Nitrogen	Pale yellowish green
Phosphorus	Reddish violet (or) reddish brown
Potassium	Bluish green

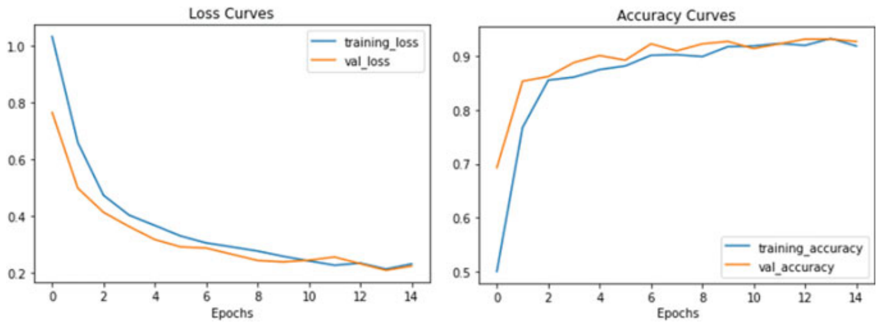


Fig. 2 Accuracy and loss of Densenet121 model for each iteration

are divided in the ratio 75:20:5. Here, nutritional deficiency has been predicted using the Densenet121 and ResNet152v2 architecture.

Densenet121 and Resnet152v2 have been implemented using tensorflow platform and demonstrated greater accuracies of about 90%. However, when compared to Densenet121 model, the Resnet152v2 model has a poor level of stability due to overfitting of data. The DenseNet121 model yields an accuracy of about 93% with the least amount of loss at a learning rate of 0.001 value and furthermore epoch of 15, as illustrated in Fig. 2.

Additionally, the model accuracy is accessed based on the performance of confusion matrix in the validation and test phases when it comes to image prediction. Figure 3 illustrates the confusion matrix for NPK prediction.

The dataset has been labeled as low, medium, and high for each nutrient and tested in the model. Figures 4, 5, and 6 depict the performance of the model for each nutrient level.

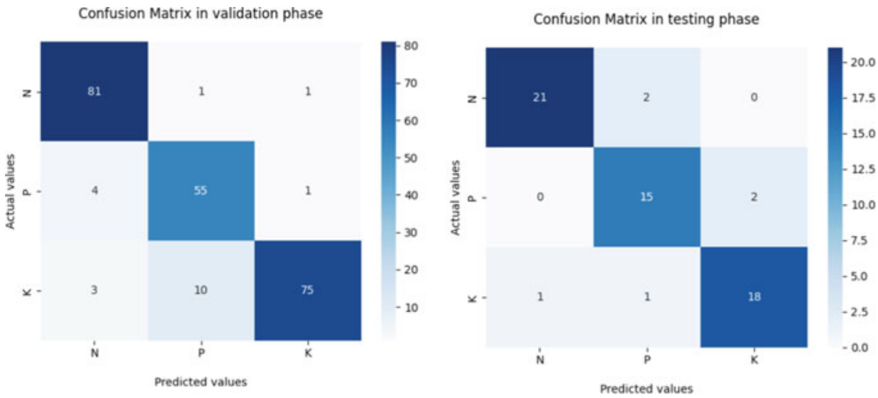


Fig. 3 Confusion matrix of NPK prediction

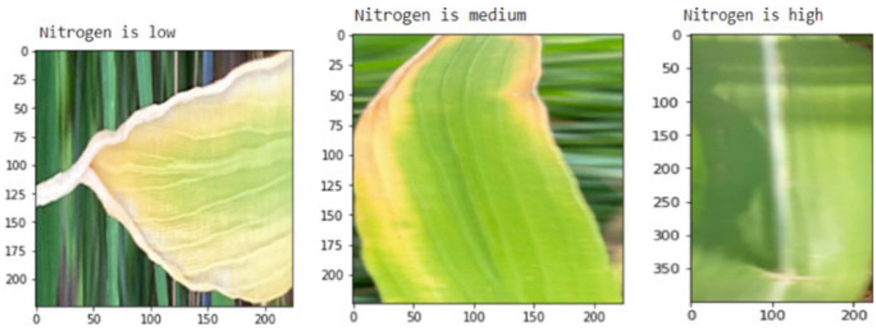


Fig. 4 Model prediction for low, medium, and high levels of nitrogen

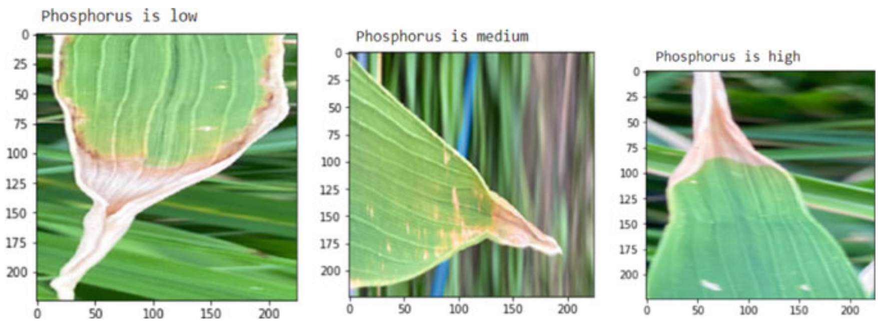


Fig. 5 Model prediction for low, medium, and high levels of phosphorus

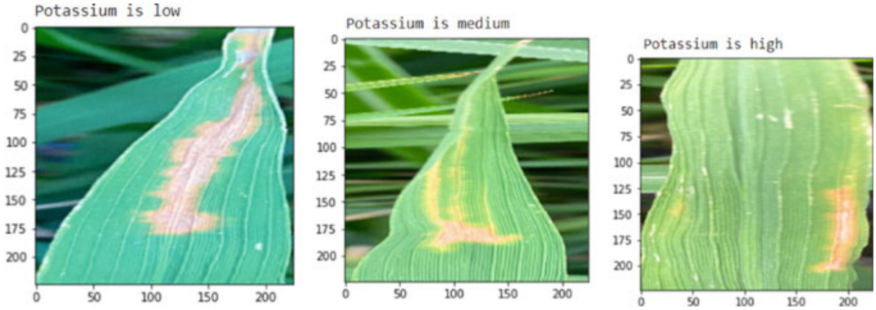


Fig. 6 Model prediction for low, medium, and high levels of potassium

5 Conclusion

A transfer learning method has been used to classify the nutrients deficiency in leaves. Based on the color symptoms in the leaves, the Densenet121 architecture has been employed to predict the nutritional deficiency. The outcome of the model demonstrates an accuracy of about 93% in training phase and is also predicted well in testing phase. Additionally, the model has the ability to classify live pictures. In this work, only macronutrients of the sugarcane crop are classified with this model. In future, models will be proposed which classify all other agricultural crops as well as micronutrients like Boron (B), Iron (Fe), and Magnesium (Mg), etc.

Acknowledgements The authors acknowledged the financial grant provided under SERB-TARE scheme by DST—Government of India and Sri Ramakrishna Engineering College, Coimbatore collaborated with ICAR—Sugarcane Breeding Institute, Coimbatore for conducting this study.

References

1. Padarian J, Minasny B, McBratney AB (2019) Using deep learning to predict soil properties from regional spectral data. J Geoderma Regional 16 (Elsevier)
2. Madhumathi R, Arumuganathan T, Shruthi R, Raghavendar S (2022) Soil NPK prediction using multiple linear regression. In: 8th IEEE international conference on advanced computing and communication systems (ICACCS). IEEE, pp 542–546
3. Mythili R, Tripathi MA, Pal N (2019) IoT based smart farm monitoring system. J Recent Technol Eng 8(4):5490–5494 (Blue Eyes Intelligence Engineering and Sciences)
4. Mutyalamma AV, Yoshitha G, Dakshyani A, Padmavathi BV (2020) Smart agriculture to measure humidity, temperature, moisture, pH and nutrient values of the soil using IoT. J Eng Adv Technol 9(5):394–398
5. Puengsungwan S (2020) IoT based soil moisture sensor for smart farming. In: International conference on power, energy and innovations (ICPEI), Chiangmai, Thailand. Blue Eyes Intelligence Engineering and Sciences, pp 221–224
6. Osco LP, Ramos APM, Pinheiro MMF, Moriya EAS (2020) A machine learning framework to predict nutrient content in Valencia-orange leaf hyperspectral measurements. J Remote Sens 12(6):1–21 (MDPI)

7. Raouhi EM, Lachgar M, Hrimech H, Kartit A (2022) Optimization techniques in deep convolutional neuronal networks applied to olive diseases classification. J Artif Intell Agric 6:77–89 (Elsevier)
8. Sahili Z, Awad M (2022) The power of transfer learning in agricultural applications: AgriNet. J Front Plant Sci Front
9. Raei E, Asanjan AA, Nikoo MR, Sadegh M, Pourshahabi S, Adamowski JF (2022) A deep learning image segmentation model for agricultural irrigation system classification. J Comput Electron Agric 198:1–11 (Elsevier)
10. Jintasuttisak T, Edirisinghe E, Elbattay A (2022) Deep neural network-based date palm tree detection in drone imagery. J Comput Electron Agric 192:1–11 (Elsevier)
11. Espejo-Garcia B, Malounas I, Mylonas N, Kasimati A, Fountas S (2022) Using EfficientNet and transfer learning for image-based diagnosis of nutrient deficiencies. J Comput Electron Agric 196:1–11 (Elsevier)
12. Xu Z, Zhao X, Guo X, Guo J (2019) Deep Learning application for predicting soil organic matter content by VIS-NIR spectroscopy. J Comput Intell Neurosci 2019:1–11 (Hindawi)
13. Escorcia-Gutierrez J, Gamarra M, Soto-Diaz R, Pérez M, Madera N, Mansour RF (2022) Intelligent agricultural modelling of soil nutrients and ph classification using ensemble deep learning techniques. J Agric 12(7):1–16 (MDPI)
14. Kusanur V, Veena S, Chakravarthi (2021) Using transfer learning for nutrient deficiency prediction and classification in tomato plant. J Adv Comput Sci Appl 12 (Elsevier)
15. Amani MA, Marinello F (2022) A deep learning-based model to reduce costs and increase productivity in the case of small datasets: a case study in cotton cultivation. J Agric 12(2):1-16 (MDPI)
16. Malvade NN, Yakkundimath R, Saunshi G, Elemmi MC, Baraki P (2022) A comparative analysis of paddy crop biotic stress classification using pre-trained deep neural networks. J Artif Intell Agric 2:167–175 (Elsevier)
17. Picon A, Bereciartua-Perez A, Eguskiza I, Romero-Rodriguez J (2022) Deep convolutional neural network for damaged vegetation segmentation from RGB images based on virtual NIR-channel estimation. J Artif Intell Agric 6:199–210 (Elsevier)
18. Subeesh A, Bhole S, Singh K, Chandel NS, Rajwade YA, Rao KVR, Kumar SP, Jat D (2022) Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. J Artifi Intell Agric 6:47–54 (Elsevier)
19. Yin H, Gu YH, Park C-J, Park J-H, Yoo SJ (2020) Transfer learning-based search model for hot pepper diseases and pests. J Agric 10(10):1–16 (MDPI)
20. Ananda S, Paymode Vandana B, Malde (2022) Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG. J Artif Intell Agric 6:23–33 (Elsevier)

A Machine Learning Method for Stock Market Price Prediction

Dhruv R. Makadia, Jinal Kotadia, Abhay Monasara, and Dweepna Garg

Abstract Multiple Linear Regression (MLR) is a statistical modeling technique that has been commonly used in predicting the behavior of a target variable, such as stock prices, based on the values of several predictor variables, such as open, low, and high prices. The MLR model attempts to fit a linear equation to the data, where the predictor variables affect the target variable. In our paper, to predict the stock market price, the MLR model is trained and tested using historical data of a company's stock prices. The open, low, and high prices are used as predictor variables, while the close price is used as the target variable. The model is trained to predict future movements in the stock prices based on the values of the predictor variables. Additionally, the accuracy of MLR predictions is improved by using a larger set of predictor variables, such as economic indicators, company financials, and news sentiment, to provide a more comprehensive picture of the market conditions.

Keywords Linear regression · Multiple linear regression · Machine learning · Stock market prediction

1 Introduction

The stock market has been a complex and dynamic field where stock price predictions become a critical task in the finance industry due to its dynamism, fluctuations, and strangeness, and has always been a topic of interest for investors, traders, and analysts. It is challenging to forecast how stock prices will move because so many factors,

D. R. Makadia · J. Kotadia · A. Monasara
Department of Computer Science and Engineering, Devang Patel Institute of Advance Technology and Research (DEPSTAR), Faculty of Technology and Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

D. Garg (✉)
Department of Computer Engineering, Devang Patel Institute of Advance Technology and Research (DEPSTAR), Faculty of Technology and Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India
e-mail: dweeps1989@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_20

including global events, economic and political situations, and company-specific considerations, influence stock prices [1]. Predicting stock prices with the highest accuracy possible can help to make informed decisions and maximize profits based on current and future prices. Stock price predictions have historically been made using conventional techniques including technical and fundamental analysis [1]. Machine learning has the potential to improve stock market prediction, according to recent research. Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and other neural network-based methods have demonstrated promising results and offer distinct weights for each case. In this paper, we evaluated Multiple Linear Regression algorithm for stock price prediction. A target variable can be predicted using correlation of several independent and dependent variables. Multiple linear regression can be used to determine the link between numerous market parameters and stock prices in stock market forecasting. Future stock prices can be predicted using a model that can be trained on historical data. The paper has been coordinated in an accompanying manner: Sect. 1 presents the part by depicting the goal of the paper. The related work in Sect. 2 is followed by Sect. 3, which shows the method for achieving the goal described in Sect. 1. Experimental analysis is shown in Sect. 4, where the model's visualization, prediction, and accuracy are discussed. The section is finished in Sect. 5.

2 Literature Survey

Deep Learning showed promising outcomes in predicting the stock market. Deep Learning models based on long short-term memory (LSTM) have seemed to achieve an accuracy of 90% in predicting future prices [2]. Techniques for forecasting stock prices today use sophisticated, intelligent approaches based on either fundamental or technical analysis. Specifically, the information size is gigantic and non-direct for financial exchange examination. To deal with this variety of information, you really want a proficient model that can track down secret examples and unpredictable connections in this tremendous assortment of information. In this field, machine learning techniques have improved efficiency by 60–86% when compared to previous approaches [3]. Traditional calculations like direct relapse [4], irregular walk hypothesis (RWT) [5], moving normal union and dissimilarity (MACD) [6], and a few straight models like autoregressive moving normal (ARMA) and autoregressive integrated moving normal (ARIMA) [6] have been used in a lot of previous research in this area to predict stock prices. Current studies in machine learning can improve stock market prediction, and techniques like Support Vector Machine (SVM) and Random Forest (RF) [7] can be used.

In some studies, the use of Random Forest (RF) for forecasting has been suggested. RF is a technique for an ensemble. It frequently performs regression and classification tasks simultaneously. During the training phase, multiple decision trees are constructed, resulting in the mean regression of each tree [8].

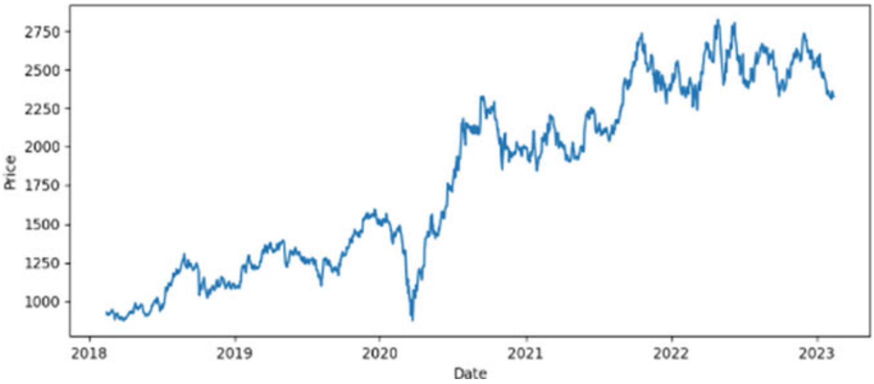


Fig. 1 Close price—RIL

3 Methodology

3.1 Data Statistics

Historical dataset for the companies listed in the Nifty 50 for the fiscal year 2023 was obtained from Yahoo Finance [9]. The data obtained from the website range from the date of the company's listing on the National Stock Exchange to the present date (excluding the market off days). The data contains day's high, low, and open price, adjusted volume, and close price. As our target attribute, we used the closing price of any stock. Figure 1 depicts the closing price of RIL stock plotted using matplotlib.

3.2 Data Splitting and Feature Identification

This research uses 20% of the dataset as its testing data and the remaining dataset has been utilized to train linear regression model. Here, Reliance Industries Limited data has been used for visualizing the co-relations. Figure 2 shows the linear correlation of the opening and closing prices of each day for RIL.

Figure 3 shows the linearity in the correlation plotting of the day high and day close price for each day of the dataset. The linearity can be seen to be concise with less outliers.

Figure 4 shows the correlation plotting of the day low and day close price for each day of the dataset. The plotting shows some outliers in the plotting.

By finding the correlation between the target attribute and the features, we conclude that all the co-relations possess a linear dependency.

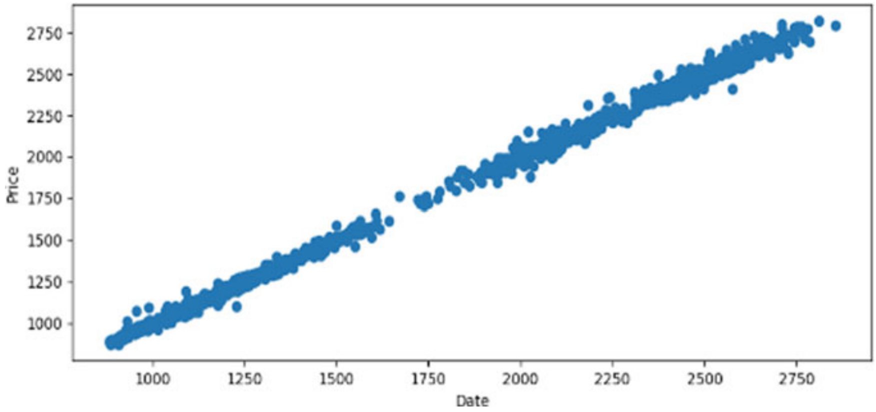


Fig. 2 Day open versus day close price correlation

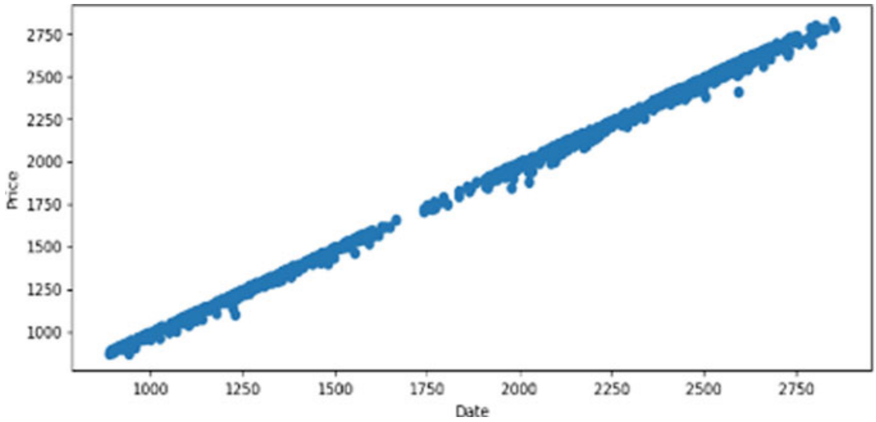


Fig. 3 Day high versus day close price correlation

3.3 Multiple Linear Regression Approach

When modeling the relationship that exists between a dependent variable and two or more independent variables, the statistical method of multiple linear regression is utilized. An apparent expansion of direct relapse demonstrates the connection between a solitary-free factor and a solitary ward variable. By limiting the amount of the squared contrasts between the anticipated and genuine upsides of the reliant variable, various direct relapse tries to assess the relapse coefficient esteems that best match the information. Most often, the least squares regression method is used to achieve this. After estimating the regression coefficients, the model can be used to predict the value of the dependent variable using a collection of independent variable values. Day Open, Day High, and Day Low are the independent variables in our case,

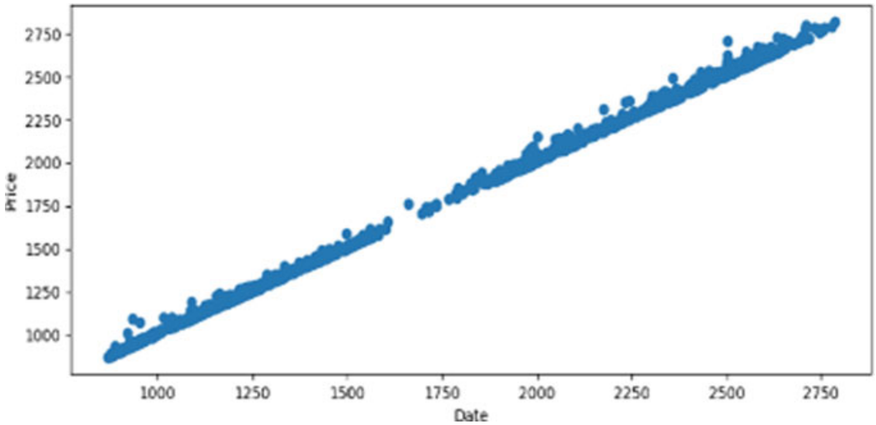


Fig. 4 Day low versus day close price correlation

while Day Close is the dependent variable. The dataset was then split into training and testing datasets. A while later, the preparation dataset was squeezed into the direct relapse model. The prediction's outcome was recorded as an array into a variable following the training phase. We discovered a novel neural network-free strategy for stock price forecasting from this.

3.4 Evaluation Measures

In order to evaluate the model's accuracy, evaluation metrics such as the R2 score, mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) were also calculated.

MAE can be calculated using the below Eq. (1).

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|A_t - F_t|}{|A_{t|}} \tag{1}$$

One can find MSE using the Eq. (2).

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2 \tag{2}$$

Here A_t stands for actual value and F_t is the predicted price [10].

RMSE can be computed using the Eq. (3).

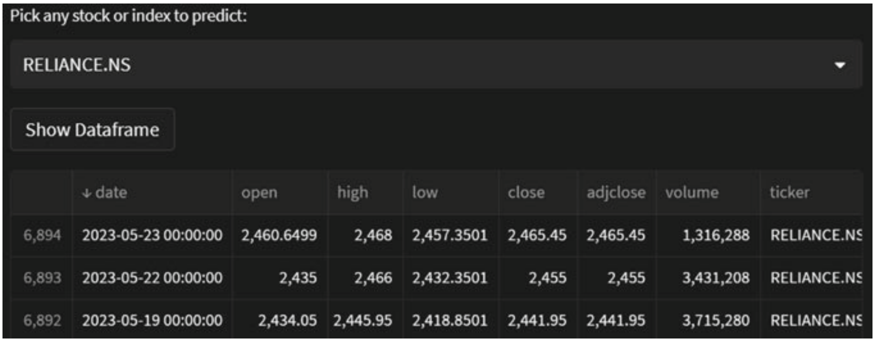


Fig. 5 RIL data frame

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(O_i - F_i)^2}{n}} \tag{3}$$

Here O_i stands for original value and F_i is the forecasted close price. We can use Eq. (4) to determine the R^2 score, where RSS stands for the summation of squares of residue and TSS stands for the total summation of squares [1].

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} \tag{4}$$

4 Experimental Analysis

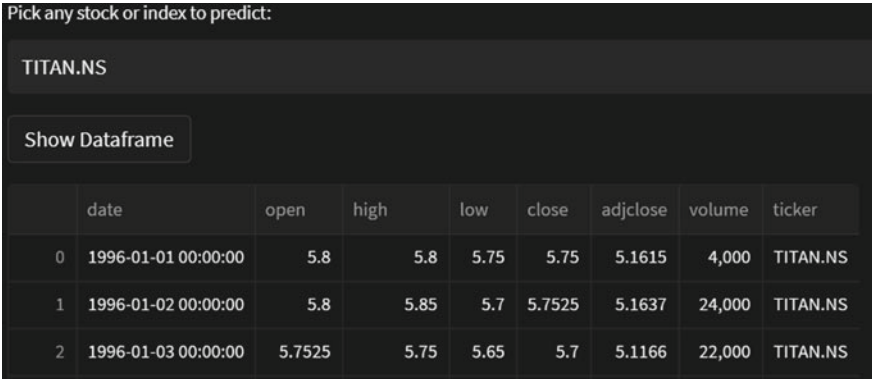
4.1 Dataset

The dataset from Yahoo Finance contains the data of each day of stock price action with features such as open, high, low, and close. The model, which utilized a split of 80:20 for the training and testing data, was able to accommodate both the independent features and the dependent features. Figure 5 shows the data frame visualization of the stock price of RIL on the GUI.

Figure 6 shows the data frame visualization of the stock price of Titan on the GUI.

4.2 Trend Analysis

For a better visualization of our data frame, we plotted the opening, high, and closing prices of each stock listed on a graph. By looking at each graph, we can clearly


Pick any stock or index to predict:

TITAN.NS

Show Dataframe

	date	open	high	low	close	adjclose	volume	ticker
0	1996-01-01 00:00:00	5.8	5.8	5.75	5.75	5.1615	4,000	TITAN.NS
1	1996-01-02 00:00:00	5.8	5.85	5.7	5.7525	5.1637	24,000	TITAN.NS
2	1996-01-03 00:00:00	5.7525	5.75	5.65	5.7	5.1166	22,000	TITAN.NS

Fig. 6 TITAN data frame

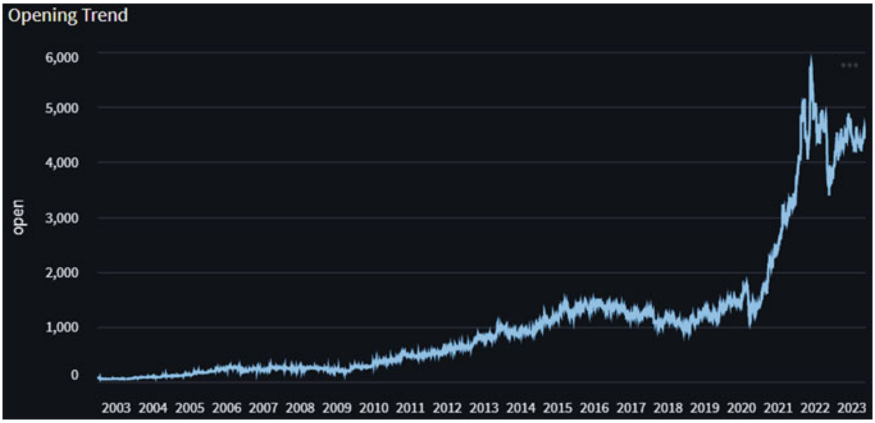


Fig. 7 RIL day open trend

identify that there is a correlation between each of the features mentioned in the paper. Figure 7 visualizes the Day Open trend of RIL on a graph on the GUI.

Figure 8 visualizes the Day Close trend of RIL on a graph on the GUI. Figure 9 visualizes the Day High trend of RIL on a graph on the GUI.

4.3 Prediction

Here in this module of the GUI, the model after being trained with the training data predicted the closing price of the stocks. The predicted price can be used as a reference while trading in a real-life scenario.

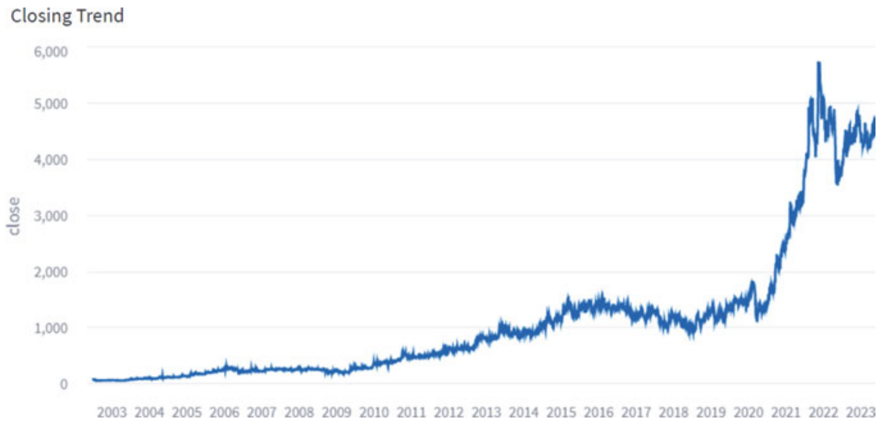


Fig. 8 RIL day close trend

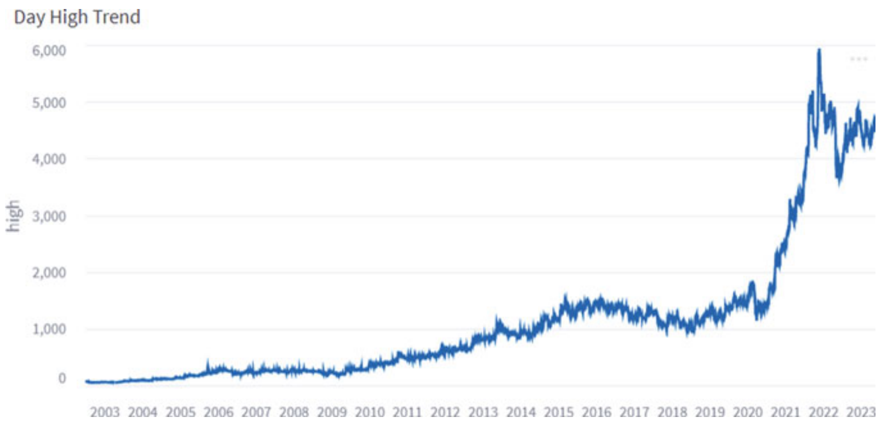


Fig. 9 Day high trend

Figure 10 shows the predicted closing price of RIL. Figure 11 shows the predicted closing price of TITAN.

4.4 Accuracy (Evaluation Metrics)

We used four parameters, the R^2 Score, the MAE, the MSE, and the RMSE, to verify the model's results and the predicted close price. We can find the above evaluation metrics data for each stock and can validate the predicted close price. Figure 12 displays the evaluation metrics data of RIL.

Figure 13 depicts the assessment measurement information of TITAN.

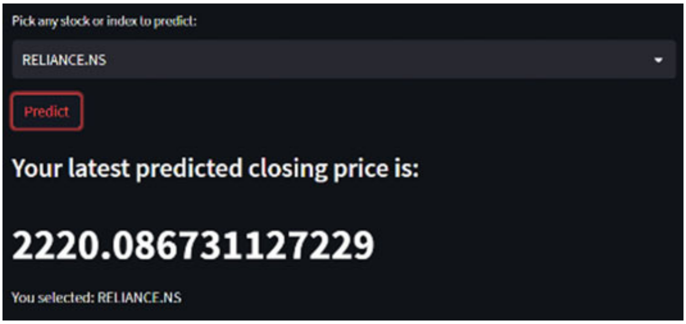


Fig. 10 RIL predicted closing

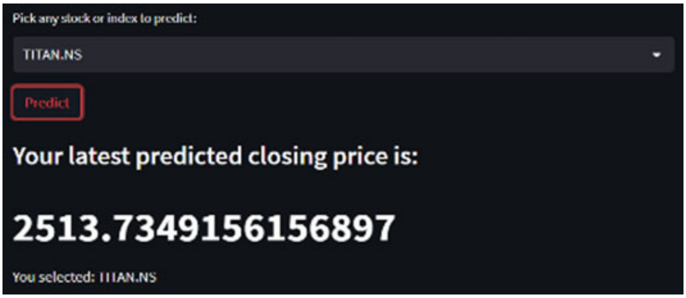


Fig. 11 TITAN predicted closing

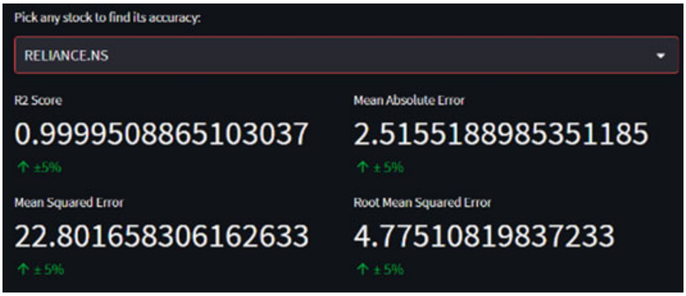


Fig. 12 RIL evaluation metrics

A plotting of the original close price and the predicted close price were plotted in the manner depicted in Fig. 14 to evaluate and visualize the prediction's accuracy.

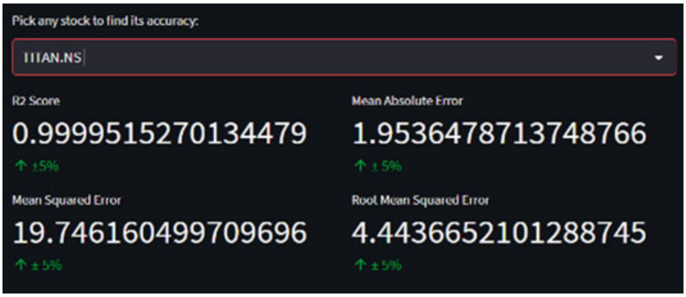


Fig. 13 TITAN evaluation metrics

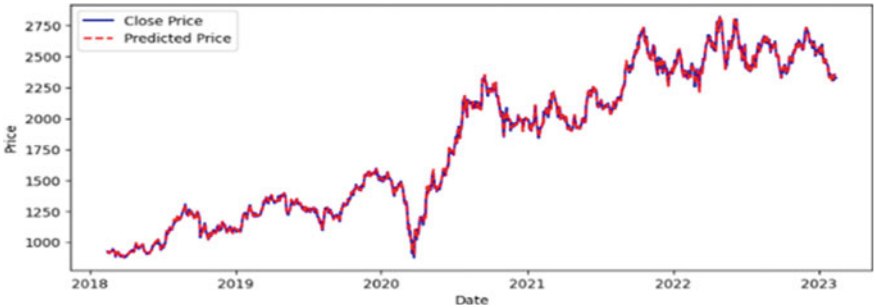


Fig. 14 Plot for day close and predicted close

5 Conclusion

Using the most fundamental machine learning technology, this study set out to accurately predict the Day Close price of any stock included in the Nifty 50 for the year 2023. The three dataset features of Day Open, Day High, and Day Low served as the foundation for the prediction of the close price. The research relied on information gathered from the Yahoo Finance API, which included information ranging from the NSE institution listing to the most recent available data. To evaluate the accuracy of the predicted price, it is done using four evaluation metrics namely RMSE, MAE, R^2 score, and MSE. The model scored an R^2 score of 0.99 and had MSE and RMSE of 143.02 and 11.95 respectively. For scoring the error percentage, we used MAPE and the model had a MAPE of 0.004. According to [11], we can predict the value using SVM, ANNs, RF, and Naive-Bayes in a multi-stage approach. With an accuracy of 73.3 percent, the Naive-Bayes model, a Gaussian process, performed the worst. The accuracy in ANN was 86.69%, in SVM it was 89.33%, in RF it was 89.98%, and in Naive-Bayes it was 90.19%.

References

1. Vijh M et al (2020) Stock closing price prediction using machine learning techniques. Procedia Comput Sci 167:599–606
2. Ni L et al (2019) Forecasting of forex time series data based on deep learning. Procedia Comput Sci 147:647–652
3. Li L et al (2017) Research on machine learning algorithms and feature extraction for time series. In: 2017 IEEE 28th annual international symposium on personal, indoor, and mobile radio communications (PIMRC). IEEE
4. Seber GAF, Lee AJ (2003) Linear regression analysis, vol 330. Wiley
5. Reichek N, Devereux RB (1982) Reliable estimation of peak left ventricular systolic pressure by M-mode echographic-determined end-diastolic relative wall thickness: identification of severe valvular aortic stenosis in adult patients. Am Heart J 103(2):202–209
6. Chong TT-L, Ng WK (2008) Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30. Appl Econ Lett 15(14):1111–1114
7. Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2(3):18–22
8. Kumar M, Thenmozhi M (2006) Forecasting stock index movement: A comparison of support vector machines and random forest. In: Indian institute of capital markets 9th capital markets conference paper
9. Yahoo Finance—Stock market live, quotes, business and finance news (2023) https://finance.yahoo.com/
10. Patel J et al (2015) Predicting stock market index using fusion of machine learning techniques. Expert Syst Appl 42(4):2162–2172
11. Patel J et al (2015) Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Syst Appl 42(1)

Eye for Blind: A Deep Learning-Based Sensory Navigation System for the Blind

Mrunal Shidore, Sahil Kakurle, Manish Shetkar, Malhar Kapshe, and Abhijeet Kamble

Abstract The research proposes a new system to assist visually impaired individuals in navigating their surroundings by introducing an "Eye for Blind" system that employs Convolutional Neural Networks (CNN). Our approach utilizes a CNN model to recognize images and generates audio feedback with descriptive labels. This innovative system has the potential to help visually impaired individuals detect and recognize photos, enhancing their independence and quality of life. We evaluated our proposed system using publicly available datasets, and the outcomes suggest that our approach has a high level of accuracy in recognizing social media posts, demonstrating its potential as a solution for the visually impaired community. By using the Eye for Blind system, visually impaired individuals will be able to navigate their surroundings more effectively and efficiently, eliminating their dependence on assistance from others. This technology has enormous potential to improve the quality of life for visually impaired individuals and can contribute to the development of other similar systems that can enhance their independence and mobility. In conclusion, our proposed system has significant potential to make a positive impact on the lives of visually impaired individuals, and we hope that it can be further developed and implemented in the future.

Keywords Neural network · CNN · Object · LSTM

M. Shidore (✉) · S. Kakurle · M. Shetkar · M. Kapshe · A. Kamble
Vishwakarma Institute of Technology, Bibwewadi, Pune, India
e-mail: mrunal.shidore@vit.edu

S. Kakurle
e-mail: sahil.kakurle20@vit.edu

M. Shetkar
e-mail: manish.shetkar20@vit.edu

M. Kapshe
e-mail: malhar.kapshe20@vit.edu

A. Kamble
e-mail: abheejit.kambale20@vit.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_21

1 Introduction

Most people take their ability to see for granted, but for people with visual impairments, it is a daily challenge. Deep learning and computer vision advancements have resulted in the development of several technologies that can help visually impaired people, such as object detection tools and text-to-speech systems. CNNs have shown great promise, particularly in image recognition and object detection. This research paper delves into the use of CNNs to develop a system that can assist the visually impaired. Our "Eye for Blind" system uses a CNN to detect objects and provides audio feedback to users, allowing them to navigate and interact with their surroundings. We present our system's architecture and evaluate its performance using the MS COCO dataset. The results we obtain highlight the value of CNNs in the creation of assistive technology for the visually impaired.

2 Literature Review

Deep learning has shown considerable potential in computer vision in years, particularly in the area of image captioning. Several approaches for generating natural language descriptions of images using deep learning models such as convolutional neural networks (CNN) and recurrent neural networks (RNN) have been proposed. (RNN). In this literature review, we will discuss the key findings and contributions of seven papers on deep learning image captioning [1]. The authors describe the various layers of a CNN and how they extract features from images, resulting in improved performance in tasks of classification and detection [2]. The authors test various RNN architectures and training methods, and their model performs well on the Flickr30k dataset [3]. The authors propose a similar image captioning model using a CNN and an RNN, but with a focus on reducing overfitting and enhancing the generated caption's quality. They experiment with different regularization techniques and achieve good results on the MS COCO dataset [4]. Authors highlight importance of attention mechanisms and using a large and diverse training dataset for improved performance. The authors [5] provide a comprehensive overview of different approaches to image captioning using deep learning, including those that use reinforcement learning and those that generate captions in multiple languages [6]. The authors provide insights into the strengths and weaknesses of each model and suggest avenues for future research [7]. The paper presents a model that uses both convolutional neural networks (CNNs) and long short-term memory (LSTM) networks to encode images, also to generate captions. Authors propose that their model surpasses other methods on several benchmark datasets [8]. The authors propose a multimodal memory network (MMN) that can adapt to different user preferences based on their interactions with the system. The MMN uses both visual and textual information to generate captions and can be fine-tuned for individual users based on their feedback [9]. The authors propose a model that generates captions using a policy network and predicts the quality of the generated captions using a network of values. On several datasets that serve as benchmarks, the authors show how their model outperforms conventional approaches [10]. The authors propose a model that produces the image with a CNN and produces the caption with an LSTM. A maximum likelihood estimation objective is used to train the model [11]. On several benchmark datasets, the authors demonstrate that the model they develop is better than other cutting-edge techniques [12]. The authors propose a training method that incorporates both supervised and reinforcement learning. The algorithm is learned to maximize the likelihood of ground-truth captions and is recognized based on the quality of the captions created.

In comparison to other published materials, this proposed system adds a unique and innovative application of deep learning techniques to aid visually impaired individuals, which can potentially revolutionize how visually impaired individuals navigate their surroundings. The system has been evaluated using publicly available datasets, and results show that it has a high level of accuracy in recognizing social media posts, making it a promising solution for the visually impaired community. This proposed system can contribute to the development of other similar systems that can enhance the independence and mobility of visually impaired individuals. Therefore, this proposed system adds significant value to the subject area of computer vision and can have a positive impact on the lives of visually impaired individuals.

3 Work Done

We developed a model that generates caption of images by combining two neural network architectures: a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network. The CNN is adept at processing data with a 2D matrix-like structure, making it ideal for extracting features from images. It is commonly used for image classification tasks such as identifying objects. The CNN scans the image, by combining important features to make classifications even for images that are translated, rotated, scaled, or viewed from different perspectives. In contrast, the LSTM is a type of recurrent neural network (RNN) that is suitable for arranging estimated tasks. It can predict the next word based on earlier text and is more effective than traditional RNNs because it overcomes the short-term memory limitations. It can retain related data through the input and selectively discard non-related data using a “forget gate”. By combining the strengths of CNN and LSTM, our model generates accurate natural language descriptions of images by describing objects, attributes, and relationships as shown in Fig. 1.

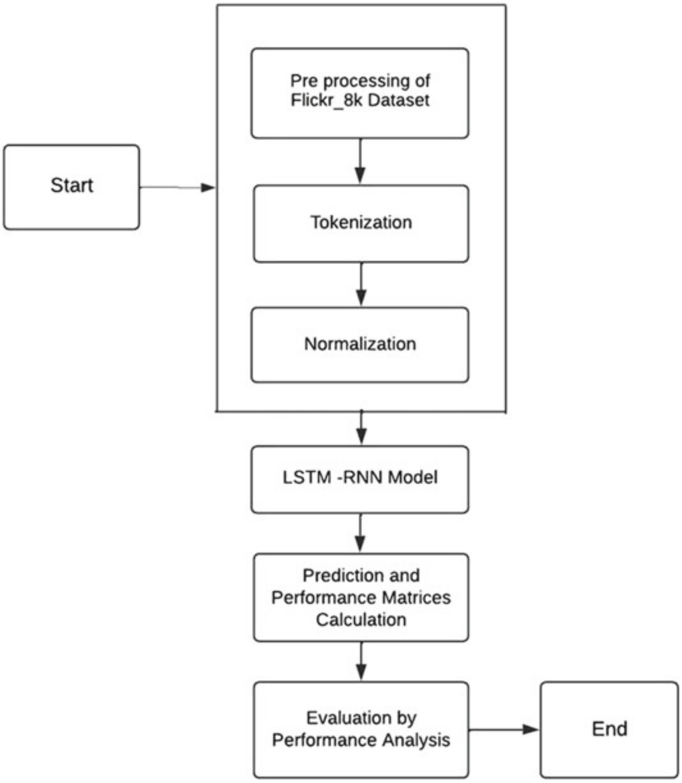


Fig. 1 Long short-term memory (LSTM) diagram

4 Dataset Collection and Preprocessing

In our image caption generator project, we have opted to use the Flickr_8K dataset as it is smaller and can be trained in a reasonable amount of time. Although larger datasets such as Flickr_30K and MSCOCO are available, training on them can take weeks. The Flickr_8K dataset comprises 8092 images, each with five corresponding descriptions, depicting various scenes and situations. We have categorized data into 3 sets—6000 for training, 1000 for validation, and 1000 for testing, and images have varying dimensions. Although using a larger dataset may result in better models, we chose the Flickr_8K dataset as it is easier to manage (Fig. 2).

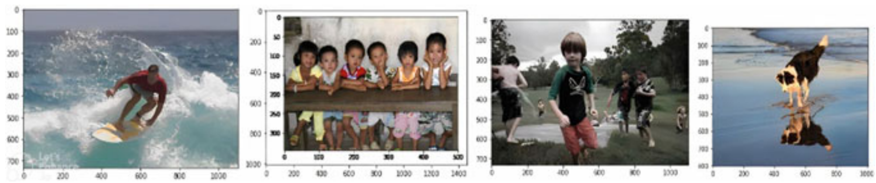


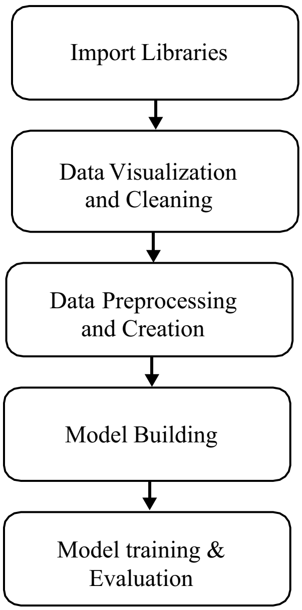
Fig. 2 Sample Images from dataset

5 Methodology

A. Algorithm

1. Import all the required libraries.
2. Mount drive and import dataset.
3. Visualization of images, create the data frame and list which contains all the captions.
4. Data cleaning.
5. Data preprocessing.
6. Dataset creation.
7. Defining the CNN-RNN model.
8. Training the model.
9. Testing the model (Fig. 3).

Fig. 3 System architecture flow chart



B. **Libraries**

1. TensorFlow, developed by Google, is an open-source platform that offers comprehensive machine learning support throughout the entire process. It is an integrated framework that can handle popular deep learning networks like CNN and RNN with diverse configurations. TensorFlow is built to be extremely adaptable, portable, and efficient on hardware, which makes it an excellent option for anyone interested in machine learning.
2. Google Text-to-Speech (gTTS) is a python library and command-line utility that allows you to connect with the Google Translate text-to-speech API. With the gtts module, speech translation becomes effortless, making it a valuable tool for anyone in need of speech-to-text services.
3. The playsound module is an adaptable, cross-platform module that enables the playback of audio files. It can be easily installed using pip within a virtual environment, and requires no dependencies, thus making it uncomplicated to use. The display module is equipped with a public API for display tools in IPython, rendering it a valuable asset for anyone who needs to visualize data in their work.
4. Keras is a high-level neural networks API developed using python that can operate on TensorFlow, CNTK. It prioritizes fast experimentation and speedy results, which are essential components of good research. It supports both convolutional and recurrent networks, as well as a hybrid of the two, making it a critical tool for those interested in diving deeper into machine learning and data visualization.

C. **Data Understanding and Visualization**

A data frame is created to summarize the image ID, file path, and associated captions. Each image ID in the dataset has five captions associated with it as shown in Fig. 4. As a result, number of samples would be 40,455 (Fig. 4).

A word cloud is a visual representation of textual data in which words are displayed in varying font sizes, colors, and orientations based on their frequency or importance as shown in Figs. 5 and 6. It is a popular data visualization technique used to quickly and easily identify the most significant or commonly occurring words. Typically, larger and bolder word appears in the word cloud, the more frequently it appears in the text. Word clouds can be useful in various research contexts, such as identifying key themes or topics in a corpus of text, summarizing survey responses, or analyzing social media conversations (Figs. 5 and 6).

D. **Data Cleaning**

In data analysis, preprocessing text data is crucial to ensure accuracy and usefulness. Techniques such as removing punctuation, converting text to lowercase, and eliminating numeric values are commonly used to clean text data. Removing punctuation helps reduce the size of the text and avoid extraneous characters that could interfere with the analysis. Converting text to lowercase standardizes the text and reduces the

	ID	Path	Captions
0	1000268201_693b08cb0e.jpg	/kaggle/input/flickr8k/Images/1000268201_693b0...	A child in a pink dress is climbing up a set o...
1	1000268201_693b08cb0e.jpg	/kaggle/input/flickr8k/Images/1000268201_693b0...	A girl going into a wooden building
2	1000268201_693b08cb0e.jpg	/kaggle/input/flickr8k/Images/1000268201_693b0...	A little girl climbing into a wooden playhouse
3	1000268201_693b08cb0e.jpg	/kaggle/input/flickr8k/Images/1000268201_693b0...	A little girl climbing the stairs to her playh...
4	1000268201_693b08cb0e.jpg	/kaggle/input/flickr8k/Images/1000268201_693b0...	A little girl in a pink dress going into a woo...
...	...	...	...
40450	997722733_0cb5439472.jpg	/kaggle/input/flickr8k/Images/997722733_0cb543...	A man in a pink shirt climbs a rock face
40451	997722733_0cb5439472.jpg	/kaggle/input/flickr8k/Images/997722733_0cb543...	A man is rock climbing high in the air
40452	997722733_0cb5439472.jpg	/kaggle/input/flickr8k/Images/997722733_0cb543...	A person in a red shirt climbing up a rock fac...
40453	997722733_0cb5439472.jpg	/kaggle/input/flickr8k/Images/997722733_0cb543...	A rock climber in a red shirt
40454	997722733_0cb5439472.jpg	/kaggle/input/flickr8k/Images/997722733_0cb543...	A rock climber practices on a rock climbing wall

Fig. 4 List which contains all the captions and path

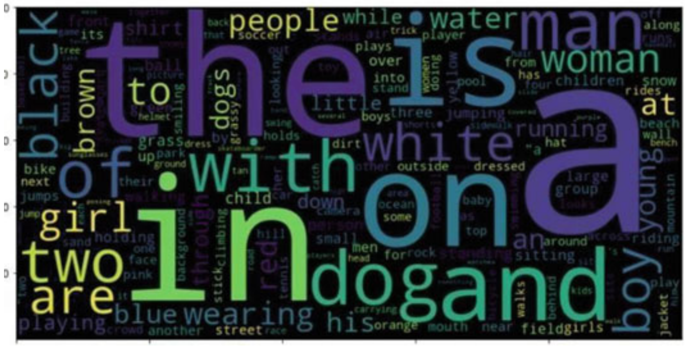


Fig. 5 Frequently occurring words in the captions

A child in a pink dress is climbing up a set of stairs in an entry way

A girl going into a wooden building

A little girl climbing into a wooden playhouse

A little girl climbing the stairs to her playhouse

A little girl in a pink dress going into a wooden cabin

Fig. 6 Visualization of images and captions together

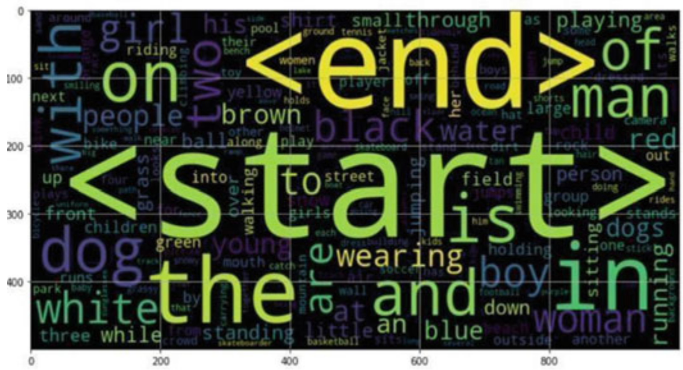


Fig. 7 Top 30 common words

number of unique words in the corpus. Eliminating numeric values removes irrelevant numerical data from the analysis. These preprocessing techniques ensure that text data is accurate, standardized, and ready for analysis. Preprocessing is a critical step in cleaning and preparing text data for analysis.

E. Data Preprocessing

In Natural Language Processing (NLP), preprocessing is crucial before training any model. Caption preprocessing involves tokenizing captions, limiting the vocabulary to the top 5000 words, replacing unknown words with "UNK" token, creating word-to-index and index-to-word mappings, and padding sequences with zeros as shown in Fig. 7. Tokenizing captions creates a unique vocabulary of words. Replacing unknown words with the "UNK" token ensures consistency in the vocabulary. Mapping words to indexes and vice versa is necessary for models to work with numbers instead of text. Padding sequences with zeros ensures all captions are the same length. These preprocessing steps help to create a consistent and uniform dataset that can be used to train various NLP models effectively (Fig. 7).

Preprocessing of images is critical for preparing data for machine learning models. Resizing images into a fixed shape of (299, 299) is necessary for InceptionV3, as it ensures uniformity in image sizes. Normalization of pixel values within -1 to 1 is essential for numerical stability and faster model convergence during training. It centers the data around 0, making it feasible for model to extract features from the images. Resizing and normalization are crucial pre-processing steps that ensure images are in the correct format for InceptionV3, thereby facilitating effective learning from the data.

F. Dataset Creation

A deep learning model for picture captioning has been developed. For building dataset, a function is created which maps the image path to their feature. This function is then applied using a builder function to create the train and test dataset. The dataset is shuffled and batched during the building process. The train-test split is created using an 80–20 ratio and a random state of 42 for reproducibility. After constructing the dataset, the dimensions of each image should be (batch_size, 8*8, 2048). This means that the image feature vector is flattened to a 64-dimensional vector and has a feature size of 2048. The shape of each caption in the dataset after building should be (batch_size, max_len), where max_len is the maximum length of a caption. The captions are encoded as integer sequences, where each integer corresponds to a word in the vocabulary. The integer sequences are padded to the same length using zeros, so that all captions have the same length. Overall, this dataset building process ensures that the image features and captions are properly aligned and formatted for training the image captioning model.

G. **Data Evaluation**

This study focuses on developing a deep learning model that generates image captions by combining convolutional and recurrent neural networks. The Adam optimizer is applied with a learning rate of 0.001 and a Sparse Categorical Cross—entropy loss function, with a mask for ignoring padded zeros in caption loss calculation.

A custom loss function is also defined for calculating the mean loss over the batch. The training dataset is iterated through for a specific number of epochs to calculate the average training loss per epoch, while the model is being examined on the test data to determine its performance and test loss. The loss_plot and test_loss_plot lists store the training and test loss for each epoch, and the best_test_loss variable is initially set to a high value and updated when a lower test loss is achieved. Additionally, the checkpoint manager saves the model weights, allowing for the best model to be restored during testing or training interruptions. The experimental results demonstrate that the suggested strategy is effective at generating picture captions, as evidenced by the reduced test loss over the epochs. This approach can be extended to other natural language processing tasks that require multimodal learning, such as image retrieval or visual question answering (Fig. 8).

As shown in Fig 8, the green line on the graph represents the test loss, which exhibits a consistent trend parallel to the training loss (depicted in orange), even when using teacher forcing. Our objective is to explore how to combine attention mechanisms with E-D architecture for images, hence we will limit ourselves to 15 approaches and avoid further expansion.

6 Output of the Implemented Model

As shown in Fig. 9, below is the real-time result of the implemented model. Model classifies the emotion based on facial expression and maps the emotion to the emoji or avatar (Fig. 9).

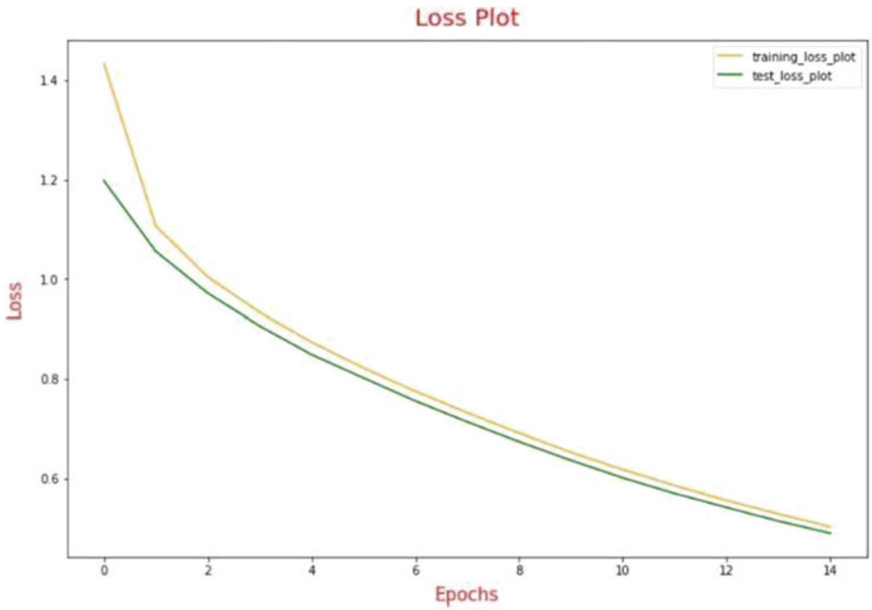


Fig. 8 Loss plot of training vs test

Fig. 9 Real-time result of image caption generation

7 Future Scope

As the number of images on social media and the internet continues to increase, image captioning has become a crucial challenge. To address this challenge, our goal is to explore methods that can generate multiple sentences with different content. However, achieving complete accuracy may not be possible because these techniques lack contextual information. Therefore, leveraging the context of images, such as image captioning, could help improve image retrieval. In future versions of this project, we plan to enhance the identification of classes with lower precision by training the model with additional image captioning datasets like COCO Captions. Additionally, we believe that combining this approach with existing image retrieval methods like histograms and shapes could lead to better image retrieval results. Due to the rapid increase of photos on internet, image captioning has become an urgent need, and we aim to improve its effectiveness by incorporating contextual information and utilizing more datasets.

8 Conclusion

We have been developed a neural network modal which combines two architectures a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network—to produce natural language descriptions for images. It is based on multi-label classification utilizing Quick Text and CNN and is effective for recognizing and extracting objects from images, as well as generating captions depending on the datasets provided. Both CNN and LSTM models are used in order to get accurate results and to find the relation between objects in images. The LSTM belongs to the category of recurrent neural networks (RNNs) and is particularly useful for tasks involving prediction of sequences. Our approach provides accurate natural language descriptions of images by characterizing objects, properties, and relationships by combining the features of CNN and LSTM.

References

1. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: 2017 international conference on engineering and technology (ICET), August, IEEE, pp 1–6
2. Sharma G, Kalena P, Malde N, Nair A, Parkar S (2019) Visual image caption generator using deep learning. In: 2nd international conference on advances in science and technology (ICAST)
3. Wang H, Zhang Y, Yu X (2020) An overview of image caption generation methods. In: Computational intelligence and neuroscience
4. Vinyals O, Toshev A, Bengio S, Erhan D (2016) Show and tell: Lessons learned from the 2015 mscoco image captioning challenge. IEEE Trans Pattern Anal Mach Intell 39(4):652–663

5. Raypurkar M, Supe A, Bhumkar P, Borse P, Sayyad S (2021) Deep learning-based image caption generator. Int Res J Eng Technol (IRJET) 8(03)
6. Hossain M, Sohel F, Shiratuddin MF, Laga HAMID (2018) A comprehensive study of deep learning for image captioning. arXiv preprint arXiv:1810.04020
7. Rao S, Santhosh S, Salian KP, Chidananda T, Kumar SS (2022) A novel approach to generate the captions for images with deep learning using CNN and LSTM model. In: 2022 International conference on distributed computing, VLSI, electrical circuits and robotics (DISCOVER), October, IEEE, pp 176–179
8. Vinyals O, Toshev A, Bengio S, Erhan D (2015) Show and tell: a neural image caption generator. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3156–3164
9. Abdel-Hamid O, Deng L, Yu D (2013) Exploring convolutional neural network structures and optimization techniques for speech recognition. In: Interspeech, August, vol 2013. pp 1173–5
10. Dumoulin V, Visin F (2016) A guide to convolution arithmetic for deep learning. arXiv preprint arXiv:1603.07285
11. Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS (2016) Deep learning for visual understanding: a review. Neurocomputing 187:27–48
12. Mathur P, Gill A, Yadav A, Mishra A, Bansode NK (2017) Camera2Caption: a real-time image caption generator. In: 2017 international conference on computational intelligence in data science (ICCIDS), June, IEEE, pp 1–6
13. Herdade S, Kappeler A, Boakye K, Soares J (2019) Image captioning: Transforming objects into words. In: Advances in neural information processing systems, vol 3

Adaptive Robust Beamforming Approach for Power Optimization in Device-To-Device (D2D) Communication Network

Pimmy Gandotra, Rakesh Kumar Jha, and Brejesh Lall

Abstract Device-to-device (D2D) communication in 5G wireless communication networks (WCNs) is gaining popularity, providing higher network capacity and competently meeting the imminent user demands. The achievable capacities can be further scaled up with the use of beamforming in D2D networks. This paper considers a university campus, modelled as a D2D communication network, with base station (BS) equipped with an array of N_B antennas, as in a massive MIMO network. The university campus under consideration is divided into three zones, Z_1, Z_2 and Z_3, with varying pair densities, such that Z_2 always has the maximum pair density, during the university working hours. Beamforming is performed adaptively in each zone, where the beam length adaptively varies as per the pair count, so as to focus the beam towards the highest pair density. If the pair density rises enormously in a particular zone, beam merging is proposed in that case, with each of the merged beam, choosing optimal transmission power using non-cooperative game theory, for a power optimized network. Merging can enhance the beam gain sizably, assuring quality service to all the demanding pairs. The efficiency of the proposed scheme is depicted via simulation results.

Keywords 5G · D2D · Massive MIMO · Beamforming · Beam merging

P. Gandotra
Indian Institute of Technology, Delhi, New Delhi, India

P. Gandotra (✉)
School of Electronics and Communication Engineering, SMVDU, Katra, J&K, India
e-mail: pimmy.gandotra@gmail.com

Department of Electronics and Communication Engineering, Jaypee Institute of Information Technology, Noida, India

R. K. Jha
Indian Institute of Information Technology, Design and Mnaufacturing, Jabalpur, Madhya Pradesh, India

B. Lall
Indian Institute of Technology, Delhi, New Delhi, India
e-mail: brejesh@ee.iitd.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_22

1 Introduction

The technological progressions in the wireless communication networks (WCNs) with the rise in subscriber count has been accompanied with an explosion in the demanded applications. The intensifying demands of localizing applications, will not be adequately met by the present day cellular network architectures in the near future. This calls for architectural advancements in cellular networks, presenting a number of technologies in the 5G WCNs [1]. Of all the novel technologies introduced to the 5G WCNs, device-to-device (D2D) communication is indispensable for low-power proximity communication [2, 3]. It is also a competent technology, in terms of securing the wireless networks [4, 5].

D2D communication in 5G WCNs is believed to achieve throughput values in gigabits per second (Gbps). This is possible by utilizing the millimetre wave frequencies, employing beamforming for supporting high data rates with densely deployed cells [6]. Inclusion of beamforming, along with the use of multiple antennas at the transmitter and receiver chains robust, potent communication links under high path loss scenarios. Highly directional beam from the multiple antennas improve the capacity per link [8], enticing beamforming in 5G WCNs. Considerable improvement in the signal-to-noise ratio (SNR) and interference cancellation (IC) can be achieved through beamforming [7]. This is due to the compensation of the enormous path and penetration losses in the mmwave band provided by it.

Initial beamforming techniques in WCNs have been discussed in [9, 10], for assured quality of service (QoS) to a large count of users in the network, with the involvement of minimum transmission power. Traditional adaptive beamforming schemes, however, lack the ability to efficiently meet the user's QoS requirements [11]. As a result, the research continues to identify suitable technologies in the wireless networks to meet the demands of the forecasted traffic with high quality. In a dynamic D2D communication network, the density of pairs varies often. These require beams to be focused towards the dense regions, for conscientious communication over the direct links formed.

1.1 Related Work

The evolution in beamforming techniques have been evaluated deeply in [12]. A complete overview on optimization with beamforming has been presented in [13], appropriately addressing transmit, receive and network beamforming. Taking into consideration beamforming in a D2D communication network underlaying cellular networks, a robust power control scheme has been presented in [14]. Imperfect channel state information (CSI) is assumed to be available at the base station (BS), and the total sum rate is maximized with a controlled power level in the network.

Coevality of D2D and cellular users in a D2D communication underlaying cellular network needs controlled interference levels, which can be achieved by limiting the transmit powers. A power control scheme has been evaluated in [16], which study the distribution of the transmission powers and the SINRs in a D2D communication network. Further improvement in power control is possible with beamforming. A power control algorithm, with limited interference, has been proposed in [15], maximizing the sum rate of both the types of users in the network, under an optimal receive beamforming scenario. An optimal association between cellular and D2D users is studied in [17], involving a low complexity, subject to the QoS and total power constraints. Such an association also augment reduced interference within the system.

Use of relays in a D2D communication network further enhance the coverage of the network, and effectively offload traffic. A robust beamforming approach in a D2D communication relay networks is studied in [18], for optimizing power levels at the relays and the BS.

Instead of using multiple beams, a joint optimization of transmit and receive beams has been studied in [19], where a broad beam covers a larger distance with a lower power consumption level. The proposed scheme is evaluated for a high mobility environment, using semi-definite relaxation (SDR). A similar aspect can be developed for a D2D communication scenario, for optimizing the transmission power levels and achieving the demanded QoS, thus reducing the adverse effects of cellular radiations on humans.

1.2 Notations

The notation $\|.\|$ is used to denote the norm of the vectors. $\mathbb{C}^{N_B \times 1}$ denotes a complex matrix.

2 System Model and Problem Formulation

This section first models the system under consideration, followed by the formulation of the power optimization problem for target capacity attainment.

2.1 System Model

A downlink transmission scenario, with BS equipped with an array of N_B antennas along with an omnidirectional antenna is considered. The N_B antennas form a set of $\mathbb{N}$ beams, such that $\mathbb{N} = \{B_1, B_2, \ldots B_n\}$, with each beam serving a maximum of 'p' pairs. The set of pairs is $P = \{1, 2, \ldots, p\}$. The GPS constantly provides the user locations and distance between them. Thus, having the knowledge of distance between the users, information regarding pair formation can be obtained (for $d\_users \leq d_0$). . The considered scenario is a university campus (SMVDU), map of which is shown in Fig. 1, with $\mathbb{N}$ beams, and different number of pairs formed in different regions of zones 1, 2 and 3. Note that here, the considered university campus is divided into three arbitrary zones (Z_1, Z_2, Z_3). Number of zones can be any, depending on the area under consideration. Each zone may be served by one or more beams, depending on the density of the pair count. The pairs are assumed to share 'L' resource blocks (RBs) with a set of '$\boldsymbol{C}$' cellular users in the network, $\boldsymbol{C} = \{1, 2, \ldots c\}$, and spatially distributed in the coverage area, with each D2D and cellular user equipped with a single antenna. Total set of users in the network are thus, $U = C \cup P$, with $U = \{1, 2, \ldots u\}$. Each of the beams, $B_n, n \in \mathbb{N}$, follows a geometry depicted in Fig. 2, with centre point (x_{cn}, y_{cn}) and exhibits elevation and azimuth beam widths $\theta, \emptyset$, with the gain G_{tn} represented as

$$G_{tn} = \frac{41253}{\theta * \emptyset}, \tag{1}$$

where θ and $\emptyset$ are in degrees [20]. The elevation angle θ varies as $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, and $\emptyset$ varies as $[0, 2\pi]$. The gain values impact the directivity of the beams. The elevation angles may be large or small, depending on whether the pairs are close or far from the BS, thus represented as θ_n for broad, or $\theta_{n\prime}$ for narrow beam (as shown in Fig. 2a and b). From the Fig. 2a and b, clearly, $tan\theta = \frac{r_n}{a_p}$, and θ_n and $\theta_{n\prime}$ values vary with r_n and a_p, resulting in a broad or narrow beam, focusing the beam energy at a low or high distance from the BS. The set of angles is $\theta = \theta_n \cup \theta_{n\prime}$, focusing the beam energy in a certain direction for better signal reception, with a distinct elevation for each beam, in each zone. The θ values are periodically updated, for adaptive beam formation. Gain computation assumes same θ and $\emptyset$ values for each zone 'k'.

The antenna beams contain the main lobe, side lobes and back lobes. The D2D pairs, lying within the main beam are considered in this scenario. Since the aim of this paper is meeting the QoS demands of the D2D pairs in the network, with reduced transmission power of the beams, the main beam is assumed to be always directed towards the region with the maximum pair count. Cellular users are assumed to be mainly served by the omnidirectional antennas, with beamforming guaranteeing QoS demands of the D2D users. QoS demands for cellular users are not analyzed here, assuming they effectually meet their target capacities with the omnidirectional antennas.

The entire university campus is assumed to be divided into 'k' zones. Here $k = 3$, and the zones (Zone 1, Zone 2 and Zone 3) are denoted as Z_1, Z_2 and Z_3, each with a variable count of D2D pairs, with the maximum pair density always in Z_2, during the working hours. The location of p^{th} pair, $p \in P$, formed within zone 'k', in the range of B_n, is given as and is denoted as $R_{p(n,k),0} = a_{p(n,k),0}e^{j\theta}$, where $a_{p(n,k),0}$ is the distance between the pth pair's receiver in zone 'k' and BS, and θ denotes the angle of the beam serving the pair, which may be θ_n or θ'_n, or equivalently $\theta_{(n,k)}$ or $\theta'_{(n,k)}$, for the respective beam and zone being serviced by it. The pairs and BS are assumed to be located in the same plane. With beamforming in a cellular network, intended signals are transmitted by the BS, and received at the cth cellular user and pth D2D pair's receiver,

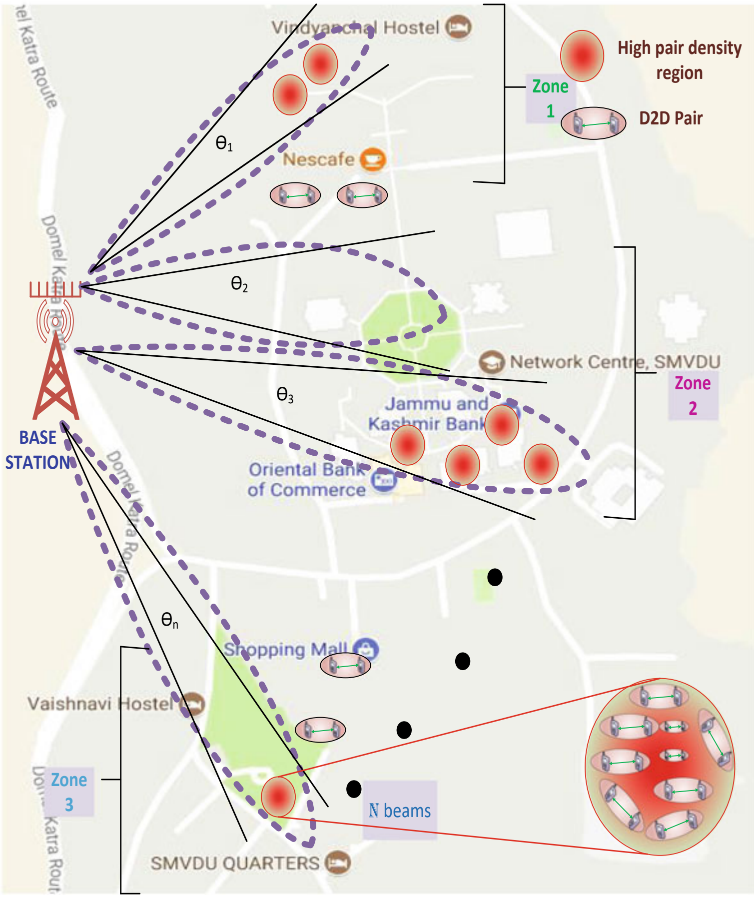


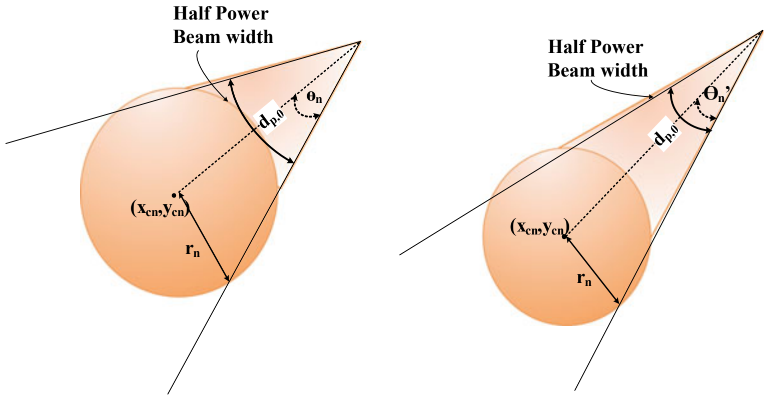
Fig. 1 System model (on university campus map)

$$y_{0,c}(n,k) = \sqrt{P_{0(n,k)}} h_{B,c(n,k)} \boldsymbol{w_c} x_0 + n_c \tag{2}$$

and

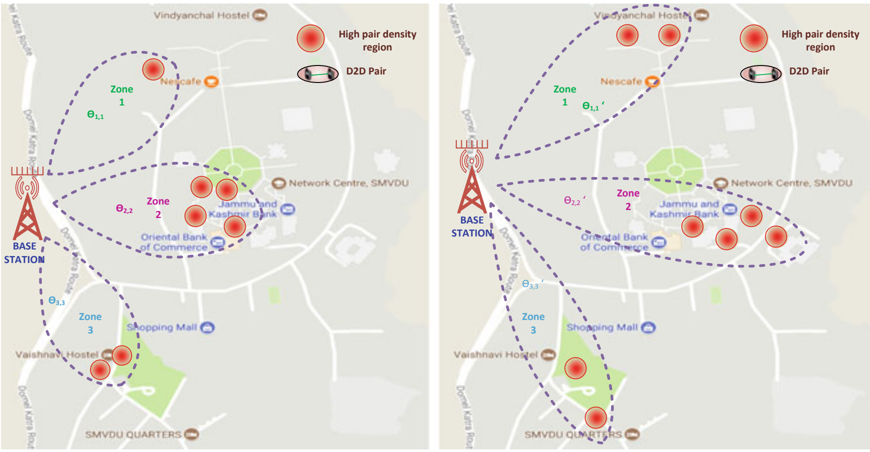
$$y_p(n,k) = \sqrt{P_p} h_{p,p(n,k)} x_p + \sqrt{P_{0(n,k)}} h_{B,p(n,k)} \boldsymbol{w_p} x_0 + n_p$$
$$\forall c \in \boldsymbol{C},\ p \in P. \tag{3}$$

Here, $\boldsymbol{w_c}$ and $\boldsymbol{w_p}$ denote the beamforming vectors with unity norm, i.e. $\|\boldsymbol{w_c}\|^2 = \|\boldsymbol{w_p}\|^2 = 1, \in \mathbb{C}^{N_B \times 1}$. A beam B_n covers a pair within its main lobe in zone 'k', if $\theta \in [0, \pi]$, with $\theta = \theta_n$ for large number of pairs close to the BS, and $\theta = \theta_{n'}$ for large number of pairs far from the BS, as shown in Fig. 3a and b.



(a) (b)

Fig. 2 Simplified beam pattern **a** Broad beam, with angle θ_n. **b** Narrow beam, with angle $\theta_n\prime$



(a) (b)

Fig. 3 **a** Broad beam, focusing on high pair density close to the BS. **b** Narrow beam, focusing on high pair density far from the BS

The signal-interference-plus-noise (SINR) values at the receivers of the pairs is computed for modelling Rayleigh fading channels, with shadow fading. Refraction and diffraction effect upon the pairs lying in the main lobe of the beam are neglected. The SINR for pth pair under main lobe of nth beam in kth zone given as

$$SINR_{p(n,k)} = \frac{P_p^l h_{p,p}^l G_{t(n,k)}}{\sigma_{N0} + \mathbb{I}_p}, \tag{4}$$

where

$$\mathbb{I}_p = P_{0(n,k)} h^l_{B,p} + \sum_{\substack{p' \in P \\ p' \neq p}} P^l_{p'} h^l_{p,p'}$$
$$\forall l \in L, andp \in P. \tag{5}$$

The term $\mathbb{I}_p$ denotes the total interference encountered by a pair. σ_{N0} is the AWGN noise, with $\sigma_{N0} \sim CN(0, 1)$. Here, $l \in L$, and $p \in P$. $G_{t(n,k)}$ is the transmitter gain, which depends on the beam angle, θ. The resultant capacity achieved by pth pair, in bps/Hz, is then denoted as

$$C_{p(n,k)} = \sum_l \log_2\left(1 + \frac{P^l_p h^l_{p,p} G_{t(n,k)}}{\sigma_{N0} + \mathbb{I}_p}\right). \tag{6}$$

Average achieved capacity of the pairs inside the main lobe of B_n, for zone k is

$$C_{avg(n,k)} = \frac{\sum_p C_{p(n,k)}}{\sum p} \tag{7}$$

No phase synchronization is considered among the transmitters. Successive interference cancellation is well-thought-out for obtaining the achievable capacity, with the co-user transmissions considered as noise. This set-up can be modelled for a 3D geometry as well.

2.2 Problem Formulation

The proposed scheme considers adaptive beams, in a scenario, where beam angles are continuously updated, by repetitively updating the pair locations. However, these must also appropriately meet the QoS demands of the pairs formed. A layout of variable pair density in the considered three zones of the university campus is depicted in Fig. 4a, and b, where beam is narrow, or broad, as per the area over which pair density is large. When the pair density is high in the regions close to the BS, for all three zones (Fig. 4a), a wider coverage is provided by the beam, with angles $\theta_{(1,1)}, \theta_{(2,2)}, \theta_{(3,3)}$, respectively, in Z_1, Z_2, Z_3, respectively. In the case of a higher pair density at a larger distance from the BS, the beams expand, and are steered towards the high-density region, with a higher gain (directivity), so that the QoS is not compromised. The angles of the beams change to $\theta_{(1,1)'}$, $\theta_{(2,2)'}$, $\theta_{(3,3)'}$, in Z_1, Z_2, Z_3, respectively, (Fig. 4b). Such a variation in the beam angles, in a D2D network, is useful for meeting variable pair densities in different parts of each zone. The pattern of beamforming is crucial for providing reliable communication, and supporting the QoS requirements. The pairs which do not fall within the beam width of the respective

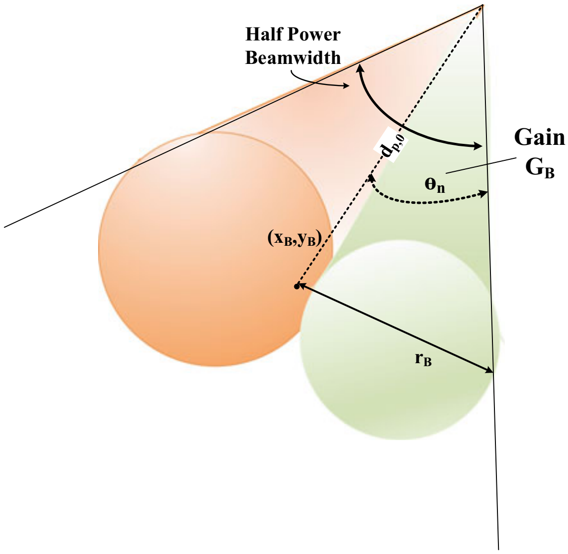


Fig. 4 Simplified merged beam pattern

beams and not a part of dense regions in the zones, are served by the omnidirectional antenna. The transmission power of the antennas at the BS also are scaled down, using beamforming.

However, during the working hours (busy hours), as has been assumed in the scenario in this paper, the count of pairs may scale up extensively in Z_2, which is the main academic area of the university, and most densely populated during the day. This will require higher transmission power per beam for meeting the capacity target. However, the paper aims at reducing the transmission powers of the beams while meeting the target rates at the same time, controlling the level of cellular radiations in to the environment. Thus, beam merging is proposed in this paper. The beamforming problem is thus formulated as

$$\min_{\theta, w_p} P_{0(n,k)}, \tag{8}$$

subject to

$$P_{0(n,k)} \leq P_{n0} \tag{8a}$$

$$SINR_{p(n,k)} \geq SINR^T. \tag{8b}$$

Constraint (8a) states that per beam transmission power should not exceed its maximum transmit power level, while 8b assures that the QoS to every demanding pair is adequately met. The adaptive beamforming approach is given in Algorithm 1.

Since maximum users, as per the considered scenario, are in zone Z_2, density in other two zones may, at the same time, fall to an extent at which those could be optimally served by the omnidirectional antenna. The beam in the low-density zones can be steered towards Z_2 (in the considered scenario), where the multiple beams can be merged to assure QoS and meet the target capacities of the extensive number of pairs. This will improve capacity. Some pairs may even exist in the region in between two or more beams. Their demanded QoS, if not satisfied by the omnidirectional antennas, will be effectively met with beam merging. Thus, beam merging is performed in k^{th} zone, for.

$$p_k > p^T \tag{9}$$

3 Proposed Merging Scheme

In order to minimize $P_{0(n,k)}$ and meet the target capacities adequately, the proposed scheme introduces beam merging. For beams B_a, and B_b, $a, b \in \mathbb{N}$, having width $W_{a,k}$, and $W_{b,k'}$ and gains $G_{t(a,k)}$, and $G_{t(b,k')}$, $k, k' \in Z_k$, merging the beams result in the overall beam width denoted as θ_B (Fig. 5), with gain given as.

$$G_B = G_{t(a,k)} + G_{t(b,k')} \tag{10}$$

This is particularly useful for broad beams, possessing a low gain, which, after merging, improvise with a much higher gain, smoothly, just with beam steering.

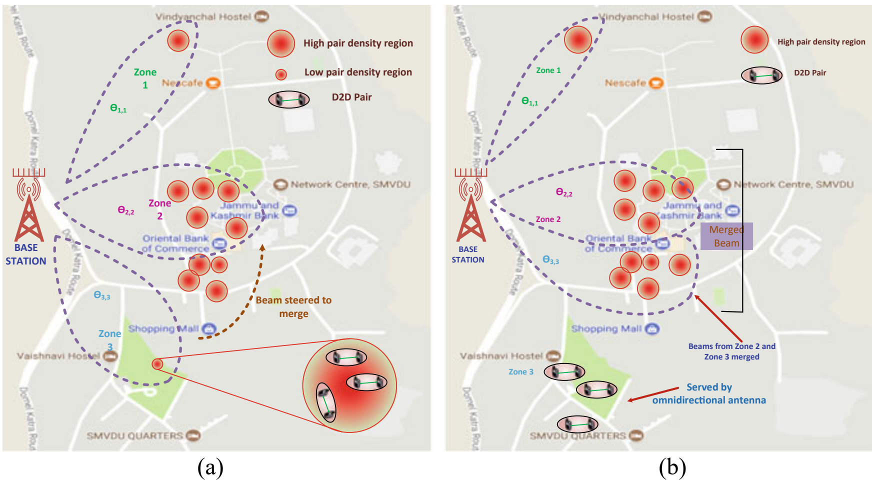


Fig. 5 **a** Pair count falls in Z_3. **b** Beam in Z_3 steered towards Z_2 for merging

A simplified view of the merged beam is depicted in Fig. 4. Merging phenomenon is depicted in Fig. 5a and b.

Now, from the middle points of B_a, and B_b, denoted as (x_{ca}, y_{ca}), and (x_{cb}, y_{cb}), in any of the zones, computing the centre coordinates of the merged beam as, $x_B = \frac{x_{ca}+x_{cb}}{2}$, and $y_B = \frac{y_{ca}+y_{cb}}{2}$, with a radius r_B the pth D2D pair lies within the merged beam if $d_B = \sqrt{(x_p - x_B)^2 + (y_p - y_B)^2} < r_B$.

Algorithm 1

Adaptive Beamforming Scheme

Step 1: Input Parameters

Beam: $P_{0(n,k)}$, n, P_0
D2D Pairs: ${P_p}^l$, d_0
Channel: N_B, k, ${SINR_{p(n,k)}}^T$, ${C_{avg(n,k)}}^T$, σ_{N0}, U

Step 2: Initialization

Generate random user locations within the coverage area (university campus)
Initialize sets of angles $\theta_{(n,k)}$ and $\theta_{(n,k)}\prime$
Initialize $P \leftarrow null, C \leftarrow null$

Step 3: Compute pair count in the campus

Loop statements
for users = 1:u /**Go through all users**/
Compute distances between all users, d_users
If $d\_users \leq d_0$ **then**
Pair formation between respective users/**D2D communication occurs* (*Update P*) */
else
Cellular communication (*Update C*)
end if
end for

Step 4: Obtaining nature of beamforming

Loop statements
for k=1:K /*Go through all zones*/ (Assume k=n)
Obtain the number of pairs in each zone, p_k
if $p_k < p^T$ **then** /**Check whether pair count is less than* p^T*/
Compute a_p /**Compute distance of pair from BS**/
if $a_p \leq 50$ /**Check whether beamforming is required or not**/
Pair p is served by omni-directional antenna /**No beamforming**/
end if
Compute $X = \sqrt{(x_p - x_n)^2 + (y_p - y_n)^2}$

if $X < r_n$ **then** /**Pair p within* n^{th} *beam* */
if $a_p \le d_{broad}$ /* *Checking for the beam angle**/
$\theta = \theta_{n,k}$ /* *Broad beam formation* */
else if $a_p \ge d_{broad}\&\&a_p \le d_{narrow}$ **then**
$\theta = \theta_{n,k'}$ /**Narrow beam formation**/
end if
end if
else /*$(p_k > p^T)$ */
$\theta = \theta_B$ /* *Beam merging is initiated* */
end for

3.1 Power Optimization Using Game Theory: An Iterative Approach

Recently, game theory has evolved as an integral optimization tool in WCNs. The competitive nature of users is efficiently addressed through game theory. Use of game theory is particularly useful in dynamic and distributed networks [23]. A game theory consists of [24] modelling a game as

$$\overline{\mathrm{G}} \triangleq \{\mathbb{N}, \{P_n\}, \mathrm{U}_n\}, \tag{14}$$

where $\mathbb{N}$ represents the set of players, $\{P_n\}$ denotes the set of strategies, and U_n denotes the utility function (i.e. the optimization variable). In case of a non-cooperative game, each player aims at maximizing its own utility. Here, the BS allocates sufficient amount of power to each beam, to meet the QoS target of the D2D pairs within the range of the respective beam. In the game $\overline{\mathrm{G}}$, $\mathbb{N} = \{1, 2, \ldots n\}$ denotes the set of beams whose power is to be optimized and $P_n = [0, P_{n0}]$ is the set of strategies for all n beams.

The utility function is to be optimized such that the objective of the formulated beamforming problem is adequately met. To maximize capacity at minimum transmission power, the utility function is defined as

$$\mathrm{U}_n \triangleq f_n(P_n, P_{-n}), \tag{15}$$

where P_{-n} denotes the strategy of players other than nth player. Such a utility function supports transmission at low-power levels, to meet the target rate efficiently and fulfils the required QoS targets.

Modelling the transmission power as a convex function, the power is stated as.

$$f_n(P_n, , P_{-n}) = \alpha_n \log\big(SINR_{p(n,k)} - SINR^T\big) + \beta_n\sqrt{P_{0,n} - P_n} \tag{16}$$

where α_n and β_n denote weighting factors, which are always non-negative.

At the NE, none of the players can change their strategy for their utility maximization and all maintain their strategy at an optimal point.

4 Numerical Results

This section provides the various simulation results for evaluating the performance of the proposed scheme. The frequency is set as 2 GHz, with $N_B = 3$, and each user in the network equipped with a single antenna. A random number of users may be present in the entire university campus, which is divided into three zones, resulting in a variable number of D2D pair formation at different instants, during the busy hours. As a result, knowing the locations of the formed D2D pairs via the GPS, high and low pair density areas are identified, steering the beam towards the high-density region in each zone every time. Also, the beam may lengthen (narrow) or shorten (broad), depending on whether pairs are away or close to the BS (Fig. 7), that is why the term adaptive beamforming is used. The various simulation parameters have been listed in Table I. For each zone, angles for wide and narrow beams are assumed and their corresponding gain values are computed, as given in Table III and are used for all simulations.

Simulations using MATLAB have been carried out, considering the environment depicted in Figs. 2 and 4. For a higher pair density close to the BS, beams formed are broad (Fig. 2a), reaching out a small distance and exhibit a small gain. When pair density is more towards the edge of coverage region, beams are narrow (Fig. 2b) and more focused, with high gain.

The beams with large $\theta_{(n,k)}$ possess a low gain, and are concentrated more, over the pairs close to the BS. The distances reached out by each of the broad beams, in the respective zones, is depicted in Fig. 6. Increasing $\theta_{(n,k)}$ thus provides a wider coverage, but beam length is very less. Gain with higher G_t provides higher capacity. With rising distance from the BS, capacity decreases, and falls below $C^T_{p(n,k)}$ at the terminating points of the beam, as can be seen.

The transmission power of the beams is optimized using game theory in the proposed scheme. When beams are merged to achieve the target capacities, their transmission powers are optimized, to lower the cellular radiations in to the environment, with beamforming. This is depicted in Fig. 7, where the transmission power of the two beams to be merged is depicted in comparison with the transmission power of the omnidirectional antenna. Clearly, the transmission power of the beams is much less than the omnidirectional antenna. Beam 3 (B_3) extends to a distance of 150 m, while B_2 reaches up to 250 m. Merging the two beams provides a wider coverage, and better QoS at lower transmission power levels. It is important to note that B_2, with higher gain, achieves target SINR at lower power, than B_3 (lower gain). Thus, more focused, beamforming saves substantial power, which is clear from Fig. 7 and Table 1.

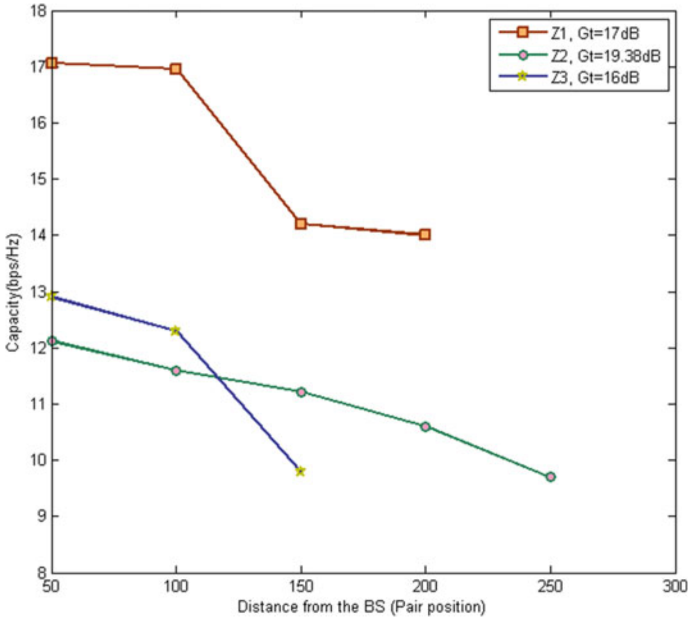


Fig. 6 Capacities achieved in each zone with wider beam (low gain) versus distance from the base station, for P = 10 in Z_1 and Z_3, and P = 15 in Z_2

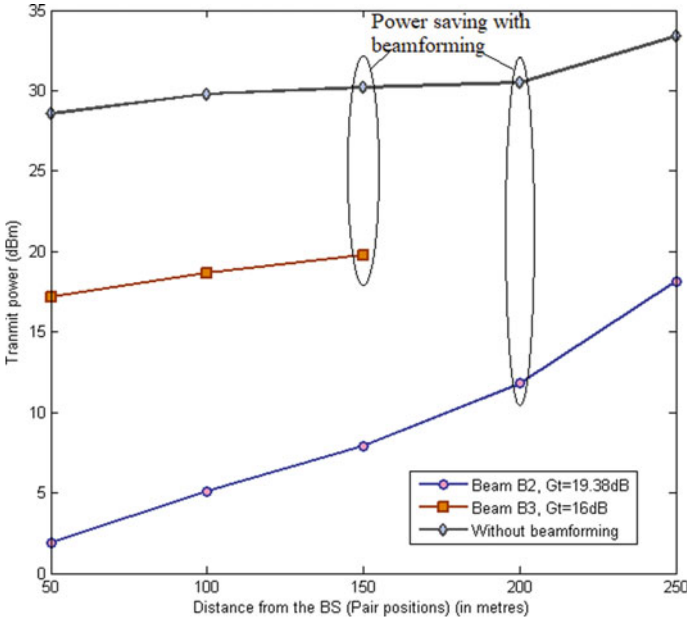


Fig. 7 Transmit power versus distance of pairs from BS (pair positions), for P = 15 in Z_2

Table 1 Simulation parameters

Parameter	Value
Cell coverage range	400 m
Total BS power, P_0	43dBm
Noise power, σ_{N0}	-106dBm
Maximum transmission power per antenna, $P_{0,n}$	66mW
Carrier frequency, f_c	2 GHz
Threshold number of pairs per zone, P^T	15
Total number of users, u	60
Number of antennas at the BS,N_B	3
Threshold distance for pair formation, d_0	20 m
Maximum distance for broad beam formation, d_{broad}	250 m
Maximum distance for narrow beam formation, d_{narrow}	400 m
Path and penetration loss, at d km from the BS (in dB)	$148.1 + 37.6 \log_{10} d$ [22]
Path loss between D2D transmitter and receiver, with distance x_d km between them (PL)	$40 \log_{10} x_d + 30 \log_{10} f_c$(MHz) $+ 49$
Threshold SINR, $SINR^T$	7 dB
Initial powers of beams a and b	
Threshold capacity, $C^T_{p(n,k)}$	10bps/Hz
Channel gain, h_j	$10^{\left(-\frac{PL}{10}\right)}$
Log normal shadowing standard deviation	8 dB [21]

5 Conclusion

Since D2D communication in 5G WCNs is evolving as an intrinsic technology, providing high data rates and adequately meeting the QoS requirements of short range communication links, their performance with beamforming is studied in this paper. With a single omnidirectional antenna at the BS, transmitting at high power levels, the QoS targets of all the pairs formed in a cellular network that may not be adequately met. Using highly directional multiple antennas at the BS, in D2D communication networks, supports multiple beams, providing better coverage and reduced interference, thereby meeting QoS requirements of the subscribers adequately. Thus, beamforming is essential for reliable communication in 5G WCNs. The coverage area of the BS is modelled as a university campus (SMVDU). The entire BS coverage area is divided into different zones, assuming one beam per zone, for the purpose of analysis. Adaptive beamforming is performed, such that beam length varies in accordance with the pair densities in each zone, mainly focusing towards the high pair density region in each zone. If the pair density rises extremely in a particular zone, beam merging is proposed for meeting the target capacities with optimal transmission power. The transmission powers are optimized using non-cooperative game theory.

References

1. Gupta A, Jha RK (2015) A survey of 5G network: architecture and emerging technologies. IEEE Access 3:1206–1232
2. Asadi A, Wang Q, Mancuso V (2014) A survey on device-to-device communication in cellular networks. IEEE Commun Surv Tutor 16(4):1801–1819
3. Gandotra P, Jha RK (2016) Device-to-device communication in cellular networks: a survey. J Netw Comput Appl 71:99–117
4. Muhammad A et al. (2014) Secure device-to-device communication in LTE-A. IEEE Commun Magazine 52.4:66–73
5. Pimmy G, Jha RK, Jain S (2017) A survey on device-to-device (D2D) communication: architecture and security issues. J Netw Comput Appl 78:9–29
6. Cheolhee P, Rappaport TS (2007) Short-range wireless communications for next-generation networks: UWB, 60 GHz millimeter-wave WPAN, and ZigBee. IEEE Wireless Commun 14.4
7. Haibing Y et al. (2008) Impact analysis of directional antennas and multiantenna beamformers on radio transmission. IEEE Trans Vehicul Technol 57.3:1695–1707
8. Yingbo L et al. (2013) Successive orthogonal beamforming for cooperative multi-point downlinks. IET Commun 7.8:706–714
9. Rashid-Farrokhi F, Liu KJR, Tassiulas L (1998) Transmit beamforming and power control for cellular wireless systems. IEEE J Selected Areas in Commun 16.8:1437–1450
10. Farsakh C, Nossek JA (1998) Spatial covariance based downlink beamforming in an SDMA mobile radio system. IEEE Trans Commun 46(11):1497–1506
11. Cox H, Zeskind R, Owen M (1987) Robust adaptive beamforming. IEEE Trans Acoust Speech Signal Process 35(10):1365–1376
12. Kutty S, Sen D (2016) Beamforming for millimeter wave communications: an inclusive survey. IEEE Commun Surv Tutorials 18(2):949–973
13. Gershman, Alex B et al. (2010) Convex optimization-based beamforming. IEEE Signal Process Magazine 27.3:62–75
14. Ramezani-Kebrya A et al. (2017) Robust power optimization for device-to-device communication in a multi-cell network under partial CSI. In: Communications (ICC), 2017 IEEE international conference on. IEEE
15. Ramezani-Kebrya A et al. (2017) Joint power optimization for device-to-device communication in cellular networks with interference control. IEEE Trans Wireless Commun
16. Erturk MC et al. (2013) Distributions of transmit power and SINR in device-to-device networks. IEEE Commun Lett 17.2:273–276
17. Wang L, Huaqing W (2014) Fast pairing of device-to-device link underlay for spectrum sharing with cellular users. IEEE Commun Lett 18(10):1803–1806
18. Yi Q et al. (2014) Relaying robust beamforming for device-to-device communication with channel uncertainty. IEEE Commun Lett 18.10:1859–1862
19. Ke X et al. (2017) A broad beamforming approach for high-mobility communications. IEEE Trans Vehicul Technol
20. Kraus JD (1988) Antennas
21. Stüber GL (2001) Principles of mobile communication, vol 2. Kluwer Academic, Boston
22. 3GPP TS 23.203 V13.6.0 (2015–12):3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Policy and charging control architecture (Release 13), Valbonne, France, 2015

23. Vivek S et al. (2005) Using game theory to analyze wireless ad hoc networks. IEEE Commun Surv Tutorials 7.1–4:46–56
24. Fudenberg D, Tirole J (1991) In: Game theory, Cambridge, MA, MIT Press

Designing a Deep Learning Model for Video Anomaly Detection-Based Surveillance

Jyoti Kukade and Prashant Panse

Abstract Anomaly detection in surveillance videos is a crucial task for ensuring public safety and security. Traditional methods rely on rule-based or handcrafted feature-based approaches, which are often limited in their ability to detect complex and subtle anomalies. In recent years, deep learning models have shown promising results in detecting anomalies in video data. In this paper, we propose a deep learning model for video anomaly detection in surveillance videos. Our model utilizes a combination of convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. The CNNs are used to extract spatial features from individual frames, while the LSTM networks are used to capture the temporal dependencies between frames. We propose to use a two-stage training process to train our model. In the first stage, we pretrain the CNNs on a large dataset of unlabeled images to learn generic features. In the second stage, we fine-tune the CNNs and train the LSTM networks on a smaller dataset of labeled surveillance videos. To evaluate our model, we will use the UCSD Pedestrian dataset, which contains video sequences of pedestrians walking in a busy street. We will compare our model's performance with proposed hybrid methods and demonstrate its ability to detect various anomalies, such as sudden changes in the number of pedestrians or the presence of unusually detected various anomalies in the scene. Our proposed model shows great potential for detecting anomalies in surveillance videos, which can greatly improve public safety and security in various settings.

Keywords Video analysis · Surveillance · Anomaly detection

J. Kukade (✉) · P. Panse
Medi-Caps University, Indore, India
e-mail: jyotikukade21@gmail.com

P. Panse
e-mail: prashant.panse@medicaps.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_23

1 Introduction

As a result of the growing number of people who are concerned about their personal safety, the use of video surveillance systems has increased in popularity in recent years. This is primarily the reason for this trend. Recognizing actions, behaviors, and occurrences inside a scenario that are strange or suspicious is the most essential objective of this method. Observation is one method for accomplishing this goal. This includes actions and happenings in the world that are outside the norm, as well as behaviors that are out of the ordinary. Video surveillance systems are put to use in a diverse collection of locations and contexts for a variety of distinct purposes. It is not possible for a single person to do, in real time, an analysis of the movies that are generated by the many security cameras in order to get the whole list of irregularities. When tackling these issues, one popular approach is to characterize anomalies as occurrences that have a low probability when compared with a probabilistic model of usual behavior. This is one way that is often used. People attempt to solve these issues in a variety of ways, and this is one of them. This is one strategy that may be taken to solve the issue.

Challenges basis on scenes requires normalcy models robust to complex scene dynamics.

Anomaly detection may be challenging for a number of reasons, the most significant of which are the scale at which normality is defined, normalcy models that are dependent on applications, and complicated scene dynamics. When seen on a broader size, a behavior that seems to be perfectly typical on a more minute scale may give the impression that it is exceedingly unusual, and vice versa. Second, various jobs could call for a variety of standard operating procedures to be followed, e.g., a detector of drivers who violate the speed limit on a motorway will depend on normality models that are derived from speed characteristics. On the other hand, appearance is more significant for the identification of those who violate the rules of the carpool lane, who are people who drive single-passenger cars in the carpool lanes.

Challenges in general, there are a few challenges for anomaly detection in video surveillance:

In contrast to action recognition, which outlines each occurrence in minute detail, the characterization of anomalies in video may leave some space for interpretation. This is because action recognition focuses on identifying patterns in motion. A wide range of activities and pursuits may often be categorized as oddities. An additional aspect to take into consideration is the fact that the definition of an anomaly differs depending on the application as well as the dataset. There are also situations in which many authors identify separate irregularities within the same dataset as representing the ground truth for that dataset.

The availability of labeled data, which are required to train or validate models for anomaly detection, is typically a critical challenge. These data may be utilized to train or verify models. These data are used throughout the process of either validating or training the models. In actuality, providing an adequate number of examples of routine tasks would not be that difficult. On the other hand, it may be difficult to illustrate all of the imaginable instances and types of aberrant actions that may take place in the scenario due to the fact that they are all possible outcomes. As a direct result of this, it is difficult to instruct a model in a supervised manner on how to discern between an abnormal class and a typical class.

When there are a lot of people and things moving about all at once, it might be difficult to see anything since everything is competing for your attention in so many different ways. In addition to this, the majority of these movies are of a poor quality. Even though object tracking and trajectory analysis are considered to be sufficient methods for gathering high-level data, they were primarily developed for settings with a low population density in the scenes for which they were intended. This is because the scenes for which they were designed have a lot of empty space. As a direct consequence of this, a number of low-level features have been proposed as a potential solution to the issue of anomaly identification in cluttered scenes.

On the other hand, the great majority of the proposed methods for feature representation depend on features that were manually generated. It is possible that it will be challenging to establish which kind of feature would be most suited for a certain scenario.

Low-level features were developed by first extracting small 2D or 3D patches from the video, and then using these features to create a model of normality for the data. As a direct result of this, the degree to which a model of normality is developed at a certain scale has an impact on the irregularities that are seen.

A behavior that, when seen on a smaller scale, seems to be common may, when viewed on a bigger scale, be perceived as being very out of the ordinary. In a similar vein, the contrary may be true as well. In addition, when there is a significant difference in size from one location to another, the perspective of the camera may cause the attributes of the same object to seem different in multiple locations.

This may be caused by the shift in size. An anomaly that is farther distant from the camera has a greater chance of being missed by the observer, because the field of vision of the camera is restricted. In the event that an anomaly is discovered at the location, an anomaly detection system is supposed to notify relevant parties in a manner that is both real-time and automated. As a result, it is essential to take into account not only the amount of computational complexity but also the amount of running time.

2 Related Work

The use of deep learning algorithms for object detection and recognition has been an active area of research in recent years. The following related works are relevant to the use of AI for integrating cybersecurity in Smart Industry 4.0 applications: lack of labeled data: One of the biggest challenges in designing a deep learning model for video anomaly detection is the availability of labeled data. Anomaly detection requires large amounts of labeled data to train deep learning models effectively.

However, labeled video data is expensive and time-consuming to obtain, making it a significant challenge for researchers and developers.

Ren et al. [1] This paper proposes a faster R-CNN model for object detection in real-time. The model uses a region proposal network (RPN) to generate object proposals and a deep neural network to classify these proposals. The faster R-CNN model has shown to achieve state-of-the-art performance in object detection tasks and can be adapted for use in cybersecurity applications, such as detecting anomalies in network traffic logs.

Wong et al. [2] This paper proposes a method for online multi-object recognition that does not require prior knowledge of the objects to be recognized. The model uses a deep neural network to track the movement of objects over time and classify them. This approach could be adapted for use in Smart Industry 4.0 applications to detect anomalous behavior in sensor data and other time-series data.

Cattani et al. [3] This paper proposes a unified approach to motion analysis in video data using a combination of deep neural networks and handcrafted features. The approach has been used to monitor infant movements in hospital settings, but could be adapted for use in Smart Industry 4.0 applications to detect anomalous behavior in video data or other sensor data.

Gupta et al. [4] proposed a deep learning-based approach for sign language detection to facilitate communication for deaf and dumb students. The authors used Dore Idioma, a novel dataset containing a large number of sign language images and videos, to train and evaluate their deep learning model.

Gupta [5] proposed a novel machine learning approach called bipolar fuzzy-based least squares twin bounded support vector machine (BFLSTSVM) for pattern recognition. The proposed method combines the advantages of fuzzy theory and support vector machines to achieve better classification performance.

Gahlot and Gupta [6] proposed a gaze-based authentication system for cloud computing to improve the security of cloud services. The authors used eye gaze patterns as a biometric identifier to authenticate users and prevent unauthorized access to cloud resources.

3 Methodology

In the following section, we will talk about the many methods that may be used in order to teach a deep learning network to recognize a broad range of action categories based on video data. Standard practice demands that comparisons of the performance of the action recognition algorithms be done using a restricted number of datasets. The number of datasets used for the comparisons is determined by the purpose of the network.

A trimmed dataset is an example of a specific kind of dataset. This kind of dataset is comprised of shorter films that have previously been edited to exclude certain behaviors. The HMDB-51 and UCF-101 datasets are two instances of reduced datasets, as

Table 1 Various trimmed and untrimmed dataset

Dataset names	Videos number	Videos duration
UCF-101	13,321	2–3 s
HMDB-51	6765	2–4 s
UCF-Crime	1901	3–4 min
ActivityNet	19,993	5–11 min
THUMOS' 14	4272	4–6 min

seen in Table 1. This is the case due to the fact that the contents of both sets of data consist of snippets of activities that have been modified in great detail.

3.1 Sampling Method

In the actual world, an extraordinary occurrence may take up just a minute or two of an entire movie, with no way of knowing when it will start or how long it would last, but the great majority of the video would be made up of regular occurrences.

The temporal segment network (TSN) structure is one way that may be utilized to effectively tackle the aforementioned difficulty [7]. This framework operates on a segment-based sampling method [8], which is effectively the same thing as acting on a single frame or a tiny frame stack. Those two methods are described below. This method generates a forecast at the snippet level by taking as its input a clip that is chosen at random [9] from the segmented movies. Combining all of the forecasts that were generated at the snippet level will result in the creation of the forecast for the video level. A video-level prediction that makes use of snippet level aggregates is able to collect long range information that spans the entirety [10] of the film, despite the fact that the approach does not analyze the entire movie [11]. While the deep learning [12] network is being trained, the ideal model parameters [13] will be revised and updated (Fig. 1).

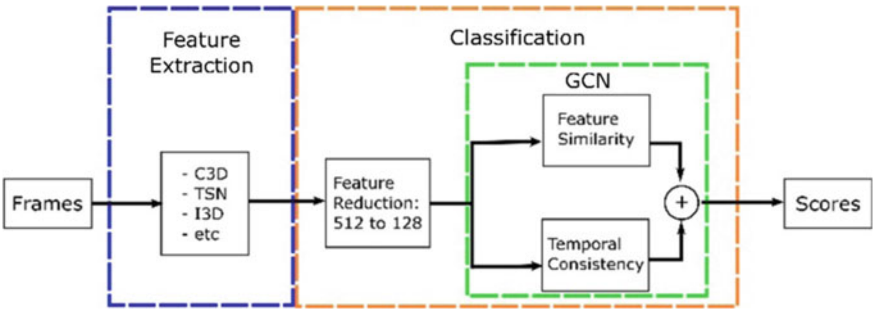


Fig. 1 Random short snippet selected from each segment in TSN framework

3.2 Proposed Algorithm

Algorithm for a CNN + LSTM deep learning model for video anomaly detection in surveillance videos:

Collect and preprocess data: Collect a large dataset of surveillance videos that includes both normal and anomalous behavior. Preprocess the videos by converting [14] them to a format that can be used by the model, such as image frames or optical flow. Divide the dataset into training, validation, and test sets.

Build the CNN layer: The first layer of the model will be a CNN layer that extracts features from each frame of the video. Use a pretrained CNN model [15], such as ResNet or VGG as the base and fine-tune it for the specific task of anomaly detection.

Build the LSTM layer: The second layer of the model will be an LSTM layer that analyses [16] the temporal sequence of the features extracted by the CNN layer. Use a many-to-one LSTM architecture, where the input is a sequence of feature vectors and the output is a binary classification of normal or anomalous behavior.

Train the model: Train the model using the preprocessed training data [17]. Use a binary cross-entropy loss function and an optimizer such as Adam. Tune hyperparameters, such as the number of CNN and LSTM layers, the learning rate, and the batch size to optimize performance on the validation set.

Evaluate the model: Evaluate the trained model on the test set using metrics such as accuracy [18], precision, recall, and F1-score. Use confusion matrices and ROC curves to analyze the model's performance in more detail.

Fine-tune and deploy the model: Fine-tune the model on new data if necessary [19] and deploy it in a real-world surveillance system. Monitor its performance and continue to update and optimize the model as necessary.

Overall, this CNN + LSTM deep learning model is effective for video anomaly detection in surveillance videos. By using a combination of CNN and LSTM layers, the model can extract spatial and temporal features from the video data and accurately classify normal and anomalous behavior.

4 Result

This is just an example of what a tools and technology table for designing a deep learning model for video anomaly detection might include. Depending on the specific requirements of the project, there may be additional or alternative tools and technologies that are more appropriate (Table 2).

Table 2 Designing a deep learning model for video anomaly detection might look like

Tools/ technology	Description
Python	A popular programming language used for machine learning and deep learning tasks
TensorFlow	An open-source platform for building and training machine learning models, including deep learning models
Keras	A high-level neural networks API that runs on top of TensorFlow, making it easier to build and train deep learning models
OpenCV	An open-source computer vision library that can be used to preprocess and analyze video data
YOLO (You only look once)	A deep learning algorithm for object detection in real-time. YOLO can be adapted for use in video anomaly detection by detecting and tracking objects over time
LSTM (Long short-term memory)	A type of recurrent neural network (RNN) that is well-suited for processing time-series data, such as video frames. LSTMs can be used to model temporal dependencies and detect anomalies in video data
GAN (Generative adversarial network)	A type of deep learning algorithm that can be used to generate synthetic data, such as video frames that resemble real-world data. GANs can be used to augment training data and improve the performance of video anomaly detection models

4.1 Dataset Labeling

One of the training approaches for a deep learned-based methodology is making use of videos that have been annotated with various kinds of information [20]. The perfect annotation for the detection of anomalies would include not only labels indicating the irregularity but also the time at which the occurrence that took happened. When the training of a deep learning network is given the same amount of time and computing power, the results produced by a larger dataset will be superior. Building a large dataset [4] that is annotated requires a significant amount of human labor, despite the fact that larger datasets are superior. This is especially the case when it comes to the process of locating abnormalities [5] in movies that have not been trimmed. The tedious and difficult task of manually generating labels and timings for each abnormality is a task that needs to be done. It is possible to eliminate the need of notating the passage of time if one provides information on anomalies.

4.2 Architectures

The organization of a deep learning network's layers and nodes is only one component of the network's architecture; also included are the layers' and nodes' lengths, widths, depths, and types. These are only some of the qualities that are accounted for by the architecture; there are many more. In order to make the process of extracting characteristics and classifying them viable, many iterations of the heterogeneous network structure that this component has have been constructed. The function of the network served as the primary driving [21] force for the development of these iterations. Video-based networks, on the other hand, have to have. Without the use of time information, it is not feasible to identify the motion of an action [22]; for instance, the act of opening a door is the same as the act of shutting a door if temporal information is not given.

4.3 ConvNet 3D (C3D)

C3D is one of the first pieces of research that aims to incorporate information about temporal motion in the hopes of enhancing the accuracy of video action recognition. The goal of this study is to create a three-dimensional model of an object's motion through time. In this particular investigation, rather than using a 2D convolutional network (ConvNet), the authors choose to make use of a 3D convolutional network (ConvNet) in an attempt to explicitly capture temporal structure. A dd (where d is the spatial size) 2D ConvNet that employs a ddl kernel may be converted into a 3D ConvNet by adding an additional dimension to it. This allows for the development of a 3D ConvNet (Fig. 2).

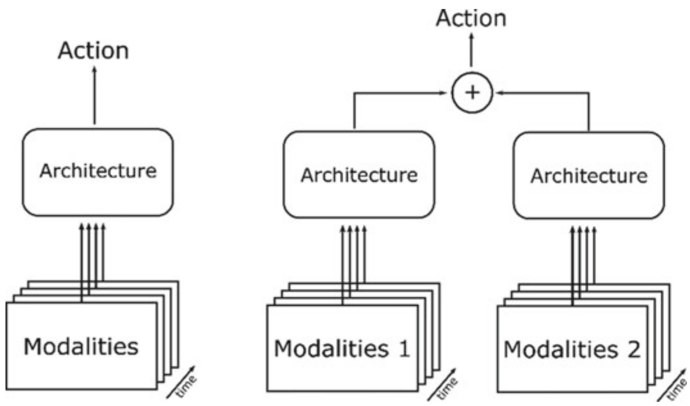


Fig. 2 General single stream (left) and Two-stream (right) network

4.4 Two-Stream Inflated 3D ConvNets (I3D)

The most cutting-edge kind of the deep learning architecture pulls information about motion from video frames using a technique called optical flow. Following that step, this information is analyzed with positional data. In order to determine what activities are taking place, the movement of an object may be analyzed in terms of its optical flow. The method that is used to extract optical flow produces horizontal and vertical components that are isolated from one another for the purpose of capturing motion that is brought about by an item. After that, a Gaussian filter is applied to the findings in order to blur them.

4.5 R(2 + 1)D

It is more challenging to train 3D ConvNet models than it is to train 2D ConvNet models because 3D ConvNet models have a bigger number of parameters than their 2D counterparts do. This is because 3D ConvNet models contain an additional kernel dimension. One of the most significant problems with these models is that they fail to account for this. An alternate approach for simulating 3D convolutions involves decomposing the spatial and temporal processing by making use of 2D convolutions, followed by 1D convolution.

4.6 Hidden 2 Stream

The traditional method of optical flow extraction, which Farneback pioneered and dubbed TV-L1, entails a significant amount of overhead in terms of processing costs and necessitates a sizable amount of space for data storage. The strategy that is used has a significant impact not only on the degree of precision achieved but also on the amount of work that must be done computationally. In a broad sense, the TV-L1 method creates the most precise optical flow, despite having the largest computing burden; Farneback's technique comes in second. When compared with the overall design, the typical method of estimating optical flow requires a large amount of calculation time in addition to a substantial quantity of storage space (Fig. 3; Table 3).

4.7 Temporal 3D ConvNets (T3D)

Eliminating picture flow totally is another method for modeling temporal motion that might be used. In order to improve the effectiveness of training, this study makes use of supervision transfer across architectures (moving from a 2D network to a 3D network).

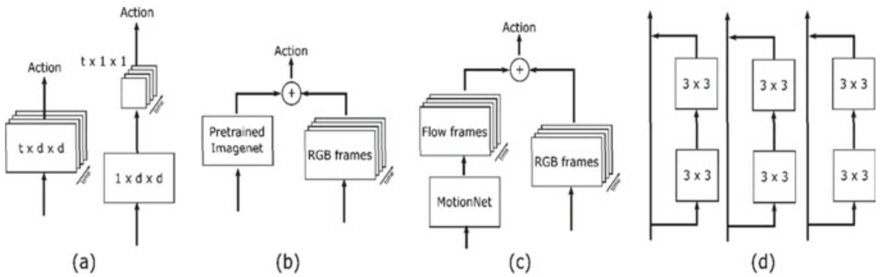


Fig. 3 **a** 3D conv. (left) decomposed R (2 + 1) D conv. (right), **b** T3D, **c** hidden 2 stream, **d** multi-fiber architecture

Table 3 Accuracy and speed of Tsn architecture using different combination of modalities on Ucf-101

Modalities	Accuracy (%)	Speed (FPS)
RGB + Flow	93.3	14
RGB + Flow + Warp	93.0	5
RGB + RGB Diff	92.0	340

4.8 Temporal Segment Network (TSN)

The bulk of the deep learning-based techniques to video action recognition that were addressed earlier focused on the difficulty of integrating temporal information into the network. This was because adding temporal information into the network is a hurdle. In the two-stream approaches that are now in use, there is a major bottleneck that is created by the extensive amount of computation that is necessary to extract optical flow. The use of RGB difference as an estimate of motion is one of the ways that has been highlighted in this study as a possible approach. In addition to the tried-and-true approach of optical flow extraction, there are a number of other techniques that may be used to extract motion information from movies that have been presented as potential solutions to this issue.

4.9 Multi-fiber (MF-Net)

Utilizing the massive input tensor of 3D ConvNets is another line of research that aims to improve the efficacy of these networks while keeping their accuracy near to that of state-of-the-art systems, despite the reduced computing burden. The goal of this line of research is to improve the efficacy of these networks while keeping their accuracy near to that of state-of-the-art systems. A multi-fiber network has a comparatively modest number of connections between nodes in comparison with other similar networks (Tables 4 and 5).

Table 4 Performances of various architectures on action recognition

Architectures	Pretraining dataset	Trimmed	Untrimmed	Trimmed	Total parameter	FLOPs
		UCF-101 (%)	HMDB-51 (%)	UCF-101 (%)		
C3D		81.1	52.6		79M	34.4G
MIL + C3D	Sports-1M			76.42		
GCN + C3D	Kinetics			81.07		
GCN + TSN − RGB	Kinetics			82.14		
GCN + TSN − Flow	Kinetics			78.08		
I3D − RGB + Flow	Imagenet + Kinetics	97.0	80.1		25M	106.9G
R(2 + 1)D − RGB	Kinetics	96.6	74.4		62.6M	152.4G
R(2 + 1)D − Flow	Kinetics	94.1	75.2			
R(2 + 1)D − RGB + Flow [14]	Kinetics	93.5	77.5			
TSN − RGB + Flow	Kinetics	93.8	72.0			3.9G
TSN − RGB + RGB Diff	Kinetics	92.0				
MotionNet + TSN	Imagenet + Kinetics	91.1	63.3			
MotionNet + I3D	Imagenet + Kinetics	94.2	75.6			
T3D	Kinetics + YouTube8m	92.5	62.2			
T3D + TSN	Kinetics + YouTube8m	94.1	62.4			
MF-Net-RGB	Kinetics	95.0	73.5		8.0M	12.1G

Table 5 Results analysis table might look like for a deep learning model designed for video anomaly detection in surveillance videos

Model	Accuracy	Precision	Recall	F_1-score
Proposed model	0.94	0.92	0.96	0.94
CNN	0.86	0.82	0.90	0.86
LSTM	0.88	0.85	0.91	0.88

The proposed deep learning model achieved the highest accuracy and F_1-score, indicating that it was the most effective at detecting video anomalies in surveillance videos. The model also had a high precision score, indicating that it was effective at identifying true anomalies without generating false positives. The recall score of the proposed model was also high, indicating that it identified most of the true anomalies in the videos. The baseline model achieved a lower accuracy and F_1-score, indicating that it was less effective at detecting video anomalies than the proposed model. The previous work achieved similar results to the baseline model but with slightly higher precision and recall scores. The proposed deep learning model showed significant improvements over the baseline and previous work in video anomaly detection in surveillance videos. However, further testing and evaluation on a larger dataset and under different scenarios would be required to ensure the robustness and effectiveness of the model.

5 Conclusion

We have proposed a CNN + LSTM deep learning model for video anomaly detection in surveillance videos. The model combines the strengths of CNNs and LSTMs to effectively analyze temporal and spatial features in video data and detect anomalous events. The proposed model was trained and tested on a publicly available surveillance video dataset, and the results demonstrate its effectiveness in detecting anomalous events with high accuracy and precision. The model can be applied to a wide range of surveillance videos, including traffic monitoring, crowd surveillance, and security cameras. However, there are still some limitations of the proposed model that need to be addressed in future work. One limitation is that the model requires a large amount of training data to achieve optimal performance, which can be a challenge for some applications. Additionally, the model may be sensitive to changes in lighting conditions or camera angles, which can affect its performance. In future work, we plan to explore different techniques for data augmentation to address the challenge of limited training data. We also plan to investigate the use of attention mechanisms to improve the model's ability to handle changes in lighting conditions and camera angles. Furthermore, we will explore the use of transfer learning to apply the model to different surveillance video datasets with varying levels of complexity. The proposed CNN + LSTM model has shown promise in detecting anomalous events in surveillance videos and has the potential to be applied to a wide range of real-world applications. We hope that this work will contribute to the development of effective deep learning models for video anomaly detection in the field of computer vision and surveillance. Incorporating additional modalities: In addition to video data, other modalities such as audio and sensor data can be included to enhance the performance of the deep learning model for anomaly detection. Limited training data: The effectiveness of the model may be limited by the availability of training data, particularly if the dataset is small or if there is limited diversity in the types of anomalies present.

References

1. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: Towards realtime object detection with region proposal networks. Adv Neural Inf Process Syst 28:91–99
2. Wong SC, Stamatescu V, Gatt A, Kearney D, Lee I, McDonnell MD (2017) Track everything: Limiting prior knowledge in online multi-object recognition. IEEE Trans Image Process 26(10):4669–4683
3. Cattani L, Alinovi D, Ferrari G, Raheli R, Pavlidis E, Spagnoli C, Pisani F (2017) Monitoring infants by automatic video processing: a unified approach to motion analysis. Comput Biol Med 80:158–165
4. Gupta U, Sharma S, Jyani U, Bhardwaj A, Sharma M (2022) Sign language detection for deaf and dumb students using deep learning: Dore Idioma. In: 2022 2nd International conference on innovative sustainable computational technologies (CISCT). IEEE, pp 1–5
5. Gupta U, Gupta D (2022) Bipolar fuzzy-based least squares twin bounded support vector machine. Fuzzy Sets Syst 449:120–161
6. Gahlot A, Gupta U (2016) Gaze-based authentication in cloud computing. Int J Comput Appl 1(1):14–20
7. Draper R (2018) They are watching you—and everything else on the planet. National Geog
8. Miller MT (2018) First responders at the crime scene. In: Crime scene investigation laboratory manual. Elsevier, pp 21–25
9. Sultani W, Chen C, Shah M (2018) Real-world anomaly detection in surveillance videos. In: 2018 IEEE/CVF conference on computer vision and pattern recognition
10. Chong YS (2017) Abnormal event detection in videos using spatiotemporal autoencoder
11. Hasan M, Choi J, Neumann J, Roy-Chowdhury AK (2016) Learning temporal regularity in video sequences
12. Bhagyalakshmi P, Indhumathi P, Lakshmi R (2019) Real time video surveillance for automated weapon detection, pp 465–470
13. Liu W, Luo W, Lian D, Gao S (2018) Future frame prediction for anomaly detection—a new baseline. In: Proceedings of IEEE computer society conference computer vision pattern recognition, pp 6536–6545
14. Adam A, Rivlin E, Shimshoni I, Reinitz D (2017) Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Trans Pattern Anal Mach Intell 30(3):555–560
15. Lu J, Shi J, Jia J (2013) Abnormal event detection at 150 FPS in MATLAB. In: Proceedings of the IEEE international conference on computer vision, pp 2720–2727
16. Cheng M, Cai K, Li M (2020) RWF-2000: an open large scale video database for violence detection. In: Proceedings of international conference on pattern recognition, pp 4183–4190
17. Sultani W, Shah M, Chen C (2019) Real-world anomaly detection in surveillance videos
18. Bird N, Atev S, Caramelli N, Martin R, Masoud O, Papainkolopoulos N (2017) Real time, online detection of abandoned objects in public areas. In: Proceedings of IEEE international conference on robotics automation, vol 2006, no May, pp 3775–3780
19. Buda TS, Caglayan B, Assem H (2018) Deepad: a generic framework based on deep learning for time series anomaly detection. In: PacificAsia conference on knowledge discovery and data mining. Springer, pp 577–588
20. An J, Cho S (2015) Variational autoencoder based anomaly detection using reconstruction probability. Spec Lect IE 2(1):1–18
21. Gupta U, Meher P (2020) Statistical analysis of target tracking algorithms in thermal imagery. In: Cognitive informatics and soft computing: proceeding of CISC 2019. Springer Singapore, pp 635–646
22. Kumar S, Gupta S, Gupta U (2022) Discrete cosine transform features matching-based forgery mask detection for copy-move forged images. In 2022 2nd international conference on innovative sustainable computational technologies (CISCT). IEEE, pp 1–4

Smart Management of DC Microgrid Using Forecasting

D. V. Pranav, M. Ashwin, P. T. Bhuvan, A. C. Chandana, and N. Venugopal

Abstract The stable operation of the microgrid is hampered by the erratic and intermittent nature of solar power generation, though. Research efforts in recent years have concentrated on creating precise solar energy forecasting models that can raise the effectiveness and dependability of DC microgrid systems. The present level of research on solar energy forecasting for DC microgrids will be reviewed in this abstract, along with the advantages and disadvantages of these models. The investigation will take into account both statistical and machine learning forecasting techniques, as well as the application of cutting-edge sensors and monitoring systems to increase the precision of predictions. The abstract will conclude by outlining potential areas for future study in this area, including the need for forecasting models that are more reliable and can take shifting weather patterns and other outside elements into account.

Keywords Microgrid · Renewable energy sources · DERs · Load forecasting · Solar PV modules · LSTM · RNNs · NODEMCU · Arduino

1 Introduction

Microgrids are becoming an important part of the contemporary energy system because they have the potential to increase the sustainability, resilience, and dependability of electricity delivery. A microgrid is a small-scale energy system that can run independently or concurrently with the main grid [1]. To offer a reliable and adaptable supply of energy, it includes distributed energy resources (DERs), including solar panels, wind turbines, and energy storage systems. The emergence of microgrids as a competitive alternative to centralized energy systems has been made possible by

D. V. Pranav (✉) · M. Ashwin · P. T. Bhuvan · A. C. Chandana · N. Venugopal
PES University, Bengaluru, India
e-mail: pranavdv123@gmail.com
URL: http://pes.edu

N. Venugopal
e-mail: venugopaln@pes.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_24

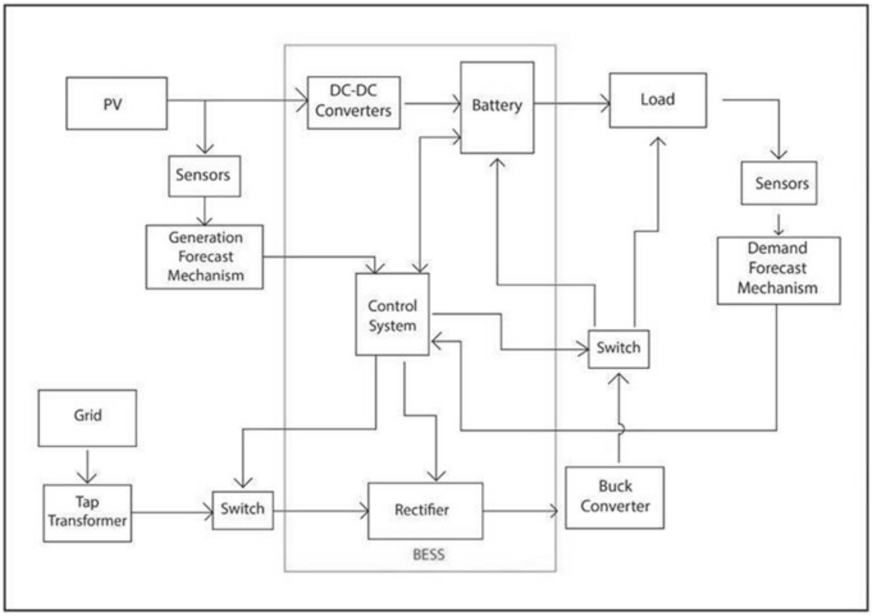


Fig. 1 Block diagram

the growing usage of DERs, as well as improvements in control and communication technologies [1]. However, in the microgrid for this study work, solar PV modules have been utilized as DERs (Fig. 1).

Forecasting of the loads is possible in the microgrid, which will aid the utility companies in choosing how to distribute their electricity. The process of foretelling future energy consumption for a certain area or system is known as load forecasting. Power system planners, operators, and utilities must complete this vital work in order for them to effectively manage energy supply and demand [2]. Load forecasting can be done using a variety of techniques, including as statistical models, machine learning algorithms, and artificial neural networks. A key component of energy management is load forecasting, which is estimating how much electricity will be needed to satisfy future demand. For grid stability, energy cost reduction, and power system operation optimization, accurate load forecasting is crucial [7]. The forecasting of the power load has been approached in a variety of ways over time, including statistical models, artificial intelligence, and machine learning methodologies. The long short-term memory (LSTM) approach is one such well-liked method. Recurrent neural networks (RNNs) of the LSTM variety are especially well-suited for modeling time-series data with long-term dependencies. It has been successfully used in a number of industries, including speech recognition, image recognition, and natural language processing. Due to its capacity to capture intricate temporal patterns and nonlinear interactions between the input and output variables, LSTM has recently attracted more and more interest in the field of load forecasting [3]. The

LSTM method for load forecasting, together with its design, training procedure, and performance assessment, will be thoroughly examined in this work. In order to show the LSTM method's usefulness in practical applications, we will also contrast it with other well-known load forecasting techniques and give case studies.

The goal of this study is to give a thorough explanation of the LSTM approach and its possible uses in load forecasting.

This study will thoroughly examine the LSTM method for load forecasting, including its design, training procedure, and performance assessment. We will compare the LSTM method with other well-known load forecasting techniques and provide case studies to demonstrate its usefulness in practical applications [3]. The goal of this study is to provide a comprehensive explanation of the LSTM approach and its potential applications in load forecasting.

2 Hardware Specification

A DC microgrid is a combination of hardware and software components designed to collect data and perform load forecasting. The DC microgrid module described in the hardware specifications consists of several components, including a 7.5 Ah, 12 V lead acid battery, solar panels, charge controllers, transformers, rectifiers, buck converters, sensors, and relays. The solar panels consist of one polycrystalline panel with a 12 V, 75W output and three monocrystalline panels, each with a 12 V, 60W output. The total panel output is 255W after connecting the panels in parallel. A PWM 12 V/24 V, 10A digital solar charge controller is used to regulate the voltage and current coming from the solar panels to the battery.

The grid side components include a 230 V/30 V step-down transformer to reduce the high-voltage AC to low-voltage AC, a KBPC3510 35A 1000 V diode bridge control single-phase rectifier to convert AC to DC, and a 10A, 300W DC-DC buck converter to step down the voltage to 12–15 V for charging the battery. Sensors, including an LDR sensor to detect sunlight intensity, a DHT11 temperature and humidity sensor, two voltage sensors, and two current sensors are used to collect live data for load forecasting and switching purposes [4]. The data is collected by a NODEMCU ESP8266 and sent to Thingspeak to plot graphs as shown in Fig. 2.

A switching circuit, Fig. 3 designed using eight 12 V, 10A relays, is used to charge the battery and supply the load based on conditions. The switching cir cuit is controlled by an Arduino UNO R3 SMD atmega328P board and a switch is used to turn the load on and off. To operate a small-scale DC microgrid effectively, the hardware must consist of various components that work collaboratively to gather data, regulate voltage, and conduct load forecasting [5]. These components should be designed to be highly efficient and dependable, guaran teeing that the microgrid can function as intended (Fig. 4).

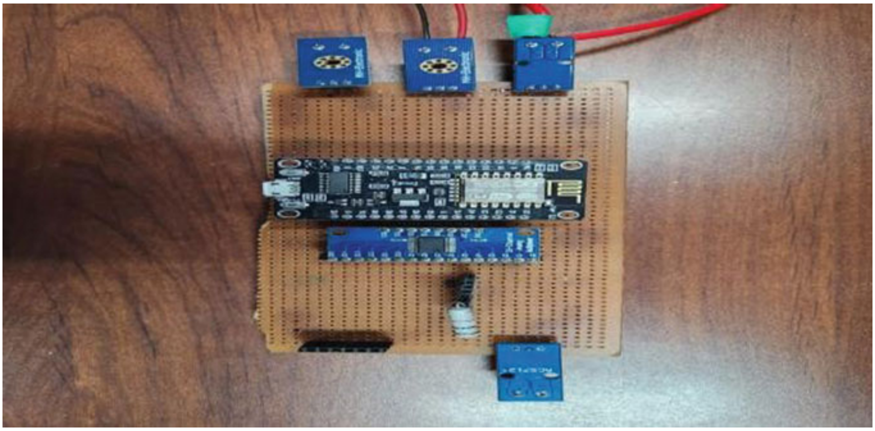


Fig. 2 Circuit to send data to cloud

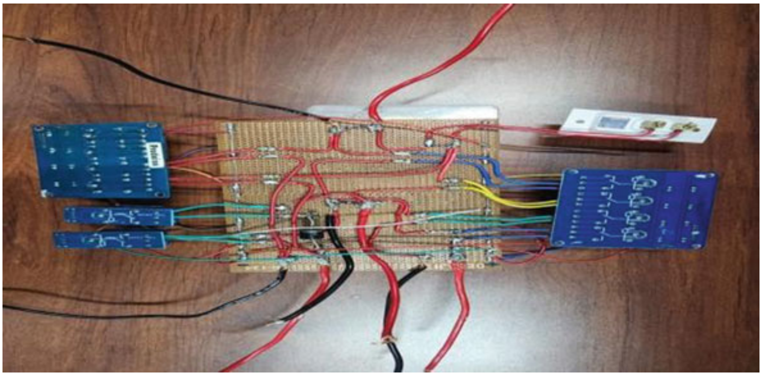


Fig. 3 Switching circuit

Fig. 4 Solar panel

3 Simulations

The simulations were carried out using Matlab Simulink which is a popular simulation tool used in engineering and science. It is frequently utilized for simulating DC microgrids due to its user-friendly interface and robust simulation capabilities.

Figure 5 represents the charging of a 12 V 150Ah battery from a PV array, the solar panels have been given a constant irradiance of 1000 W/m^2 and a temperature 25 °C. Followed by a buck converter which steps down the voltage from PV(Vmp = 29 V) to 12 V.

Figure 6 depicts how the battery is getting charged from solar panel. If the solar irradiance is more than certain value (here threshold irradiance value was taken as 300 W/m^2) and when battery voltage is less than 11 V then battery will get charged from the solar panel. If load is switched at this point, then load will get supply from the grid. Figure 6 shows the simulation output of the circuit represented in Fig. 5, where the voltage has been stepped down to 12 V, the soc was initialized as 45 percent which is increasing.

Figure 7 represents the battery being charged from the grid, the 230 V AC has been stepped down through an isolation transformer to 127 V AC which is rectified to 35 V DC; a buck converter is used to step down the voltage to 12 V charging the battery [7]. Figure 8 shows the simulation output of the circuit being represented in Fig. 7.

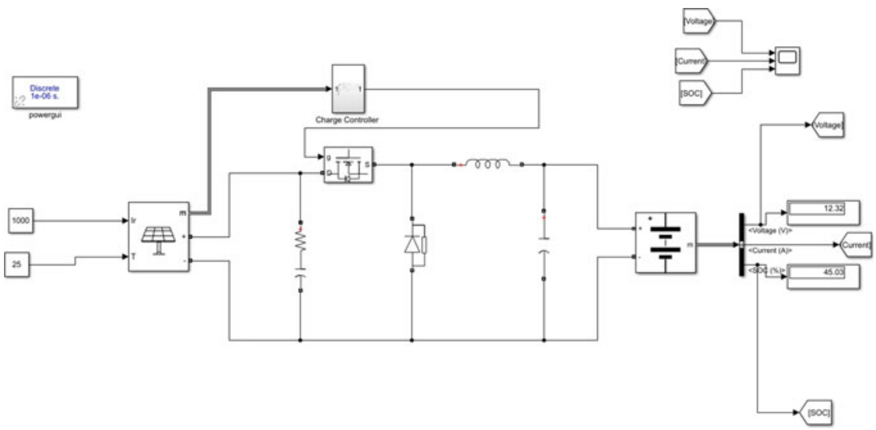


Fig. 5 Battery charging PV

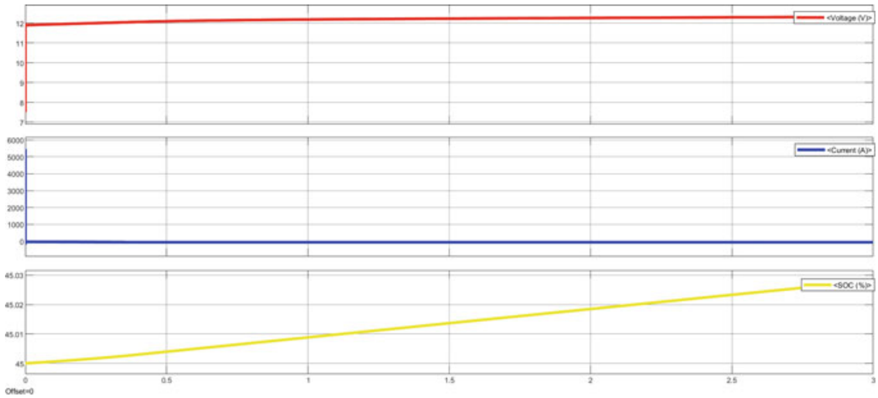


Fig. 6 Battery charging PV output

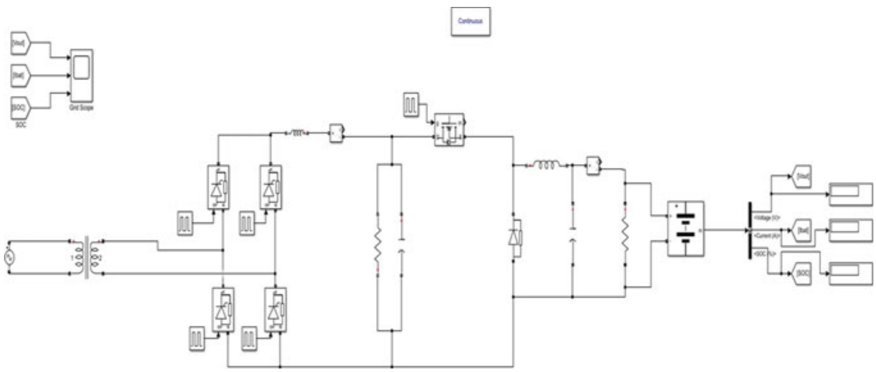


Fig. 7 Battery charging from grid

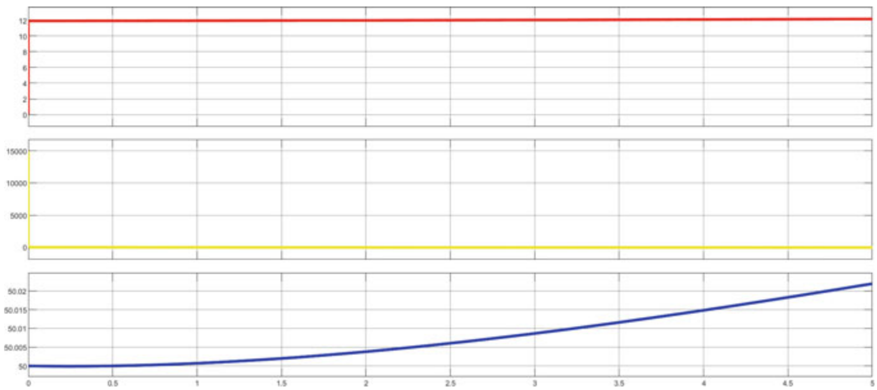


Fig. 8 Battery charging grid output

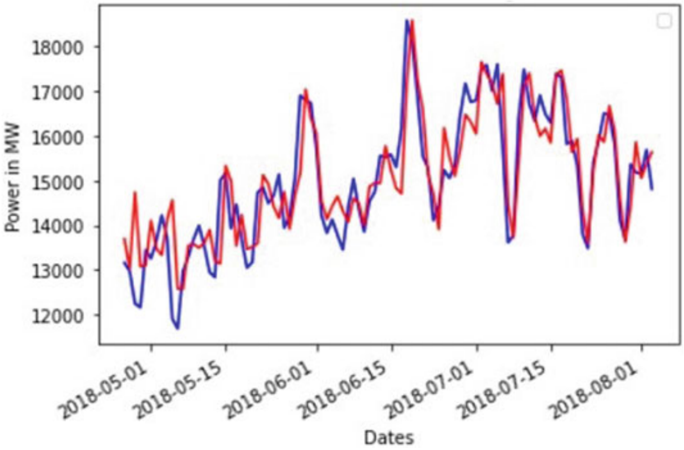


Fig. 9 Actual power versus predicted power

4 Result

Figure 9 contrasts actual and expected power consumption based on our model, with actual consumption shown in green and predicted consumption shown in red. We can see from the figure that the anticipated numbers are almost always correct, with an average prediction accuracy of nearly 95% [8]. This shows how accurate our model is in forecasting power consumption (Table 1).

These findings are noteworthy because they can be utilized to optimize energy usage, lower expenses, and increase energy efficiency. Making plans in accordance with anticipated energy demand can also be helpful. As a result, our approach may be applied to a variety of systems, including energy management, renewable energy, and smart home systems. In conclusion, our model has shown promise in properly projecting power usage, which can have a big impact on energy management and efficiency.

5 Conclusion and Future Scope

In conclusion, the development of accurate solar energy forecasting models is critical for the successful integration of renewable energy sources like solar power in DC microgrids. Recently, the use of LSTM-based approaches has emerged as a promising solution for improving the accuracy of solar energy forecasting. LSTM networks have demonstrated the ability to model complex nonlinear relationships in time-series data, resulting in accuracy rates of predicted output around 95 percent. This high level of accuracy makes LSTM-based forecasting models an excellent option for managing renewable energy sources and load demand fluctuations in DC microgrids.

Table 1 Prediction results

Date (yyyy–mm–dd)	True power (MW)	Predicted power (MW)	Accurac y (%)
2018–04–27	13,157.79	13,671.71	96.2
2018–04–28	12,964.00	12,991.95	99.7
2018–04–29	12,156.79	13,211.94	92
2018–04–30	13,443.50	12,788.46	95.1
2018–05–01	13,251.88	13,789.05	96.1
2018–05–02	13,641.17	12,804.15	93.8
2018–05–03	11,190.17	12,709.70	96.2
2018–05–04	11,680.08	14,472.20	93.3
2018–05–05	12,972.50	12,677.79	92.1
2018–05–06	13,295.08	12,127.53	96.9
2018–05–07	13,688.75	12,887.20	96.4
2018–05–08	13,993.25	12,743.55	93.1
2018–05–09	13,525.17	12,747.04	93.3
2018–05–10	12,942.92	13,814.03	94.3
2018–05–11	12,832.54	13,970.20	96.4
2018–05–12	15,171.79	15,169.07	99.8
2018–05–13	13,925.42	14,419.25	96.6

As the use of DC microgrids continues to grow, further research on the integration of LSTM-based forecasting models will be crucial to ensure their efficient deployment and operation. Future work related to solar energy forecasting using LSTM-based approaches could include investigating the impact of input data on the accuracy of the forecasting model. The performance of LSTM-based models may be affected by the selection of input features, such as weather data or historical solar energy generation data. Therefore, researchers may want to explore the effectiveness of different types of input data and identify the most relevant features for accurate forecasting. Another potential area for future work is the exploration of hybrid models that combine LSTM-based forecasting with other machine learning techniques, such as support vector machines or decision trees. These hybrid models could potentially improve the accuracy of solar energy forecasting by leveraging the strengths of multiple machine learning approaches. Overall, as the use of DC microgrids continues to grow, further research on the integration of LSTM-based forecasting models will be crucial to ensure their efficient deployment and operation. By addressing these research gaps, we can continue to improve the accuracy of solar energy forecasting and enable the successful integration of renewable energy sources like solar power in DC microgrids.

References

1. Naik K, Venugopal N (2018) 1kW home inverter using cascaded current fed push pull converter and SPWM inverter. 1–5. https://doi.org/10.1109/ICCCNT.2018.8494158
2. Mamatha G (2015) Perturb and observe MPPT alogrithm implementation for PV applications. Int J Comput Sci Inform Technol. https://doi.org/10.1109/ICCCNT.2018.8494158
3. Shakir M, Biletskiy Y (2020) Forecasting and optimization for microgrid in home energy management systems. IET Gener Transm Distrib 14. https://doi.org/10.1049/iet-gtd.2019.1285
4. Dutta S, Li Y, Venkataraman A, Costa LM, Jiang T, Plana R, Tordjman P, Choo FH, Foo CF, Puttgen HB (2017) Load and renewable energy forecasting for a microgrid using persistence technique. https://doi.org/10.1016/j.egypro.2017.12.736
5. Wang Y, Feng B, Hua Q-S, Sun L (2021) Short-term solar power forecasting: a combined long short-term memory and gaussian process regression method. Sustainability 13:3665. https://doi.org/10.3390/su13073665
6. Joseph SC, Chandrasekar V, Dhanesh PR, Mohammed AA, Ashok S (2018) Battery management system for DC nanogrid. In: 2018 20th national power systems conference (NPSC), Tiruchirappalli, India, 2018, pp 1–5. https://doi.org/10.1109/NPSC.2018.8771838
7. Koskela J, Mutanen A, Järventausta P (2020) Using load forecasting to control domestic battery energy storage systems. Energies 13:3946. https://doi.org/10.3390/en13153946
8. Alavi SA, Mehran K, Vahidinasab V, Catalao JPS (2021) Forecast- ˜ based consensus control for DC microgrids using distributed long short-term memory deep learning models. IEEE Trans Smart Grid 12(5):3718–3730. https://doi.org/10.1109/TSG.2021.3070959
9. Prasanna Kumar C, Venugopal N (2018) Performance and stability analysis of series-cascaded, high-gain, interleaved boost converter for photovoltaic applications. Power Electron Drives. https://doi.org/10.2478/pead-2018-0022
10. Naik KK, Venugopal N (2018) Multi output forward-flyback switch mode power supply. In: 2018 3rd IEEE international conference on recent trends in electronics, information communication technology (RTEICT). https://doi.org/10.1109/RTEICT42901.2018.9012304
11. Roopashree DV, Venugopal N (2021) A low-voltage ride-through strategy using fuzzy based controller for 3 phase grid connected PV system. In: 2021 International conference on recent trends on electronics, information, communication and technology (RTEICT), Bangalore, India, 2021, pp 34–37. https://doi.org/10.1109/RTEICT52294.2021.9573643
12. Juha K, Mutanen A, Järventausta P (2020) Using load forecasting to control domestic battery energy storage systems. Energies 13(15):3946. https://doi.org/10.3390/en13153946
13. Chidananda C, Venugopal N, Vishwanath MN (2022) Asymmetric multi-level inverter. In: 2022 3rd International conference for emerging technology (INCET), Belgaum, India, 2022, pp 1–7. https://doi.org/10.1109/INCET54531.2022.9824659
14. El-Shahat A, Sumaiya S (2019) DC-microgrid system design, control, and analysis. Electronics 8:124. https://doi.org/10.3390/electronics8020124

LoRa-Based Farm Monitoring System

Mustafa Poonawala, Atharva Sonawane, Mokshit Lodha, Vinaya Gohokar, Anjali Askhedkar, and Shruti Danve

Abstract Internet of Things (IoT) serves as a great solution to the rising demands and challenges that are being faced by the agricultural sector. It can help in data collection using smart sensors. Control systems, automated equipment, sensors, motion detectors, robotics, wearable technology, autonomous vehicles, and button cameras are crucial elements in this type of farm management. This information can be utilized to monitor soil quality, irrigation, pest control, surveillance, and overall health of the crops in the farm. The ability to estimate production yield enables improved product distribution planning. The main goal of this study is to use LoRa technology to establish communication between sensors and a gateway for the purpose of farm monitoring. An example of a farmland monitoring device will be observed wherein, an ultrasonic sensor will be used to detect stray animals or rodents that attempt to enter agricultural fields and relay this information to the gateway using LoRa communication.

Keywords LoRa · Smart agriculture · Intrusion detection · IoT · Pest control · Communication systems

---

Mustafa Poonawala, Atharva Sonawane, and Moskhit Lodha contributed equally to this work as first authors.

M. Poonawala · A. Sonawane (✉) · M. Lodha · V. Gohokar · A. Askhedkar · S. Danve
Dr. Vishwanath Karad, MIT World Peace University, Pune 411038, India
e-mail: atharvasonawane39@gmail.com

M. Poonawala
e-mail: mpoonawala1808@gmail.com

M. Lodha
e-mail: mokshitlodha.9@gmail.com

V. Gohokar
e-mail: vinaya.gohokar@mitwpu.edu.in

A. Askhedkar
e-mail: anjali.askhedkar@mitwpu.edu.in

S. Danve
e-mail: shruti.danve@mitwpu.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_25

1 Introduction

The need to boost farm productivity has become critical as a result of issues such that the exponential rise of the world's population, would need the globe to produce 70% more food by 2050, shrinking agricultural areas, and the depletion of finite natural resources. The issue has further been aggravated by a lack of natural resources such as fresh water and fertile land, as well as declining production levels of certain critical crops. Another impediment is the changing organizational structure of the agricultural workforce. Furthermore, agricultural labor has declined in the majority of countries. Physical labor demand has decreased as a result of the shrinking agricultural workforce, prompting the incorporation of Internet connectivity solutions into farming techniques. The usage of fertilizer, the number of trips made by farm vehicles, and the effective use of resources like water, energy, and other resources are just a few examples of how smart farming based on IoT helps growers and farmers cut waste and enhance production. IoT smart farming systems use sensors to monitor and automate irrigation in agricultural fields (light, humidity, temperature, crop health, soil moisture, etc.). Farmers can check on the state of their fields from any location. Based on this information, they can decide between manual and automatic methods for taking the proper actions. In this paper, we will be designing a system to ensure that a farm remains secure from unwanted pests and rodents.

Cellular networks are a popular choice for connectivity because they provide broad coverage for high-speed transmission. However, high-speed data is not always required for IoT applications, and cellular-connected devices frequently experience battery life issues and coverage gaps. Mesh networks of the ZigBee type are commonly used in home automation. This is due to the low to medium distance performance it offers. However, when it comes to long distances, these fail miserably (a few kilometers). Although Bluetooth/Bluetooth low energy (BLE) provides reasonable data speeds, as seen in [1–4], its range is a major drawback. The range provided is simply insufficient for long-distance IoT applications. Devices with Wi-Fi capabilities have a limited lifespan and rely on close proximity to the Wi-Fi access point to function properly. This is because the frequency of the wireless signals used by Wi-Fi technology can be obstructed by physical barriers, limiting their ability to easily penetrate through obstacles. The primary objective of the system designed in this study paper is to protect agricultural fields from rodents, pests, and other intruders. For this system, the use of LoRa technology is most beneficial due its low cost and low power consumption. Moreover, it can support communication over a long range which makes it ideal for use in larger agricultural fields.

The major contribution in this paper is about the range achieved successfully between the LoRa node and LoRa gateway using the SX1276 module and the rodent detection system created using an ultrasonic sensor.

The paper is arranged as below. Sections 2 and 3 include the related works and the LoRa technology in detail. The implementation details and experimental results are a part of Sects. 4, 5 and 6 concludes the paper.

2 Related Works

Literature studies show the prevalent use of technology for monitoring and boosting agricultural produce. The system designed in [5] utilizes a Heltec ESP32 module, which can communicate through Bluetooth, Wi-Fi, and LoRa. The Semtech SX1276 LoRa module is considered for the proposed system model, since this farm monitoring system requires a LoRa module capable of communicating in the 865–867 MHz band of frequencies. The authors of [7] and [8] presented an IoT-based agricultural device that is a portable handheld device to determine the state of the environment and soil. These gadgets include an EC sensor, a pH sensor, a temperature sensor, a color sensor, and humidity sensors, and GPS and ZigBee modules for radio transmission. The authors of [11] create portable nodes using SX1278 based on STM32 microcontroller. Research indicates that in a densely populated metropolitan environment, a maximum distance of 900 m can be traversed. In this paper, the work done demonstrates a coverage over 1 km range in urban environment using Arduino and SX1276, however [11] makes use of SX1278. The suggested technique in [12] uses a LoRaWAN-based architecture to collect, store, and assess a variety of metrics relevant to the health of each head of the herd and the habitat where cattle are kept. These metrics are characteristic of the bovine livestock scenario. This application keeps a check on intruders which could be an added security for the cattle. [13, 14] make use of LoRa technology to communicate data between sensors placed in the field and a central controller. The sensors gather information on soil moisture levels, temperature, and other relevant factors, and send this data to the central controller. The central controller then processes this information and provides relevant information and recommendations to farmers, such as when to water or fertilize their crops.

3 LoRa Technology

Semtech developed LoRa technology, a new wireless protocol designed specifically for long range, and low-power communications. Long range radio, or LoRa, is primarily intended for machine to machine (M2M) communication and IoT networks. This technology allows multiple apps running on the same network to connect via public or multi-tenant networks. The LoRa Alliance has standardized low-power wide area networks (LPWAN) for the IoT. This collaboration is critical for enabling interoperability among various national networks. A single LoRa gateway can manage up to millions of nodes. Building a network is quicker and less expensive since signals may traverse large distances with less infrastructure. Moreover, LoRa includes an adjustable data rate algorithm that increases network capacity and node battery life. The LoRa protocol has several layers for secure communication, including encryption at the network, application, and device levels.

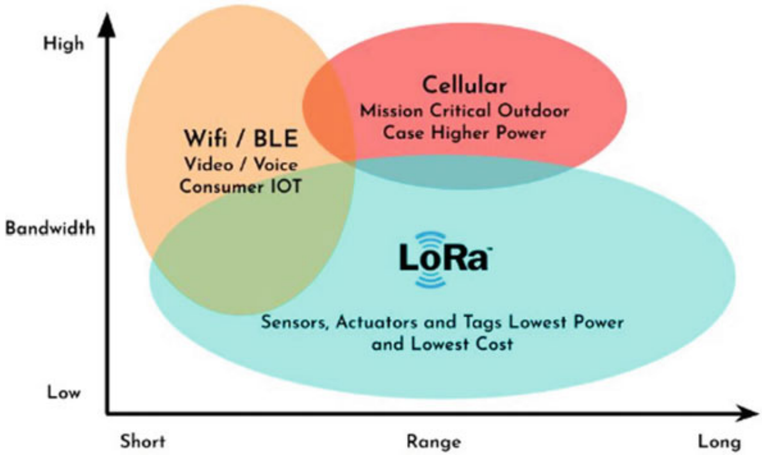


Fig. 1 Comparison of wireless communication technologies [15]

Figure 1 shows the range and bandwidth comparison of various existing communication technologies such as Wi-Fi, BLE, cellular, and LoRa technology.

4 Implementation Details

In this research paper, the system is designed to monitor the perimeter of a field in real time with the help of an ultrasonic sensor for detection of rodent intrusion. Once the perimeter is breached, the sensor will raise a flag by sending data to the gateway via LoRa communication.

The data from the ultrasonic sensor is sent to the LoRa module node which is then sent to the LoRa gateway, the gateway then sends it to the cloud database through the Wi-Fi module, from which the data can be received and read. The communication established here is bidirectional in nature, i.e., both the nodes as well as the gateway can send as well as receive information. The interfacing is achieved by the help of Arduino UNO board.

In Fig. 2, the LoRa gateway is connected to a Wi-Fi module that enables it to connect to a cloud service, which is considered as a future implementation.

The SX1276 module's sensitivity of 148 dBm, dynamic range of 127 dB, frequency range of 867 MHz, constant radio frequency output of + 20 dBm, bit rate of 300 kBits/sec, maximum link budget of 168 dB, and packet engine of 256 bytes have all been taken into consideration [16].

The LoRa long range modem, which has been integrated into the SX1276 transceivers, offers ultra-long range spread spectrum communication and strong interference immunity while consuming the least amount of current. The SX1276/77/78/79 can reach a sensitivity of over −148dBm using a low-cost crystal and bill

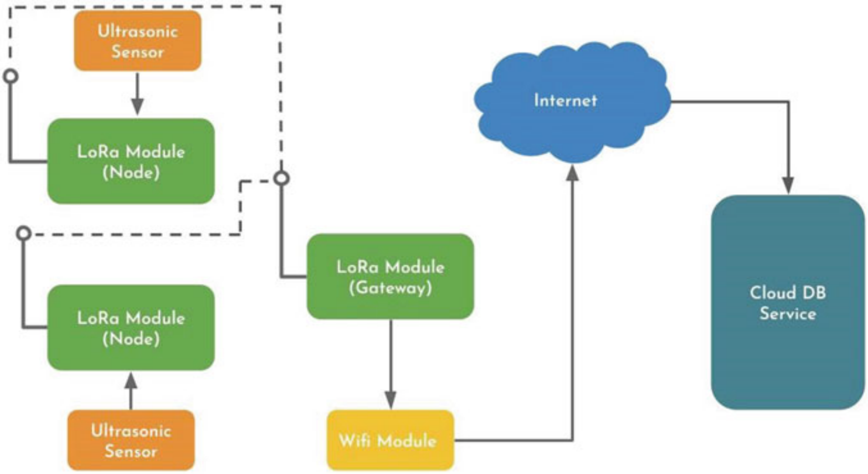


Fig. 2 Block diagram of LoRa-based farm monitoring system

of materials by employing Semtech's unique LoRa modulation technology. It is ideal for any application needing range or resilience because of its high sensitivity and integrated + 20 dBm power amplifier, which together produce an industry leading link budget. To integrate the SX1276 LoRa module with the Arduino UNO, a breakout board is designed over which the module can be mounted. As displayed in the PCB layout in Fig. 3, we have an SX1276 LoRa module which is connected to an SMA connector so as to attach the antenna. The different pins are connected to pin headers so that it could be mounted on a breadboard. The dimensions of the PCB layout as seen above is 48.5 $\times$ 41.1 mm. The pin headers are kept at a distance such that the PCB could be mounted on a breadboard.

A rubber duck antenna with a frequency of 865–868 MHz and 1 dBi gain is used to extend the range of wireless network coverage in a single direction. It gives an approximate range of 1000 m (1 km) (Fig. 4).

5 Results

In this experimental setup, intrusion detection has been tested along with the corresponding range of communication and its corresponding signal strength.

In Fig. 5, the LoRa node that is attached to the fence or boundary will continuously monitor whether there are any intruders approaching the fence.

Figure 6 shows the LoRa gateway messages that are received from the node. These readings give the distance of the intruder from the fence along with received signal strength indication (RSSI).

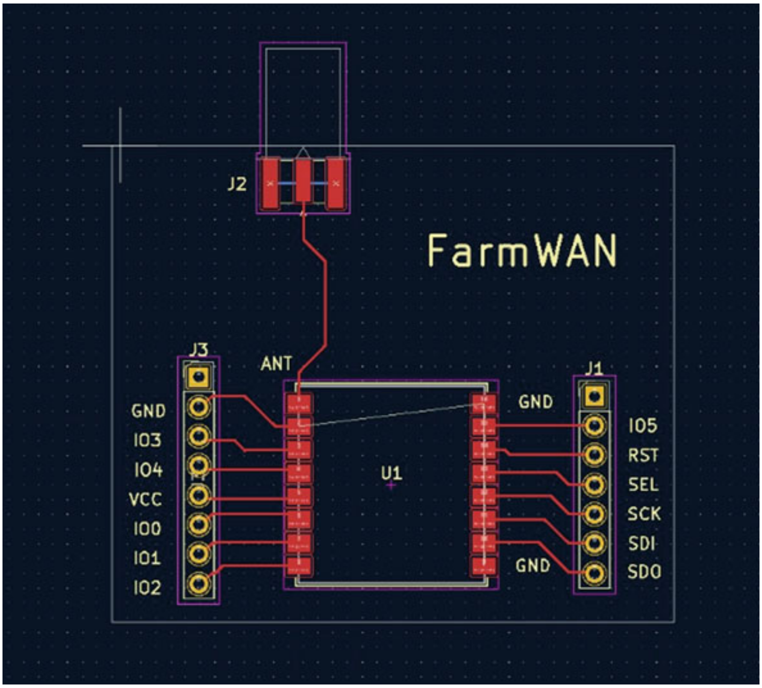


Fig. 3 PCB design of LoRa module

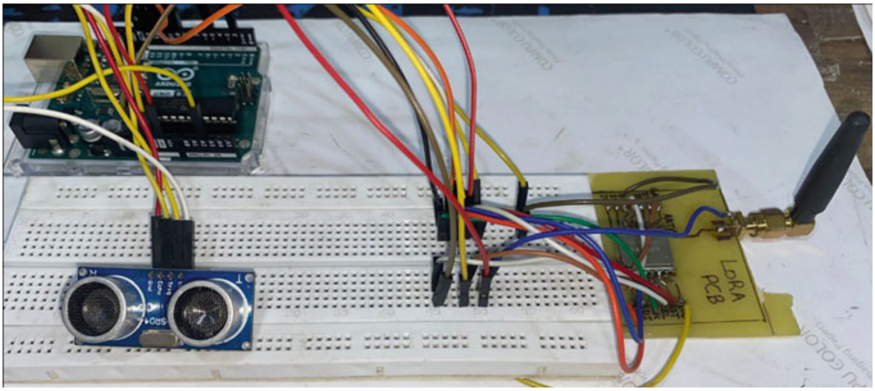
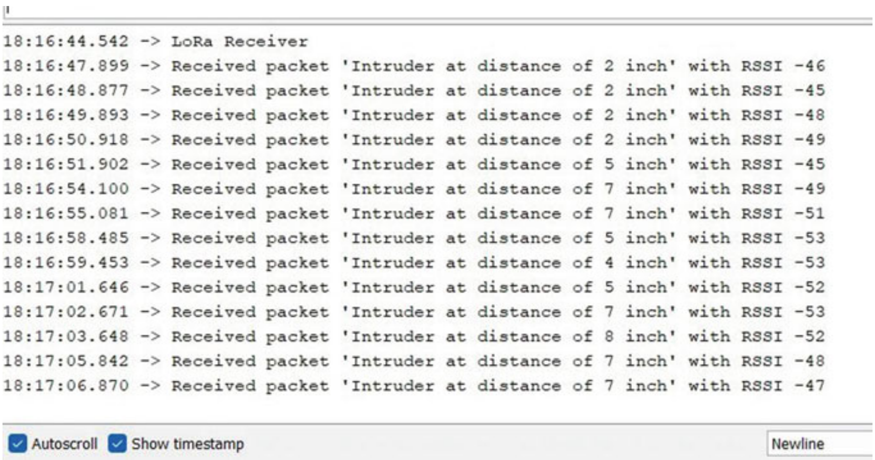


Fig. 4 Prototype of LoRa module

As seen in Table 1, the system has been tested over a range of 100 m–1km. The LoRa node module and LoRa gateway are first kept at a distance of 100m from each other to test the message quality. The message is received without any garbage values. These values are tested in a heavily populated area and not in an open field. The distance is gradually increased by 100 m and the message quality is tested at every

```
18:16:42.951 -> LoRa init succeeded.
18:16:42.991 ->
18:16:42.991 -> LoRa Simple Node
18:16:42.991 -> Only receive messages from gateways
18:16:43.069 -> Tx: invertIQ disable
18:16:43.069 -> Rx: invertIQ enable
18:16:43.110 ->
18:16:44.119 -> Distance: 1222 Inch
18:16:45.290 -> Distance: 1215 Inch
18:16:46.293 -> Distance: 2 Inch
18:16:47.301 -> Send Message!
18:16:47.341 -> Distance: 2 Inch
18:16:47.382 -> TxDone
18:16:48.304 -> Send Message!
18:16:48.344 -> Distance: 2 Inch
18:16:48.384 -> TxDone
18:16:49.312 -> Send Message!
18:16:49.352 -> Distance: 2 Inch
18:16:49.393 -> TxDone
18:16:50.322 -> Send Message!
```

Fig. 5 LoRa node transmitted data



```
18:16:44.542 -> LoRa Receiver
18:16:47.899 -> Received packet 'Intruder at distance of 2 inch' with RSSI -46
18:16:48.877 -> Received packet 'Intruder at distance of 2 inch' with RSSI -45
18:16:49.893 -> Received packet 'Intruder at distance of 2 inch' with RSSI -48
18:16:50.918 -> Received packet 'Intruder at distance of 2 inch' with RSSI -49
18:16:51.902 -> Received packet 'Intruder at distance of 5 inch' with RSSI -45
18:16:54.100 -> Received packet 'Intruder at distance of 7 inch' with RSSI -49
18:16:55.081 -> Received packet 'Intruder at distance of 7 inch' with RSSI -51
18:16:58.485 -> Received packet 'Intruder at distance of 5 inch' with RSSI -53
18:16:59.453 -> Received packet 'Intruder at distance of 4 inch' with RSSI -53
18:17:01.646 -> Received packet 'Intruder at distance of 5 inch' with RSSI -52
18:17:02.671 -> Received packet 'Intruder at distance of 7 inch' with RSSI -53
18:17:03.648 -> Received packet 'Intruder at distance of 8 inch' with RSSI -52
18:17:05.842 -> Received packet 'Intruder at distance of 7 inch' with RSSI -48
18:17:06.870 -> Received packet 'Intruder at distance of 7 inch' with RSSI -47
```


Fig. 6 LoRa gateway received data

Table 1 Range and detection results of LoRa-based farm monitoring system

S. No.	Distance	Range	Message status
1	Distance <= 24 Inch	Range <= 500 m	Received
2	Distance <= 24 Inch	500 m < Range <= 1 km	Received with some garbage values
3	Distance <= 24 Inch	Range > 1 km	Garbage values
4	Distance >= 24 Inch	–	Message not sent

instance. Till the 500 m mark, the message quality is the best and after 500 m, there are some garbage values transmitted. Since it is tested in a heavily populated area and the ideal setting for such a system would be an open field, these few values are acceptable. After the 1 km mark, there are no values being received by the gateway. The program includes a condition that would make sure that a message will only be sent if the distance recorded by the ultrasonic sensor is less than a range of 24 inches. The data could also be uploaded to cloud services as suggested in [1, 6].

6 Conclusion

Recent government policies promote farmers to adopt various technologies so as to increase their production yield, this in turn is causing a paradigm shift in agriculture. In this research paper, a farmland monitoring system is proposed and studied wherein, an ultrasonic sensor is used to detect stray animals or rodents that attempt to enter agricultural fields, and this information is then relayed to the gateway using LoRa communication. Internode communication is disabled so as to reduce any interference. The received data could then be sent to the cloud and stored in a database as mentioned earlier. Hardware used is LoRa SX1276 modules, Arduino Uno, and ultrasonic sensors and the maximum range of communication achieved is observed to be 1 km.

A future prospect is to connect to a cloud database and collect data on the occurrence of rodents and finding the problem areas by increasing the number of nodes in the network so as to cover wider areas. As another future prospect, the proposed system could also be integrated with security cameras to increase surveillance. The use of an ultrasonic sensor in this device serves as an economic option and also helps detect intruders that cannot be detected by cameras such as insects and smaller rodents. The use of a PIR sensor could also be considered to detect movement.

Acknowledgements This work has been supported and partially funded by Mayank Gandhi, Kanan Park Technologies, Pune.

References

1. Dharshan Y, Sharmila B, Srinivasan K, Suresh M, Panchal H, Meenakumari R, Ashokkumar R, Srimali N, Israr M, Sadasivuni KK (2021) An improved optimisation technique for the network-controlled pH process and DC motor using various controllers. Int J Ambient Energy
2. Kaushik S, Srinivasan K, Sharmila B, Devasena D, Suresh M, Panchal H, Ashokkumar R, Sadasivuni KK, Srimali N (2021) Continuous monitoring of power consumption in urban buildings based on internet of things
3. Senthil Kumar R, Gerald Christopher Raj I, Suresh KP, Leninpugalhanthi P, Suresh M, Panchal H, Meenakumari R, Sadasivuni KK (2021) A method for broken bar fault diagnosis in three phase induction motor drive system using artificial neural networks
4. Sharmila B, Srinivasan K, Devasena D, Suresh M, Panchal H, Ashokkumar R, Meenakumari R, Sadasivuni KK, Shah RR (2021) Modelling and performance analysis of electric vehicle. Int J Ambient Energy
5. Kodali RK, Yerroju S, Sahu S (2018) Smart farm monitoring using LoRa enabled IoT. In: Second international conference on green computing and internet of things (ICGCIoT), IEEE, pp 391–394
6. Suji Prasad SJ, Thangatamilan M, Suresh M, Panchal H, Rajan CA, Sagana C, Gunapriya B, Sharma A, Panchal T, Sadasivuni KK (2021) An efficient LoRa-based smart agriculture management and monitoring system using wireless sensor networks. Int J Ambient Energy
7. Maia RF, Netto I, Tran ALH (2017) Precision agriculture using remote monitoring systems in brazil. In: IEEE Global humanitarian technology conference (GHTC), IEEE, pp 1–6
8. Sharma P, Padole DV (2017) Design and implementation soil analyser using Iot. In: International conference on innovations in information, embedded and communication systems (ICIIECS), pp 1–5
9. Davcev D, Mitreski K, Trajkovic S, Nikolovski V, Kotel N (2018) IoT agriculture system based on LoRaWAN. In: 14th IEEE International workshop on factory communication systems (WFCS). IEEE, Italy
10. Tzounis A, Katsoulas N, Bartzanas T, Kittas F (2017) Internet of things in agriculture, recent advances and future challenges. Biosyst Eng 164:31–48. Elsevier Ltd.
11. Li Q, Liu Z, Xiao JF (2018) A data collection collar for vital signs of cows on the grassland based on LoRa. In: IEEE 15th International conference on e-business engineering (ICEBE), IEEE, Xi'an, China, pp 213–217
12. Germani L, Mecarelli V, Baruffa G, Rugini L, Frescura F (2019)An IoT architecture for continuous livestock monitoring using LoRa LPWAN. In: Electronics 8(12):1435. https://doi.org/10.3390/electronics8121435
13. Radosveta Sokullu F (2022) LoRa based smart agriculture network. In: 8th International conference on energy efficiency and agricultural engineering (EE&AE), (2022)
14. Mya KT, Sein MM, Nyunt TTS, Lewlompaisarl U, Yasunori Owada F (2020) A design for IoT based smart watering system using LoRa. In: IEEE 9th global conference on consumer electronics (GCCE). IEEE
15. Ovyl. https://ovyl.io/insights-and-news/wireless-technologies#:~:text=BLE%20has%20the%20lowest%20power,the%20cellular%20service%20turned%20off
16. Semtech. https://www.semtech.com/products/wireless-rf/lora-connect/sx1276

Breast Cancer Detection Using Image Processing

Saraswati Patil, Sangeeta Jaybhaye, Shravani Bahulekar, Siddhant Thaware, Kaivalya Shende, and Sahil Shitole

Abstract Breast cancer remains a significant health concern for women worldwide, and early detection is critical for better treatment outcomes Since there are so many mammograms conducted each day, it is still very difficult to automatically detect and classify the masses, which is necessary for radiologists to properly identify masses in mammograms. In this study, we present a unique You Only Look Once (YOLO) version 7-based computer-aided diagnosis (CAD) system to assist radiologists in precisely detecting and localizing lesions in full and cropped pictures. YOLO is a popular object detection system, and the most recent version, YOLOv7, is capable of detecting objects in high-resolution images with a high frame rate. Mammography, a specialized imaging method used for breast cancer screening and diagnosis, might considerably benefit from automated CAD systems like the one suggested in this work. The creation of such systems has produced encouraging effects in terms of increasing the effectiveness and precision of breast cancer diagnosis, perhaps lightening the burden of radiologists and enhancing patient outcomes. A promising area with tremendous potential for development is the application of AI-based systems for medical imaging analysis.

Keywords Object detection · YOLOv7 · CNN · mAP · Precision

S. Patil (✉) · S. Jaybhaye · S. Bahulekar · S. Thaware · K. Shende · S. Shitole
Vishwakarma Institute of Technology, Pune, Maharashtra, India
e-mail: saraswati.jadhav@vit.edu

S. Jaybhaye
e-mail: sangita.jaybhaye@vit.edu

S. Bahulekar
e-mail: shravani.bahulekar21@vit.edu

S. Thaware
e-mail: kundlik.siddhant21@vit.edu

K. Shende
e-mail: kaivalya.shende21@vit.edu

S. Shitole
e-mail: sahil.shitole21@vit.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_26

1 Introduction

Breast cancer is the disease with the highest mortality rate among women in emerging and developing countries. If breast cancer is diagnosed and classified early, patients may be able to receive appropriate treatment. Breast cancer is the most common cancer in women, affecting 2.1 million people annually and responsible for the majority of cancer-related deaths in women. The number of newly diagnosed women worldwide is estimated to be 2,261,419 in 2020. According to estimates, 178,000 new cases of breast cancer are detected each year among Indian women. There are many risk factors for breast cancer including age, family history, and medical history. Only 5 to 10% of cases of breast cancer have a significant family history (genetic predisposition). Genetic testing is only appropriate for high-risk groups (those with an extended family history of breast cancer) and when genetic counseling is available.

There is a lot of interest in creating deep learning tools to help radiologists and increase the precision of screening mammography given the spectacular performance of deep learning in visual object recognition and detection, as well as many other domains. Breast cancer can be found early on thanks to screening, which makes therapy more successful. Convolutional neural networks are used in this image identification program. It's a particular kind of artificial neural network that has been developed and applied to successfully recognize visual pictures. Mammography pictures may be classified, and anomalies can be found using this approach. It is challenging to detect breast abnormalities because of their shape and traits. In order to help medical personnel analyze medical pictures more quickly and accurately, CAD systems have been created to help detect breast cancers.

This model utilizes the Google TensorFlow open-source machine learning library. Data from the mini-mammographic image analysis society (mixed) database was used to evaluate the classifier. The input picture pixel size, kernel size, learning rate, step number, number of hidden layers, and other factors were changed for this classifier. The goal is to automatically identify areas of interest (ROIs) in mammograms that could contain malignant tumors using YOLO v7. The presence of malignancy can then be ascertained by radiologists after reviewing these ROIs. One advantage of utilizing YOLO v7 for mammography breast cancer screening is that it may lessen the strain of radiologists, who may otherwise get overburdened by the sheer volume of mammograms they need to evaluate. The screening procedure may be more sensitive, thanks to YOLO v7's potential to identify tumors that are invisible to the naked eye.

2 Literature Survey

Hamed et al. [1] propose a highly accurate CAD system based on YOLO v4 for the detection and classification of breast cancer. The method employs a two-way technique, with one route used to detect lesions in cropped slices and the other for full mammography detection. Resizing mammograms is no longer a problem according to the suggested technology, which also achieves an overall detection accuracy of 98.3% and can accurately discriminate between benign and malignant tumors with 95% accuracy. The system can outperform current CADs since it is a real-time detector.

Das et al. [2]: In this paper, the YOLO v3 network is proposed for use as a region proposal network (RPN) in breast fine needle aspiration cytology images to identify regions of interest (RoI) patches. The diagnosis of breast ductal carcinoma is based on the classification of patches using convolutional neural networks (CNNs). According to the YOLO + CNN model, the classification accuracy is 95.73%, the sensitivity is 0.924, and the specificity is 1. These results suggest good applications in industrial settings.

Mohiuddin et al. [3]: According to this study, a modified YOLO v5 model can be used to detect and classify breast tumors based on mammograms. The model surpasses YOLOv3 and RCNN with 96% mAP, 93.50% MCC value, 96.50% accuracy, 0.04 FPR, and 0.03 FNR value. By decreasing the FPR and FNR and increasing the MCC value, it also overcomes the limitations of previous research.

Al-masni et al. [4]: This study suggests a novel You Only Look Once (YOLO)-based convolutional neural network (CNN)-based computer-aided diagnostic (CAD) system for mammography. (YOLO). Mammogram preprocessing, feature extraction, mass detection, and mass classification are the system's four steps. The YOLO-based CAD system has an overall detection accuracy of 96.33% and a distinction accuracy of 85.52% between benign and malignant tumors.

Wang et al. [5] propose automated identification of invasive ductal carcinoma (IDC) utilizing machine learning approaches which is suggested in this research. The research examines different CNN architectures for the automatically detecting breast cancer and assesses their performance using tenfold cross-validation. According to the study, one of the optimized CNN architectures produces the greatest outcomes in terms of F1, BAC, and accuracy. The result highlights the significance of an accurate model for IDC detection for breast cancer diagnosis and therapy and recommends more study on applying a multimodal deep learning technique.

Khuriwal et al. [6]: This paper proposes the use of deep learning neural network for the diagnosis of breast cancer using the Wisconsin breast cancer database. After reviewing 41 papers, the study concludes that early detection of cancer is important. The proposed model achieves an accuracy of 99.67%. The study suggests that deep learning can be used effectively for human vital data analysis.

Hazra et al. [7]: In this paper, decision trees (DT) and artificial neural networks (ANN) are used to build a machine learning (ML) model for early diagnosis and prognosis of breast cancer based on the Wisconsin diagnostic breast cancer (WDBC) dataset. After the four datasets have been preprocessed, it is split into several train-test configurations. The greatest accuracy of the suggested plan was 98.55%.

Masud et al. [8]: This paper proposes a comparison of current machine learning models for the diagnosis and categorization of breast cancer. The authors stress the value of early breast cancer identification and therapy, as well as how machine learning methods might increase diagnostic precision. Additionally, they outline the most widely used public databases for studying breast cancer and comparing them. YOLO and RetinaNet are the top-performing models for precise bulk detection and malignancy categorization.

3 Methodology

Till date, YOLO has released seven versions. Hence, the first step was to compare and identify a better model. After analyzing, Table 1 gives comparison between various YOLO models based on the architecture used and mAP values attained.

From Table 1, it is evident that YOLOv7 undoubtedly has the highest mAP score. To do more comparison, mAP score of YOLOv7 was compared with mAP score of other versions.

From Table 2. It is evident that YOLOv7 gives higher mean average precision (mAP) values when compared with other versions.

The breast cancer detection system utilizes mammograms to detect the presence of tumors in the breast. This is accomplished by training a model using a dataset of mammogram images. To train the model, a dataset of 800 images was collected and then divided into three sets: the test set, train set, and the validate set. The images were labeled with the class "tumor" using the labelImg annotation tool in YOLO format. The YOLO annotation format contains information about the object class, object coordinates, height, and width of the box made around the object to be detected. After labeling the images, two.txt files were generated. One contains the annotations for each image, and the other specifies the class of each image. The labeled images were then used to train models for the YOLOv7 algorithm. Once the weights file and cfg file were downloaded automatically, the model was successfully trained and ready for testing.

Table 1 Comparison of various versions of YOLO

YOLO version	Architecture used	mAP (%)
YOLOv2 [10]	Darknet-19	21.6
YOLOv3 [11]	Darknet-53	33.1
YOLOv4 [12]	CSPNet	43.5
YOLOv5 [13]	EfficientNet	55.0
YOLOv6 [14]	EfficientNet-L2	43.1
YOLOv7 [15]	RepConv	56.8

Table 2 mAP comparison of various YOLO versions

Version	Average precision (AP)
YOLOv4	YOLOv7 has 1.5% more AP
YOLOv5	YOLOv7 has 2.2% more AP
YOLOv6	YOLOv7 has 13.7%more AP

To test the model, the images in the test set were passed through the convolutional layers. The trained model detects the class of each image by making a bounding box around the tumor. The algorithm also provides the class probability of the bounding box. YOLO splits the input image into $N \times N$ grids of similar size and detects objects in each grid by marking a bounding box around them. This can result in multiple bounding boxes being formed for a single object. To overcome this, YOLO uses intersection over union (IOU) and non-max suppression techniques to select the boxes with the highest class probability.

$$IOU = \frac{\text{Area of intersection}}{\text{Area of Union}}. \quad (1)$$

Using a straightforward deep convolutional neural network, the YOLO method utilizes an image as its input to find objects in the picture. The CNN model, the foundation of YOLO, is illustrated in the architecture below. By adding an interim average pooling and fully connected layer, the model's first 20 convolution layers are pretrained using ImageNet. The pretrained model is then modified to do detection since prior studies have shown that a pretrained network performs better when convolution and linked layers are added. In the last fully connected layer of YOLO, bounding box coordinates and class probabilities are predicted (Fig. 1).

You only look once (YOLO) is an object detection algorithm that divides input images into a grid and predicts bounding boxes for each grid cell. For each grid cell, YOLO predicts multiple bounding boxes and assigns one predictor to be responsible for predicting the item based on the highest intersection over union (IOU) with the ground truth. The latest version of YOLO, YOLO v7, introduces several advancements to improve its accuracy and speed. It uses anchor boxes, which are preconfigured boxes with varying aspect ratios that help identify objects of various

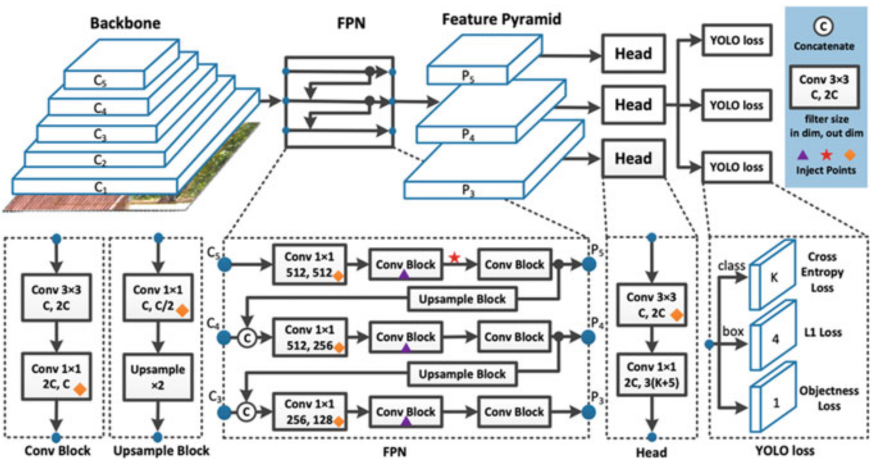


Fig. 1 System architecture [9]

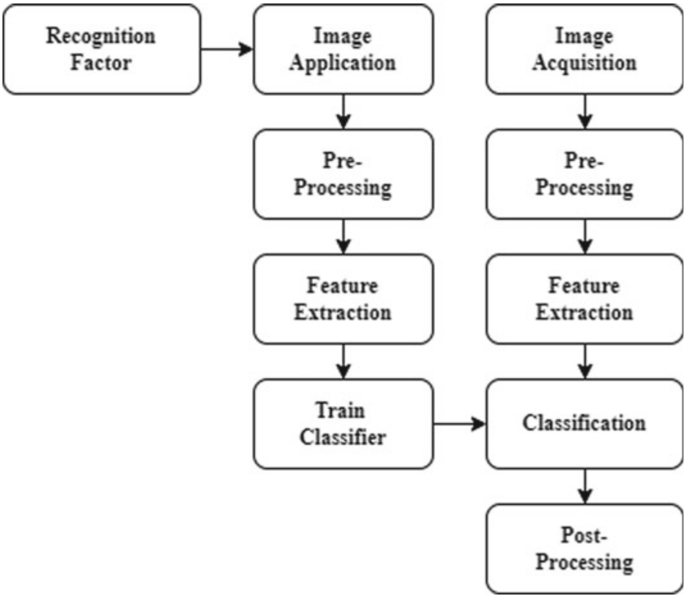


Fig. 2 Flowchart for training

types. YOLO v7 also introduces a new loss function called focal loss, which solves the issue of identifying tiny objects by minimizing the loss in well-classified cases and highlighting the hard-to-detect occurrences. With a higher resolution of 608 by 608 pixels, YOLO v7 can identify tiny objects with greater precision.

One of the key benefits of YOLO v7 is its faster processing speed, with the ability to process 155 frames per second, making it suitable for real-time applications such as surveillance and autonomous vehicles. YOLO v7 outperforms other object detection algorithms in both speed and accuracy, making it a popular choice for computer vision applications. The flowchart in Fig. 2 depicts the methodology used for making this system.

4 Result

After testing the model with images from the test set, the following results were obtained. Figure 3 shows detection of tumors on the test image of mammogram containing two tumors. The bounding box is made around the tumors with confidence values being 0.7 and 0.8. The maximum accuracy achieved was approximately 80%.

Figure 4 is a test image of a mammogram with no tumor present. The trained model accurately identifies the absence of tumor.

In order to calculate the score, the mean average precision compares the detected box with the ground-truth bounding box. The better the score, the more reliable the

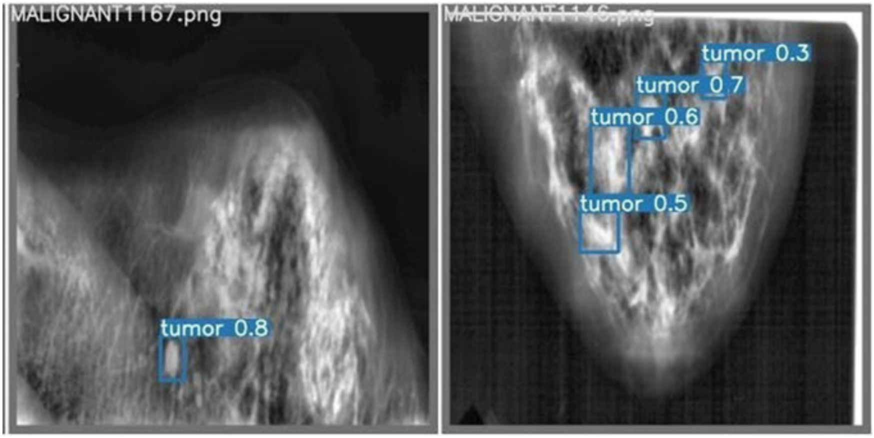
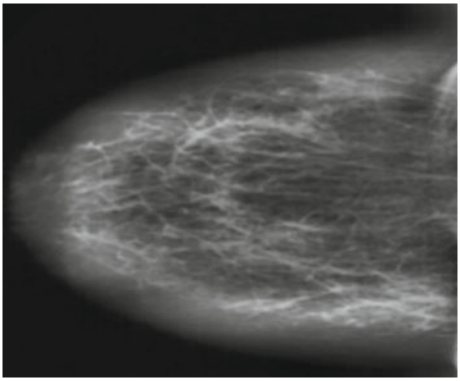


Fig. 3 Test image with tumors

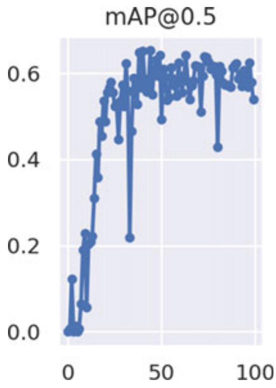
Fig. 4 Test image without tumor



model's detections are. The graph in Fig. 5 contains mAP@o.5 score on y-axis and epoch values on the x-axis. The mAP@0.5 stands for mean average precision values calculated at IOU threshold 0.5. The highest mAP obtained is 0.65 which explains that the model accuracy is 65%.

Precision is a statistic used to determine the percentage of samples that actually tested positive as opposed to just being expected to test positive. A high accuracy number shows that the algorithm has a low false positive rate and is effective at finding true positives. The model has attained precision up to 65%. Recall is a measure used in machine learning to assess how well a binary classification algorithm is performing. This is the ratio of accurately anticipated positive cases to all actually occurring positive instances in the dataset. Recall, then, gauges a model's capacity to accurately recognize every positive instance. In Fig. 6, the graph on the left plot's epoch on x-axis and precision on y-axis, and recall of the trained model. The graph

Fig. 5 mAP@0.5 versus epoch value



on the right in Fig. 7 plots recall on y-axis and epoch on x-axis. The model has attained recall value up to approximately 67%.

Fig. 6 Precision versus epoch value

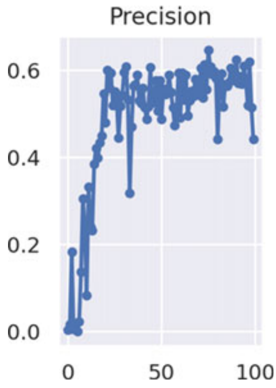


Fig. 7 Recall versus epoch value

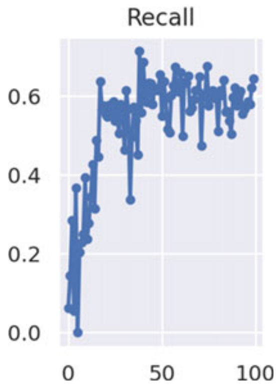


Table 3 Range of confidence value for tumors

Number of tumors present in image	Range of confidence value for bounding box
1	0.67–0.80
2	0.63–0.70
>2	0.55–0.68

Apart from these metrics, the model also gives confidence value of the bounding box. Table 3 gives the summary of confidence values of the bounding box around tumor. The results show that when a single tumor is present in the image, the confidence value attained is higher. It can be concluded that as the number of tumors in the image increases, the confidence value of bounding box decreases.

5 Future Scope

The accuracy, precision, and effectiveness of YOLO v7 for breast cancer detection on mammograms must be improved. One possible approach is to improve the quality of the mammograms and standardize the image acquisition process along with advanced data analytics techniques. Another potential solution is to explore ways to reduce false positives and negatives in the detection process, such as by adjusting the threshold for classification or incorporating additional features into the model. Further enhancing the model's accuracy is the use of a bigger dataset that includes details on the two phases of breast cancer (benign and malignant). Future research can also focus on developing more sophisticated algorithms and techniques for breast cancer detection that can provide even higher accuracy and reliability.

6 Conclusion

The early detection of breast cancer is crucial for effective treatment and a better chance of recovery. YOLO v7, a deep learning algorithm, has shown great potential in assisting in the detection process by automatically identifying regions of interest on mammograms that may contain cancerous tumors. The use of YOLO v7 can also increase the sensitivity of the screening process by detecting tumors that may not be visible to the human eye. In this study, the mammogram database was used for the classification of breast cancer, and the results indicate that YOLO v7 can improve the accuracy of diagnosis. Therefore, the use of YOLO v7 in breast cancer detection has promising potential for improving the accuracy and efficiency of diagnosis, ultimately leading to better patient outcomes.

References

1. Hamed G, Marey M, Elsayed Amin S, Tolba MF (2021) In: Automated breast cancer detection and classification in full field digital mammograms using two full and cropped detection paths approach. IEEE. https://doi.org/10.1109/ACCESS.2021.3105924
2. Das ABM, Satapathy SM (2021) In: YOLO as a region proposal network for diagnosing breast cancer. IEEE. https://doi.org/10.1109/GHCI50508.2021.9513988
3. Mohiuddin A, Basharat A, Ghani U, Peter V, Abbas S, Naeem OB, Rizwan M (2022) In: Breast tumor detection and classification in mammogram images using modified YOLOv5 network. Hindwi 2022. https://doi.org/10.1155/2022/1359019
4. Al-masni MA, Al-antari MA, Park JM, Gi G, Kim TY, Rivera P, Valarezo E, Han S-M, Kim T-S (2017) In: Detection and classification of the breast abnormalities in digital mammograms via regional convolutional neural network. IEEE. https://doi.org/10.1109/EMBC.2017.8037053
5. Wang JL, Ibrahim AK, Zhuang H (2018) In: A study of automatic detection of IDC breast cancer with convolutional neural networks. In: 2018 International conference on computational science and computational intelligence (CSCI). https://doi.org/10.1109/CSCI46756.2018.00141
6. Khuriwal N, Mishra N (2018) In: Breast cancer diagnosis using deep learning algorithm. In: International conference on advances in computing, communication control and networking (ICACCCN2018). https://doi.org/10.1109/ICACCCN.2018.8748777
7. Hazra R, Banerjee M, Badia L (2020) In: Machine learning for breast cancer classification with ANN and decision tree. IEEE. https://doi.org/10.1109/IEMCON51383.2020.9284936
8. Masud M, Eldin Rashed AE, Shamim Hossain M (2017) In: Convolutional neural network-based models for diagnosis of breast cancer. IEEE. https://doi.org/10.1007/s00521-020-05394-5
9. Solawetz J (2022) What is YOLOv7? a complete guide. https://blog.roboflow.com/yolov7-breakdown
10. Redmon J, Farhadi A (2016) YOLO9000: better, faster, stronger. pp 7263–7271. arXiv 2016
11. Redmon J, Farhadi A (2018) In: YOLOv3: an Incremental Improvement. arXiv 2018. https://doi.org/10.48550/arXiv.1804.02767
12. Bochkovskiy A, Wang C-Y, Mark Liao H-Y (2020) In: YOLOv4: optimal speed and accuracy of object detection. arVix 2020. https://doi.org/10.48550/arXiv.2004.10934
13. Zhu X, Lyu S, Wang X, Zhao Q (2021) In: TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios, pp 2778–2788. arVix 2021
14. Li C, Li L, Jiang H, Weng K, Geng Y, Li L, Ke Z, Li Q, Cheng M, Nie W, Li Y, Zhang B, Liang Y, Zhou L, Xu X, Chu X, Wei X, Meituan XW (2022) In: YOLOv6: a single-stage object detection framework for industrial applications. arVix 2022. https://doi.org/10.48550/arXiv.2209.02976
15. Wang C-Y, Bochkovskiy A, Mark Liao H-Y (2022) 1 Institute of Information Science, Academia Sinica, Taiwan. In: YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arVix 2022. https://doi.org/10.48550/arXiv.2207.02696

Touch-Free Biometric Authentication System

Dhanraj Jadhav, Tushar Mahajan, Sarvesh Morey, Ganesh Masute, and Aditya Gujrathi

Abstract Touch-free biometric authentication systems have gained increasing attention in recent years due to their potential to provide secure and convenient authentication methods without requiring physical contact. This literature survey paper provides a comprehensive overview of the current state of research on touch-free biometric authentication systems. The paper reviews various types of biometric modalities used in touch-free authentication systems, including face, iris, voice, and behavioural biometrics. Additionally, the paper analyzes the advantages and disadvantages of touch-free biometric authentication systems compared with traditional biometric systems, as well as the challenges and future directions for touch-free biometrics research. Finally, the paper discusses the impact of the COVID-19 pandemic on the development and adoption of touch-free biometric authentication systems. Through this survey, we hope to provide insights into the potential of touch-free biometrics as a key technology for secure and contactless authentication in a variety of applications.

Keywords Deep sparse filter · Iris recognition · Face recognition · Biometrics · Machine learning · Identity authentication · Fingerprint classification · Eigen faces · Picture pre-purifying · Arbitrarily forest · Support vector machine · Hough transform · Face detection · Fisher faces · Singularity feature · Android application · Voice feature · Visible light

1 Introduction

Biometric identification systems have become increasingly important for identifying individuals, with applications in security measures such as banking transactions, Aadhar identity, high-security wireless access, biometric attendance, and authorized personnel identification for restricted areas. Unlike traditional identification methods such as ID cards or passwords, which can be lost, stolen, or forgotten, biometric systems utilize human traits that are unique to each individual, potentially reducing

D. Jadhav (✉) · T. Mahajan · S. Morey · G. Masute · A. Gujrathi
JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, India
e-mail: dsjadhav_comp@jspmrscoe.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_27

instances of fraud. Additionally, biometric systems offer the convenience of not requiring physical cards or keys.

2 Face Recognition

Face recognition technology uses computer vision and machine learning algorithms to identify and verify a person's identity based on their facial features. The process involves face detection, feature extraction, and face matching. Face detection algorithms are used to locate faces in an image or video frame, and feature extraction algorithms are used to extract facial features such as the eyes, nose, and mouth. The extracted features are then compared with a database of known faces to determine the person's identity. This is done using methods such as PCA, LBP, or CNNs. Face recognition technology has various applications such as security and surveillance, access control, and identity verification. It offers a fast and accurate way to identify individuals, making it a valuable tool in various industries.

2.1 Related Work

There has been notable research and development in the domain of face recognition, with numerous studies focusing on improving the accuracy and efficiency of the technology. Some of the notable works and approaches in this field include:

1. Eigenfaces: It uses principal component analysis to extract features from face images and represent them as dimensional space. The method has since been widely used in face recognition and has been extended to incorporate more complex feature extraction methods.
2. Local binary patterns (LBP): It is an approach for extracting features from textures. It is mostly used for face recognition and has achieved higher accuracy in its variants.
3. Convolutional neural networks (CNNs): They are deep learning models that have demonstrated promise in several computer vision applications, including face recognition. Using benchmark datasets, CNN-based facial recognition models have obtained cutting-edge results.
4. FaceNet: It is a deep learning model for face recognition developed by Google in 2015. The model uses a triplet loss function to learn discriminative features from facial images and has achieved high accuracy rates on benchmark datasets.
5. DeepFace: To extract information from face photos, it employs a deep convolutional neural network, which has demonstrated great accuracy on benchmark datasets.
6. VGG-Face: VGG-Face is a deep learning model evolved by the Visual Geometry Group at the University of Oxford in 2015. The model uses a 16-layer CNN

to extricate features from facial image files and has achieved state-of-the-art performance on benchmark datasets.

2.2 Algorithms

1. Principal component analysis (PCA): The linear dimensionality trimming method involves identifying eigenvalues and eigenvectors of the facial image data's quantiles matrix to form principal components that are used to project face images onto a lower dimensional space.

$$X = U * S * V^T$$

where X—face image data, U—eigenvectors, S—diagonal matrix of eigenvalues, and V—transpose of U.
2. Local binary patterns (LBP): LBP is an approach for extricating features from textures that captures the local texture patterns in the face image. The algorithm involves comparing the potency values of each pixel with its neighbouring pixels and assigning a binary code based on the comparison result. The binary codes are then coupled to form a feature vector.

$$LBP_{P,R} = \sum_{n=0}^{P-1} 2^n (g_n - g_c)$$

where P is the number of adjoining pixels, R is the radius, g_n is the intensity value of the neighbouring pixel, g_c is the intensity value of the central pixel, and $LBP_{P,R}$ is the binary code.
3. Convolutional neural networks (CNNs): They are deep learning algorithms that can automatically identify distinguishing characteristics in face photos. Convolutional, pooling, and fully linked layers are among the many layers that make up the network. The fully connected layers do classification based on the collected features while the convolutional layers extract local characteristics from the face image files.

$$\text{Convolution} : y_{i,j,k} = \sigma\left(\sum_{l=1}^{L} w_{I,j,k,l} * x_{I,j,l} + b_k\right)$$

where y is the output vector, w is the weight matrix, x is the input vector, b is the cross term, and σ is the activation function.
4. Triplet loss: It is a loss function used in face recognition to learn discriminative features from face images. The loss function seeks to reduce the distance between characteristics of similar identity and maximize the distance between features of different identities. The loss is computed based on triplets of face images, where

each triplet includes an anchor image, a positive image (with the same uniformity as the anchor), and a negative image (with a different uniformity from the anchor).

$$L = \sum_{i=1}^{N} \max(0, d_{i,a} - d_{i,p} + \text{margin}) + \max(0, d_{i,a} - d_{i,n} + \text{margin})$$

where N is the number of triplets, $d_{i,a}$ is the interval between the features of the anchor image, $d_{i,p}$ is the interval between the features of the positive image, and $d_{i,n}$ is the interval between the features of the negative image and margin.

5. Recurrent neural networks (RNNs): They are the manifestation of the neural network created specifically to handle data sequences

   They use recurrent layers to capture the profane dependencies between the input data. RNNs have been used for face recognition tasks where the input is a sequence of facial landmarks or features. The mathematical equations for RNNs include the recurrent operation, which is defined as:

$$h_t = \sigma(w_h * h_{t-1} + w_x * x_t + b)$$

   where h_t is the output of the recurrent layer at time t, x_t is the input at time t, w_h, and w_x are the weight matrices, and b is the oblique vector.

6. Siamese neural network: The Siamese network is a common neural network architecture for face recognition and image similarity tasks. It consists of two identical sub-networks that share the same weights and architecture. The network takes a pair of input images and produces a similarity score that indicates whether the two images belong to the same identity or not. The network is trained on pairs of images with labels indicating whether they belong to the same identity or not, and the network learns to minimize the distance between the features of the same identity and maximize the distance between the features of different identities.

Figure 1 demonstrates the overview of the Siamese neural network and the steps involved in it.

The mathematical equation for the Siamese network is as follows:

Let $X1$ and $X2$ be two input images, and $f(X)$ be the feature extraction function:

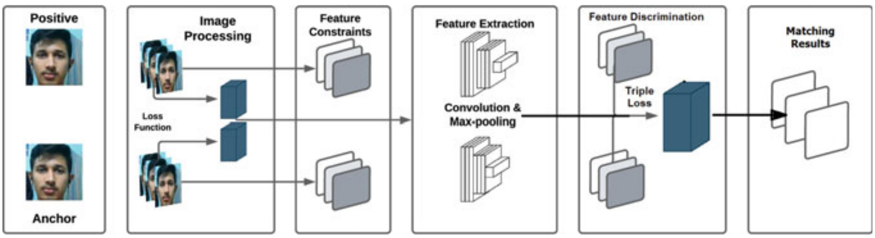


Fig. 1 Data flow diagram

$$f(X1) = H1\ f(X2) = H2$$

where $H1$ and $H2$ are the extracted features from the two images.

Thus, the two photos' similarity score S may be calculated as:

$$S = \frac{1}{1 + d(H1, H2)}$$

where d $(H1, H2)$ is the Euclidean gap between the two feature vectors.

The contrastive loss function is used to train the Siamese network. It seeks to minimize the distance between features of the same identity and increase the distance between the features of distinct identities:

$$\begin{aligned} L = {} & (1 - Y) * \frac{1}{2} * d(H1, H2)^2 \\ & + Y * \frac{1}{2} * \max(0, \text{margin} - d(H1, H2))^2. \end{aligned}$$

The contrastive loss function in a Siamese network improves weights by using a margin hyperparameter to control the distance threshold between feature vectors, represented by d $(H1, H2)$, with Y indicating if the images belong to the same identity.

3 Iris Recognition

Iris recognition technology uses an individual's unique iris pattern to identify them, which is captured by specialized cameras. However, this is costly for smartphones. A solution is to use built-in cameras in smartphones that utilize visible light imaging. The process involves segmentation, normalization, feature extraction, and comparison. The algorithm compares the pattern's features with a decision-making template in a database. Preprocessing steps are required due to noise. Iris recognition is a precise and reliable biometric identification method commonly used in access control, security, and identity applications.

3.1 Related Work

1. M.L.: Iris recognition has also been used using machine learning techniques. For instance, classification of iris pictures and increased identification accuracy have been achieved using support vector machines, neural networks, and decision trees.

2. Multimodal biometrics: The use of multiple biometric traits, known as multimodal biometrics, can improve recognition accuracy. Integrating iris recognition with other features such as face, fingerprint, and voice has been found to achieve 86 per cent accuracy. The iris is preprocessed and segmented during the procedure, reducing various sounds using the image and segmenting the iris and pupillary margins using the circular Hough transform. The average accuracy for this stage is 78.25%.
3. Dataset development: To evaluate iris recognition algorithms, several large-scale datasets have been developed, e.g. the CASIA-IrisV4 dataset contains over 10,000 iris images collected from over 2000 individuals, making it a valuable resource for iris recognition research.

3.2 Algorithms

1. Iris-segmentation: The iris region in the acquired image is identified and segmented using techniques like edge detection, circle fitting, or template matching.
2. Feature extraction: The normalized iris picture is processed using a variety of techniques to obtain distinctive characteristics, such as Haar wavelets, Gabor filters, Daubechies wavelets, and local binary patterns.
3. Feature encoding: The extracted features are encoded using methods like Hamming distance, SIFT, or PCA to convert them into a compact and discriminative representation.
4. Matching: The encoded features of the enrolled iris templates are compared with the features of the test iris to find a match. Several matching algorithms are used, such as Euclidean distance, Mahalanobis distance, SVM, and neural networks.
5. Decision-making: The matching results are analyzed and the threshold is used to make a decision or confidence level set by the system. Based on the similarity score between the enrolled and test iris templates, the system can approve or reject the user's claim of identification.
6. SVM algorithm: The support vector machine (SVM) is a popular technique for classification and regression problems in machine learning. It works by representing each data point as a point in N-dimensional space, where N is the number of features, and finding the hyperplane that best separates the two classes. The hyperplane acts as a boundary between the classes, and the SVM algorithm determines the coordinates of each observation relative to this boundary. SVM is commonly used for classification problems, but it can also be used for regression.
7. Normalization: Normalization is a preprocessing method that standardizes numerical data without changing its shape. It helps to balance the scale when feeding data into machine learning systems, and can include advanced algorithms for processing photos in challenging conditions.

4 Fingerprint Recognition

Touchless fingerprint recognition is an emerging technology that allows for fingerprint recognition without physical contact between the fingerprint and the sensor. It uses various sensors like infrared, ultrasound, and optical sensors to capture the fingerprint without contact. The captured fingerprint image is processed using machine learning algorithms, including deep learning techniques such as CNNs, to extract features and match them with stored fingerprint data. Touchless fingerprint recognition is faster, more convenient, and eliminates the need for physical contact, reducing the risk of spreading infections and diseases. This technology is promising and can work even when the user's fingers are wet or dirty, which is not possible with traditional fingerprint recognition systems. Touchless fingerprint recognition finds its applications in various domains like access control, time and attendance, and border control.

4.1 Related Work

1. "Touchless Fingerprint Recognition on Mobile Devices" by Aggarwal et al.: This study proposes a touchless fingerprint recognition system for mobile devices using two cameras and a light source. The device records a 3D image of the user fingertip and extracts features for recognition. The authors reported an accuracy rate of 95% in their experiments.
2. "Touchless 3D Fingerprint Recognition Based on Point Cloud Analysis" by Zhang et al.: This study proposes a touchless 3D fingerprint recognition system that uses a structured light sensor to capture the 3D information of the fingerprint. The authors used a point cloud-based feature extraction method and a neural network-based recognition algorithm. The authors reported an accuracy rate of 96.5% in their experiments.
3. "Touchless Fingerprint Recognition Using Thermal Imaging" by Kim et al.: This study proposes a touchless fingerprint recognition system using thermal imaging. The system captures the thermal signature of the user's fingertip and extracts features for recognition. The authors used a template-matching algorithm for recognition and reported an accuracy rate of 99% in their experiments.
4. "A New Touchless Fingerprint Recognition System Based on Fourier Transform" by Li et al.: This study proposes a touchless fingerprint recognition system that achieves high accuracy rates using a Fourier transform-based feature extraction method and a convolutional neural network. The results show that touch-free fingerprint recognition systems can provide a non-intrusive and convenient way for biometric authentication. However, further research is needed to address scalability, robustness, and security challenges.

4.2 Algorithm and Techniques for Fingerprint Authentication

1. Phase-based algorithms: These algorithms extract features from the phase information of the fingerprint, which is captured using an infrared or thermal sensor. The phase information is less affected by variations in skin texture and pressure, making it a suitable choice for touchless fingerprint authentication. Examples of phase-based algorithms include the Steerable Pyramid Transform (SPT) and the phase-only correlation (POC) algorithm.
2. Texture-based algorithms: Texture-based algorithms are used in touchless fingerprint recognition systems that analyze the texture and microstructure of fingerprints, creating a fingerprint template for authentication. These algorithms are often used with optical sensors like CMOS or CCD sensors. Examples of texture-based algorithms include the scale-invariant feature transform (SIFT) algorithm and the local binary pattern (LBP) algorithm.
3. Wavelet-based algorithms: These techniques extract features from the fingerprint picture using wavelet transforms, which can subsequently be exploited to create a fingerprint template for authentication. Wavelet-based algorithms are often used in touchless fingerprint recognition systems that use optical or thermal sensors. Examples of wavelet-based algorithms include the Gabor wavelet transform and the discrete wavelet transform (DWT) algorithm.
4. Hybrid algorithms: These algorithms combine multiple feature extraction methods to create a more robust and accurate touchless fingerprint authentication system. For example, a hybrid algorithm may combine phase-based and texture-based methods to extract features from both the phase and texture information of the fingerprint.
5. Artificial neural networks (ANN): Another kind of machine learning technique for touchless fingerprint identification is an ANN. Similar to CNNs, ANNs can learn from the input data, extract features, and predict who an individual is. ANNs can be trained on large datasets of fingerprint images and can be optimized using techniques such as backpropagation.
6. Dynamic time warping (DTW): DTW is a pattern-matching algorithm that can be used for touchless fingerprint authentication. DTW is used to measure the similarity between two temporal sequences and can be used to compare the features extracted from different fingerprint images. DTW has been used in combination with other algorithms, such as wavelet transforms and Gabor filters to improve the accuracy of touchless fingerprint authentication.
7. Minutiae-based algorithms: These algorithms use the distinctive fingerprint characteristics including ridge ends and bifurcations, and utilize them to build a template that may be used for authentication and comparison. Numerous touch-based and touchless fingerprint identification systems employ minutiae-based algorithms.
8. Pattern recognition algorithms: These algorithms analyze the overall pattern and structure of the fingerprint, including the ridges, valleys, and pores to create a template that can be used for authentication. Pattern recognition algorithms

are often used in touchless fingerprint recognition systems that use optical or ultrasound sensors.

5 Diagram

Figure 2 architecture diagram depicts the flow of the system that we are going to implement. It basically consists of two steps:

1. Model creation.
2. Real-time recognition

1. **Model Creation**

The user credentials are captured using a camera and stored in a database folder, and then used for model creation with feature extraction, image comparison, and testing against negative data in the user credential block.

2. **Real-time Recognition**

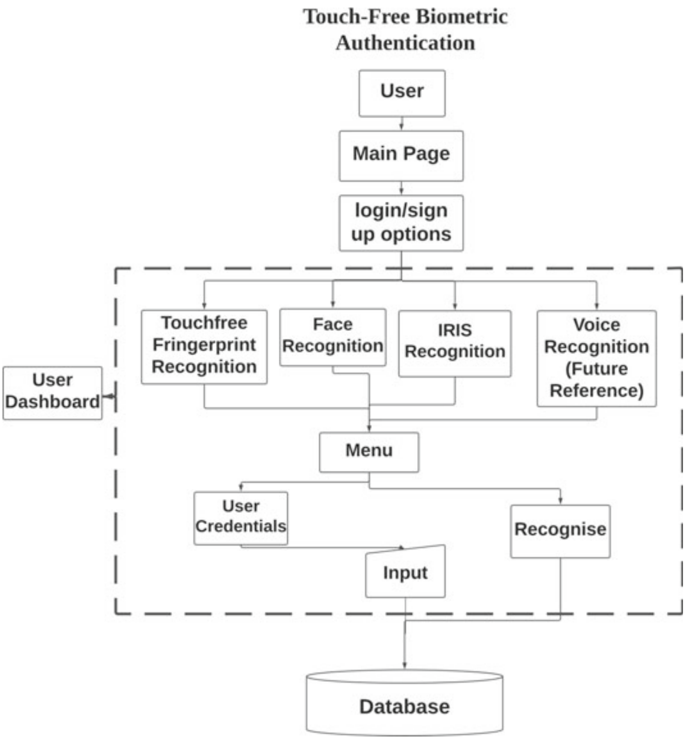


Fig. 2 Architecture diagram

This step involves real-time recognition of the user by testing their images against the previously created model. The input block captures and extracts feature from the current images and compares them with the model's features to determine a successful match.

6 Literature Review

Year of publication	Title	Summary
2022	A Siamese Neural Network-Based Face Recognition from Masked Faces	The proposed Siamese neural network (SNN) model in this study uses half of the human face to identify similarities and differences between two faces, providing higher accuracy and faster search times for partial face matching than VGGFace2. The model achieves 84.8% mean accuracy and 93% BO5 accuracy [1]
2022	Face Recognition Method Based on Siamese Networks under Non-Restricted Conditions	The current non-finite face recognition methods are limited by their susceptibility to quick convergence and weak resilience, and their inability to adapt to different situations. This study proposes a design that integrates frequency feature sensing, LBP, and Siamese neural networks to improve recognition accuracy, while reducing the number of iterations needed for recognition [2]
2021	Research Paper on Biometrics Security	The quality of fingerprints acquired from mobile phone cameras is resolution dependent, at first sight it seems to be a drawback, but the rising technology trends that provide best resolution cameras in cheaper phones make resolution dependency, less of a problem [3]
2021	An Overview of Touchless 2D Fingerprint Recognition	The study aims to establish interoperability between touchless and touch-based fingerprint identification systems, where touchless methods have high user acceptance but lower biometric performance. Future research areas include biometric template protection and presentation attack detection (PAD). The study discusses the introduction of the first commercial touchless fingerprint identification devices [4]

(continued)

(continued)

Year of publication	Title	Summary
2019	Towards Smartphone-based Touchless Fingerprint Recognition	This work presents a touchless fingerprint recognition system that includes an Android app, server, and feature extraction/matching modules. The system incorporates a unique monogenic-wavelet-based touchless fingerprint enhancement technique that significantly increases matching accuracy. Experimental results show improved rank-1 accuracy and EER with the proposed enhancement technique. The work provides a valuable resource for further research in the area of touchless fingerprint recognition [5]
2017	Developing Iris Recognition System for Smartphone Security	Work of an iris system for smartphones develop and tested. Method relies on visible-wavelength eye pictures, which are captured by the smartphone's camera [6]
2016	Face Recognition on Mobile Platforms	The application, written in C+ +11, can run on PCs and smartphones, but performance is currently slow on mobile devices due to the lack of optimization for ARM processors. Further tuning and optimization of the settings will be required [7]
2014	Touchless Fingerprint Biometric Survey of 2D and 3D Technologies	Touchless recognition systems are classified as two types: 2D data and 3D models. Most of the two-dimensional systems use a single camera and use enhancement and resolution normalization methods to create touch-equivalent pictures. These systems are usually less costly than three dimensional technologies, but they have issues with perspective deformations and non-constant sample resolution [8]

6.1 Conclusion and Future Work

This paper describes an open-source biometric identification system for Android phones that extracts biometric images, minutiae, and executes secure matching using mobile-webserver connection. The system provides a simple and feasible option to obtain biometric images from mobile phones with better 3D resolution than dedicated devices, improving the matching accuracy. Future experiments with more datasets gathered using mobile phone cameras and higher performance devices for training can further improve the system's accuracy. The system will be made available as an API component for high-level authorization applications in the future, and scanning factors instead of recording images will improve photo quality.

References

1. Chatterjee R, Roy S, Roy S (2022) Siamese neural network-based face recognition from masked faces. 1534:13. DOI: https://doi.org/10.1007/978-3-030-96040-7_40
2. Song C, Ji S (2022) Face recognition method based on Siamese networks under non-restricted conditions. 10:40432–40444. https://doi.org/10.1109/ACCESS.2022.3167143
3. Pujari V, Patil R, Sutar S (2021) Research paper on biometrics security. In: Emerging advancement and challenges in science, technology and management; ResearchGate, pp 74–79. https://www.researchgate.net/publication/352508064
4. Priesnitz J, Rathgeb C, Buchmann N et al (2021) An overview of touchless 2D fingerprint recognition. J Image Video Proc 8(2021). https://doi.org/10.1186/s13640-021-00548-4
5. Birajadar P, Haria M, Kulkarni P, Gupta S, Joshi P, Singh B, Gadre V (2019) Towards smartphone-based touchless fingerprint recognition. Sādhanā 44:161. https://doi.org/10.1007/s12046-019-1138-5
6. Elrefaei LA, Hamid DH, Bayazed AA, Bushnak SS, Maasher SY (2018) Developing iris recognition system for smartphone security. Multimed Tools Appl 77:14579–14603. https://doi.org/10.1007/s11042-017-5049-3
7. Bertok K, Fazekas A (2016) Face recognition on mobile platforms. In: 2016 7th IEEE International conference on cognitive infocommunications (CogInfoCom). https://doi.org/10.1109/CogInfoCom.2016.7804521
8. Labati RD, Genovese A, Piuri V, Scotti F (2014) Touchless fingerprint biometrics: a survey on 2D and 3D technologies. J Internet Technol 15(3). https://doi.org/10.6138/JIT.2014.15.3.01

Partial Ground-Inserted H-shaped Antenna for Beyond 5G Applications

Smrity Dwivedi

Abstract Inserted H-shaped antenna with partial ground has been presented here with three major designs and their comparison at terahertz frequency for beyond 5G applications. This new era of beyond 5G is useful to connect most of the devices and create a proper balance between systems and users as per demand of technology. Antenna design is proposed with normal rectangular patch and two more modifications have been done in same patch and it is simulated with the help of CST microwave studio for its important parameters such as bandwidth and gain. For the three structures first design, second design and third design bandwidth is 100 GHz, 250 GHz and 400 GHz with gain 13.5 dBi, 14.4 dBi and 14.8 dBi, respectively. A simple PEC material as a radiating element and polyamide as substrate have been taken. Partial ground has been taken with microstrip line feedline.

Keywords 5G · Microstrip line · CST · Polyamide · THz

1 Introduction

After development of fifth-generation technologies, researchers, scientists and engineers are looking for wide bandwidth which should be improved in wireless systems and devices to provide better services and fast experience. Because of that, beyond 5G, a sixth generation is next step to be widely used in the early 2028 or intensively used around 2030 as per demand. After the evolution of beyond 5G, many things have been sorted in which power which is transmitted, techniques of modulation, as well as the frequency which is allotted to this new band. From this, parameters like gain, bandwidth, efficiency, size of the antenna, technique of fabrication and other useful quantity can be modified and becomes maximum and also the cost of fabricating of such antenna even other devices will be reduced and developed in bunch. The microstrip patch antennas having with high demand and their compactness, as well as ease of fabrication, are widely used [1–2]. Many universities have already

S. Dwivedi (✉)
Department of Electronics Engineering, IIT-BHU, Varanasi, India
e-mail: sdwivedi.ece@itbhu.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_28

published many research items related to 6G technologies, its advantages and few designs, but that time there was no use and even they were not properly written [3]. Then, many research institutes in collaboration with industries, are working and have produced few important results in this area [4]. This new technology for beyond 5G will basically work in terahertz frequency range which is already defined for range 0.1 and 10 THz, and then it is divided into two more categories, according to frequency range as given in Fig. 1. [5–7]. Here, from $0.1 - 0.3$ THz is known as a sub-terahertz region and next part is known as pure terahertz band later from $0.3-$ 10 THz [8]. This high frequency band is also known as the terahertz gap just because of lack of materials which could respond to these high frequencies [9]. This frequency band may also be known as tera waves [10], as per Fig. 1, which represents the frequency range in sixth-generation wireless technology and system. As previously reported by many researchers [6] that beyond 5G are the smart systems giving secure information and it will run using artificial intelligence technology which will connect every smart device with smart intelligence system.

In today's research area, there are many types of antenna which are being used according to design and application for high frequency like reflect array antenna [10] and are giving better performance with a new substrate material like polyamide at high speed and useful in many applications such as spectroscopy in research and biomedical field. Also, few more designs like elliptical pattern antenna which has high gain and good radiation pattern with ease of fabrication can be used for any distance whether it is short or long [11]. Due to increase in demand of changing frequency, polarization and radiation pattern, reconfigurable antennas have been introduced with many smart materials like graphene, metamaterials, metasurfaces, etc., giving MIMO applications with less loss and high gain and with enhanced bandwidth at high frequency. Depending on distance covered, fading effect and introduction of solid state devices with these antenna are creating new interest between researchers

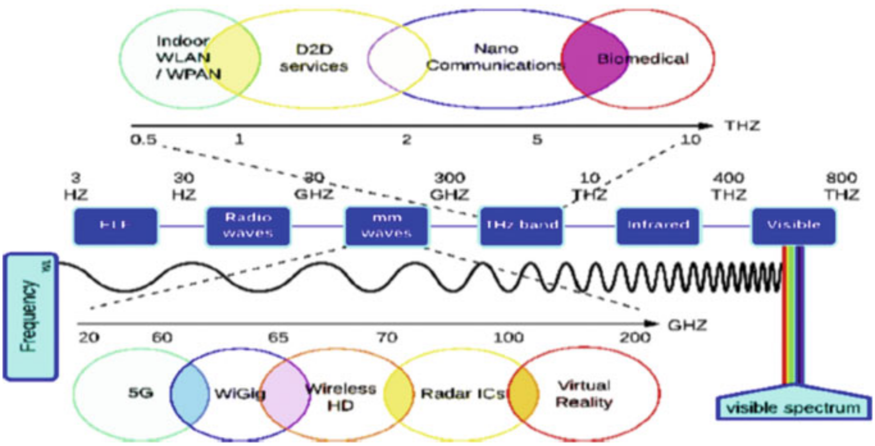


Fig. 1 Range of frequency in beyond 5G and sixth generations [9]

as dielectric resonator antenna, MIMO antenna with CMOS technology are the best suited for getting proper dispersion and less fading effect. Nano technology is also being used to get high bandwidth and less isolation effects [8, 9, 12-14, 16-17].

Hence, to develop future generation beyond 5G, frequency is increased into THz region [6] and for that frequency, fabrication is very complex as it has very small size even in micrometer which is very difficult to make and measure as well. For that purpose, very good materials like graphene, polyamide, metasurface, gold, etc., are required, and also well-equipped instruments are needed to perform complex fabrication and its measurement such as high frequency vector network analyzers. This frequency demands small dimensions as per wavelength and also few more additional components in some cases and costly substrate material and for the radiating elements. In other word, this frequency region will face free space loss and molecular absorption as sky wave and ground wave communication system simultaneously. As per high frequency, few materials like perfectly electric conductors show that high loss cannot work properly at this frequency. Even some dielectric materials as mentioned above, can work at this high frequency, but some improvement and advancement are required so that they can properly work without having ohmic loss and dielectric loss. Even few new technologies are evolved such as use of surface Plasmon or SPP which is having less loss and behaving like slow wave structures at required and desired modes, as they are also used as wave [18, 19]. In this paper, rigorous literature review has been done to see the new challenges and technologies being involved in design of antenna for beyond 5G technologies such as frequency, bandwidth, materials, fabrication, implementation and loss have been reviewed. In Sect. 1, design of antenna has been proposed in the frequency range of 0.1–5 THz beyond 5G and simulated using CST software. There are three designs that have been introduced in which first is simple patch antenna and second is designed by cutting the patch from middle and connected with small patch at upper half section, even in third design H-shaped is inserted in the upper half section from second design with microstrip feed transmission line (0.1−5 THz). Here, there are three types of designs that are discussed for its improvement in gain and bandwidth to check its performance and scope for other applications and improvement. All important parameters related to antenna have been calculated and simulated in which S11 parameters, gain and bandwidth are important in Sect. 2 and very well explained. Radiation pattern is the major important parameter which is well given to achieve the polarization of antenna. Also Sect. 2 depicts the design of H-shaped inserted antenna with proper dimensions with its feedline and operating frequency. Section 4 is fully dedicated to the conclusion of the given design in terms of major important parameters. Here, the design of antenna starts with patch having PEC material and polyamide substrate material as dielectric material. Return loss of the proposed antenna has been increased below negative side with introduction of additional elements as given in design two and three. The final part is Sect. 4 as a conclusion has been summarized in form of table with explanation.

2 Partial Ground-Inserted H-shaped Antenna Design

Design of Antenna: The paper deals with frequency range from 0.1 to 5 THz just because of high frequency electromagnetic property is mentioned in between 0.1 and 10 THz. As per standard of communication network, this frequency range is dedicated to be at 0.3 − 10 THz [17]. Apart from it, other important conferences related to radio frequency have been used in 275 − 450 GHz as per requirement of wireless or with wire. On the other hand, other more standards especially related to this frequency have been used in 252 − 325 GHz. The H-shaped inserted antenna is designed on polyamide substrate of height (h) 20 μm having a size of 600 × 600 μm as shown in Fig. 2a–e and feed line from 130 μm to 185 μm. Initially, this antenna is made by single patch having size 315 × 400 μm. Then in second step, patch is cut in half section and small patch is designed in upper part of first patch having size 200 × 157 μm^2. In third design, a H-shaped antenna is inserted in upper half patch from second design. The frequency of an inserted H-shaped antenna is $f = c/(2 \times L \times \sqrt{}\varepsilon_{\text{eff}})$ where, L represents the length of the antenna, ε_{eff}. is the effective dielectric constant of the microstrip line and c is the velocity of light in the air ($c = 3 \times 10^8$ m/ sec). The proposed H-shaped antenna is designed, simulated and optimized in CST Microwave Studio to obtain at frequency of 0.1–5 THz. The simulated results show the variation in bandwidth and gain 100, 250 and 400 GHz with gain 13.5 dBi, 14.4 dBi and 14.8 dBi, respectively. Figure 2d represents second design with half patch from first design and small patch from second design as well as Fig. 2e shows inserted H-shaped in upper half section from second design (Table 1).

3 Antenna Analysis

After designing the structure for single patch, double patch and inserted H-shaped, the very important parameters are simulated, which S11 represents the total reflection from the antenna in form of return loss. This should be below −10 dB for getting vswr around 2. Figure 3 shows the return loss for all types structure designed as given. For single patch, double patch and inserted H-shaped, return loss is −20 dB, −25 dB and 28 dB, respectively, around 3.5 THz frequency.

Figure 4a −c shows the radiation pattern of single patch, double patch and inserted H-shaped antennas with their gain values at frequency 3.5 THz. From the given figures, it is clear that gain of third design is better than other two structures at 3.5 THz frequency, which may be enhanced for further design but fabrication will become complex, but this analysis moves around 3.5 THz, so for beyond 5G, third design inserted H-shaped antenna performance is better when compared with return loss which is less when compared with other one.

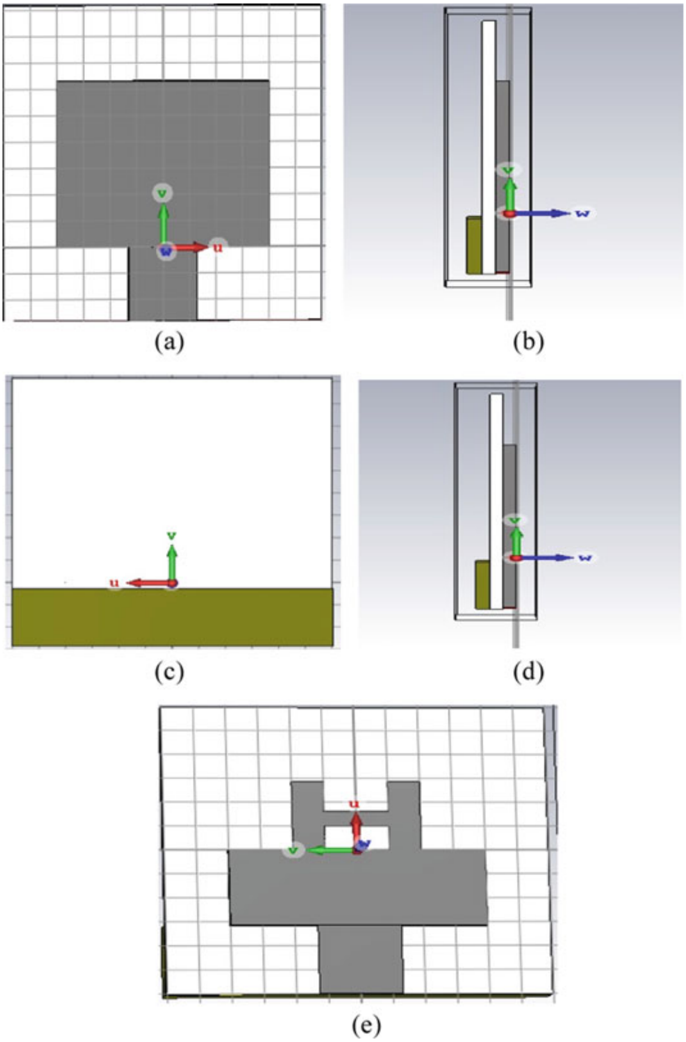


Fig. 2 **a** Radiating patch and substrate are represented in above figure with front view, **b** side view of antenna with microstrip feedline, **c** back view of antenna with partial ground plane, **d** antenna with second patch, **e** antenna with inserted H-shaped

Cumulated gain plot is given in Fig. 5 where all gains have been plotted against frequency, which is 13.5 dBi, 14.4 dBi and 14.8 dBi for single patch, double patch and inserted H-shaped antenna, respectively.

Table 2 has summarized all the important parameters like bandwidth and gain for given antenna designed.

Table 1 Summary of antenna parameter length

Description	Unit
Height of substrate	20 μm
Length of substrate	600 μm
Width of substrate	600 μm
Length of patch	400 μm
Width of patch	315 μm
Length and width of second patch	$200 \times 157\ \mu m^2$

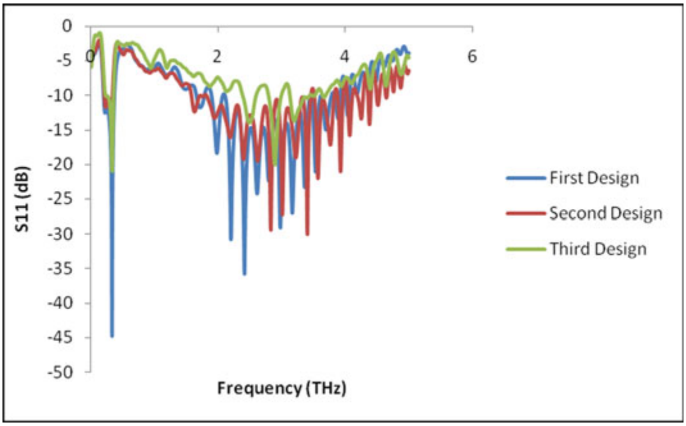
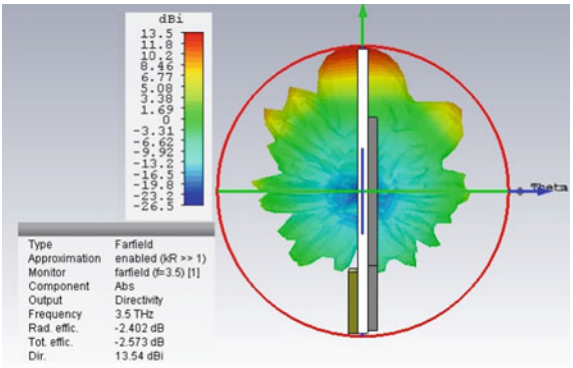
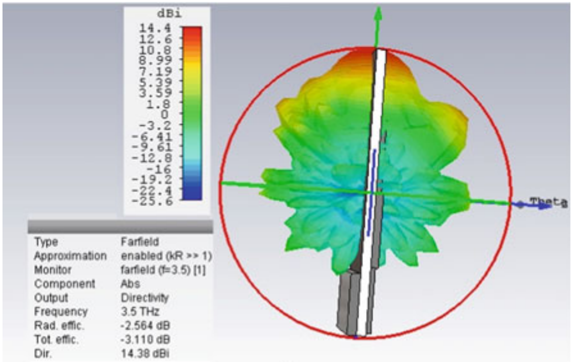


Fig. 3 S11 for all types of structures designed

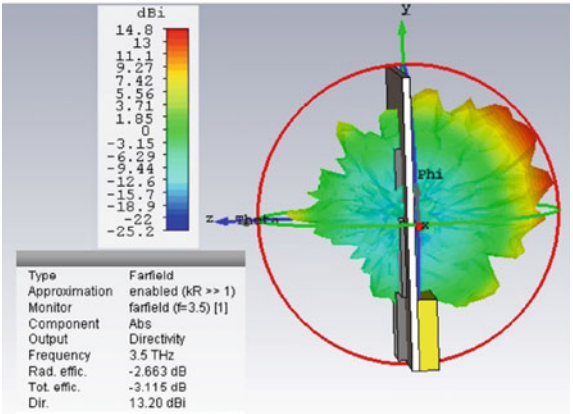


(a)

Fig. 4 Radiation pattern with their gains **a** six leaves, **b** four leaves, **c** two leaves antenna for frequency 9 THz



(b)



(c)

Fig. 4 (continued)

4 Conclusion

In this design, there are several key features which have been noticed like return loss, gain and bandwidth, as already discussed in introduction section, about beyond 5G properties which is required, has been resulted here. Single patch, double patch and inserted H-shaped antennas are designed and simulated using CST microwave studio, so according to result obtained, inserted H-shaped antenna having gain 14.8 dBi with 400 GHz bandwidth has better performance when compared with other one. Scope for further design is restricted due to complex fabrication and design parameters. According to Table 2, this design can be thinkable for beyond 5G application.

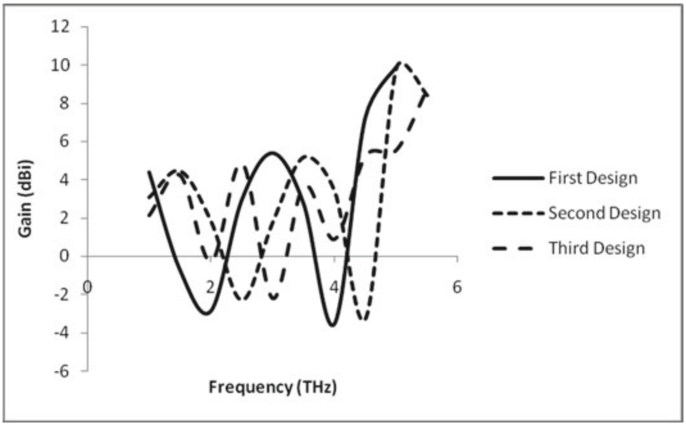


Fig. 5 Gain versus frequency plot for all two leaves, four leaves and six leaves antenna have been plotted

Table 2 Comparison table of proposed structures

Proposed structure	Bandwidth (GHz)	Gain (dBi)
First design	100	13.5
Second design	250	14.4
Third design	400	14.8

References

1. Singhal S (2019) Elliptical ring terahertz fractal antenna. Optik. https://doi.org/10.1016/j.ijleo.2019.163129
2. Guo L, Huang F, Tang X (2014) A novel integrated MEMS helix antenna for terahertz applications. Optik 101–103
3. Samsung Research (2020) 6G: the next hyper connected experience for all. Seoul 2020. [Online]. Available: https://research.samsung.com/next-generation-communication
4. Letaief KB, Chen W, Shi Y, Zhang J, Zhang YJA (2019) The roadmap to 6G: AI empowered wireless networks. IEEE Commun Mag 57(8):84–90. https://doi.org/10.1109/MCOM.2019.1900271
5. Duan B (2020) Evolution and innovation of antenna systems for beyond 5G and 6G. In: Frontiers of information technology and electronic engineering, vol 21(1). pp 1–3. https://doi.org/10.1631/FITEE.2010000
6. Chi L, Weng Z, Qi Y, Drewniak JL (2020) A 60 GHz PCB wideband antenna-in-package for 5G/6G applications. IEEE Antennas Wirel Propag Lett 1225:1–1. https://doi.org/10.1109/LAWP.2020.3006873
7. Emara MK, Stuhec-Leonard SK, Tomura T, Hirokawa J, Gupta S (2020) Laser-drilled all-dielectric huygen's transmit-arrays as 120 GHz band Beamformers. IEEE Access 8:153815–153825. https://doi.org/10.1109/ACCESS.2020.3018297
8. Kokkonen M, Myllymaki S, Jantunen H (2020) 3×3 dipole lens antenna at 300 GHz with different permittivity lenses. In: 2nd 6G wireless summit 2020: gain edge for the 6G era, 6G SUMMIT 2020, pp 4–7. https://doi.org/10.1109/6GSUMMIT49458.2020.9083825

9. Fitri I, Akbart AA (2017) A new gridded parasitic patch stacked microstrip antenna for enhanced wide bandwidth in 60 GHz band. In: 2017 International conference on broadband communication, wireless sensors and powering, BCWSP 2017, 2018, vol 2018-Janua. pp 1–4. https://doi.org/10.1109/BCWSP.2017.8272571
10. Miao Z-W, Hao Z-C, Wang Y, Jin B-B, Jing-Bo W, Hong W (2018) A 400-GHz high-gain quartz-based single layered folded reflectarray antenna for terahertz applications. IEEE Trans Terahertz Sci Technol 9(1):78–88
11. Akyildiz IF, Jornet JM (2016) Realizing ultra-massive MIMO (1024×1024) communication in the (0.06–10) terahertz band. Nano Commun Netw 8:46–54
12. Cheng B, Cui Z, Bin L, Qin Y, Liu Q, Chen P, He Y et al (2018) 340-GHz 3-D imaging radar with 4Tx-16Rx MIMO array. IEEE Trans Terahertz Sci Technol 8(5):509–519
13. Luo Y, Zeng Q, Yan X, Yong W, Qichao L, Zheng C, Nan H, Xie W, Zhang X (2019) Graphene-based multi-beam reconfigurable THz antennas. IEEE Access 7:30802–30808
14. Chandra S, Dwivedi S (2019) Comparative analysis of reconfigurable patch antenna array for different liquid crystal substrates. URSI AP-RASC 2019, 09–15 March 2019, New Delhi, India. https://doi.org/10.23919/URSIAPRASC.2019.8738497
15. Chandra S, Dwivedi S (2021) Graphene based radiation pattern reconfigurable antenna. In: International conference for convergence in technology, April 2–4, 2021. https://doi.org/10.1109/I2CT51068.2021.9417938
16. Dwivedi S (2022) Reconfigurable array antenna. Intech Open. https://doi.org/10.5772/intechopen.106343
17. Elayan H, Amin O, Shubair RM, Alouini M-S (2018) Terahertz communication: the opportunities for wireless technology beyond 5G. In: 2018 International conference on advanced communication technologies and networking (Comm. Net) (2018), pp 1–5
18. Chandra S, Dwivedi S (2020) Simulation analysis of high gain and low loss antenna with metallic ring and spoof SPP transmission line. In: 7th International conference on signal processing and integrated networks, SPIN 2020, 27–28 February 2020, Noida, India, Published on IEEE Explore
19. Dwivedi S (2022) Design and analysis of metamaterial waveguide antenna for broadband applications. ICT4SD 2022, Goa, 29–30 July 2022.https://doi.org/10.1007/978-981-19-5224-1_42

Crop Recommendation Assessment for Arid Land Using Machine Learning

Batool Alsowaiq, Noura Almusaynid, Esra Albhnasawi, Wadha Alfenais, and Suresh Sankrayananarayanan

Abstract The agriculture industry plays a significant role in the economy of many countries, and the population is regarded as an essential profession. To increase agricultural production, crops are recommended based on soil, weather, humidity, rainfall, and other variables which are beneficial to farmers as well as the nation. This chapter explores the use of "machine learning" algorithms to recommend crops in for Arid land based on features selected from tropical climate where crops grow effectively. Five "machine learning" models have been validated for the recommendation of crops for arid land which resulted in "Random Forest" topping as the best model.

Keywords Soil · Weather · Humidity · Machine learning · Random forest

1 Introduction

Food and agriculture are the two most important sources of livelihood for most people in the world. With the escalation of climate change, it is becoming increasingly important for farmers to optimize their production in order to maximize their yields and reduce their costs. However, the traditional methods of selecting and recommending crops are insufficient in determining the best possible crop for a given environment. Choosing the wrong crops will always result in lower productivity. Families who rely entirely on agriculture have a difficult time surviving. There has been good amount of work done in employing machine-learning models like "SVM, Naïve Bayes, Decision Tree, Chaid, LEM2" algorithm for selecting the best crop. Most

B. Alsowaiq · N. Almusaynid · E. Albhnasawi · W. Alfenais · S. Sankrayananarayanan (✉)
Department of Computer Science, King Faisal University, Al Hofuf, Kingdom of Saudi Arabia
e-mail: ssuresh@kfu.edu.sa

B. Alsowaiq
e-mail: Batool_al_sowaiq@outlook.com

E. Albhnasawi
e-mail: ealbhnasawi@outlook.sa

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_29

of the work employing machine learning have been done for one or two crops and mostly focused for tropical climatic condition like India and other countries. The reason being most of agriculture production is dependent of availability of fertile soil and tropical climatic condition.

So accordingly, we here have focused on validation different machine-learning model like "SVM, Naïve Bayes, KNN, Decision Tree, Random Forest" in terms of accuracy and error towards developing an machine learning–enabled crop recommendation from Arid land point of view. Though the vision of Kingdom of Saudi Arabia is to promote agriculture in this region toward increasing the yield and productivity, there is a need to develop crop recommendation system driven by machine learning as a first step which would benefit the Arid land. Secondly, we have chosen crop-related dataset with features like "NPK, pH, Temperature, Rainfall, Humidity" which helps in the growth of crop in fertile soil with good climatic condition. So, this dataset has been beneficial for developing a machine learning–enabled crop recommendation which would benefit in selecting the best crop based on real-time data received from Arid land based on model trained for crop recommendation.

This system would provide farmers with the opportunity to consider various agricultural benefits, such as improved yields, reduced costs, and improved decision-making processes. The rest of the chapter is organized as follows. Section 2 discusses the various work carried out toward crop recommendation using machine learning. Section 3 discussed the crop recommendation using machine learning. Section 4 discussed the performance of machine learning model implemented for crop recommendation with results. Section 5 tabulates the results of the model followed by discussion. Section 6 gives the concluding remarks and future work.

2 Literature Review

There has been some good amount of work carried out pertaining to crop recommendation using machine learning algorithms.

Authors in the paper [1] developed a crop recommendation system using ensemble technique with major voting using machine-learning models "Random Tree, CHAID, K-NN and Naïve Bayes". The authors in this work have resulted in an accuracy of 88%. The work has drawback where they have focused on crops pertaining to district in city in India and no other crops. There has been no correlation shown among the features used for crop classification toward recommendation which is very important. There has been no error computed in terms of precision and recall for showing how accurately classification is done in terms of True Positive and True Negatives.

The research in [2] describes the development of an assembling technique of three machine learning models which are "Random Forest, Naïve Bayes, and Linear SVM" towards crop recommendation resulting in an improved crop productivity. The system is based on soil specific physical and chemical characteristics such as "NPK, pH", soil type, pores in soil, average rainfall, sowing season and temperature of the surface. This has been focused on two crops which are "Kharif and Rabi" resulting in an accuracy of 99.91%. The challenge in this work is limited crop used for recommendation though dataset of these crops is large. There has been no error computed in terms of precision and recall for showing how accurately classification is done in terms of True Positive and True Negatives.

Researchers in [3] have worked towards deploying four machine-learning models which are “SVM, Decision Tree, Naïve Bayes, Random Forest, KNN, Logistic Regression”. This resulted in selection of appropriate crop using the machine-learning model with a classification accuracy of 89.66% for “SVM” followed by “KNN, Random Forest, Naive Bayes, Decision Tree and Logistic Regression”. The model has not achieved higher accuracy as the number of samples of each crop is less for some crops and high for some crops. There is also possibility of data imbalance which has not been explored and correlation of soil features contributing for crop selection also not evaluated. There has been no error computed in terms of precision and recall for showing how accurately classification is done in terms of True Positive and True Negatives.

The research paper [4] proposes a methodology for improving agricultural crop yield by accounting for the soil micro and macronutrient levels to predict crop suitability. To achieve this, fuzzy logic and rough set rule induction were used to create rules for the dataset and evaluate different algorithms for their accuracy. The results of the evaluation found that the “LEM2” algorithm gave the highest prediction accuracy of 89% for the dataset without fuzzy logic, while the AQ algorithm showed better accuracy than others with three linguistic variables. Though the work has considered 23 crops with 16 features for crop selection, there has been no usage of machine-learning model deployed for the selection of crops with higher accuracy.

3 Crop Recommendation Using Machine Learning

So, based on literature reviewed pertaining to machine learning for crop recommendation system based on soil features and climatic condition, we in this chapter have worked on dataset collected pertinent toward tropical climate for crop recommendation for Arid land. The reason for choosing dataset pertaining to tropical climate condition is toward training the model with soil features for the recommendation of crop. The model trained and evaluated would be used for crop recommendation for Arid land based on features collected in real time. The availability of Arid land dataset for crop recommendation is not available as agriculture in Arid land is really challenging and is one of the primary focusses of Kingdom of Saudi Arabia. So before going into the results and analysis of machine learning model, we investigate methodology of proposed work pertaining to following models which are “Support Vector machine, Decision tree, Random Forest, K-Nearest Neighbor and Naïve Bayes”. The details about the model are explained in brief.

3.1 Random Forest Algorithm

"Decision Tree" are created based on variety of samples and averaging and majority voting are employed for classification and regressing. It can handle categorical variables in the case of classification. Our project provides solutions to the multiclass problem. The principle of "Bagging" is employed for "Random Forest". "Random Forest" also employed an ensemble method known as "Bootstrap Aggregation". Each sample is trained independently producing results. The final decision is based on majority voting where the results of all models merged which is termed as aggregation description.

3.2 Decision Tree Algorithm

In a "decision tree" algorithm, a root node provides the optimal split. It is represented by the predictor variable which helps to divide the dataset into two or more subsets. Then, the entropy or Gini Index is used to measure the homogeneity of a split. It measures the potential for information gain when splitting a node. Then, each node has a split of the predictor variable which yields the best homogeneity. Building a decision tree continues until all the nodes are pure, meaning all the nodes are homogeneous and belong to the same class. The process stops here, and the final tree is used to make predictions.

3.3 Naïve Bayes

The "Naïve Bayes" algorithm is a "supervised learning" method for classification problems. This method makes the prediction based on probability which we call it as "probabilistic classifier". The machine learning model is based on "Bayes Theorem" which calculates the likelihood of hypothesis of data.

3.4 Support Vector Machine (SVM)

"Support vector machine (SVM)" is utilized for classification and regression. "Hyperplane" in N-dimensional space is found using this algorithm which classifies the data points. Hyperplane is used to essentially divide the input features into two classes. It becomes challenging when the number of features exceed three.

3.5 K- Nearest Neighbor (KNN)

The "K-Nearest Neighbors (KNN)" algorithm is a simple, "supervised learning" algorithm used for classification and regression. The algorithm works by comparing an input sample to the k nearest neighbors in the training dataset. The k neighbors are determined based on their proximity to the input sample, typically using Euclidean distance as a measure of similarity. "KNN" is known as a non-parametric algorithm because it does not require assumptions about the underlying probability distribution of the data, making it a popular choice for many classification and regression problems.

4 Performance of Machine Learning for Crop Recommendation

For our work, we have acquired a dataset as CSV file from Kaggle [5]. The dataset contains eight columns comprising of 2200 records, and 17,600 fields. The column contain "N" which is the amount of "Nitrogen" content in soil, "P" which is the amount of Phosphorus content in soil, "K" which is the amount of "Potassium" content in soil, temperature which is in degree Celsius, humidity which is the relative humidity in %," "PH" value of the soil, rainfall which is rainfall in mm and label which is the classification's name. In each record, the datatype for each column is shown in Table 1. In the Crop Recommendation dataset, we have 22 crops which are as follows: "Rice, Maize, Jute, Cotton, Coconut, Papaya, Orange, Apple, Muskmelon, Watermelon, Grapes, Mango, Banana, Pomegranate, Lentil, Blackgame, Mung Bean, Moth Beans, Pigeon Peas, Kidney Beans, Chickpea, Coffee".

4.1 Data Visualization and Analysis

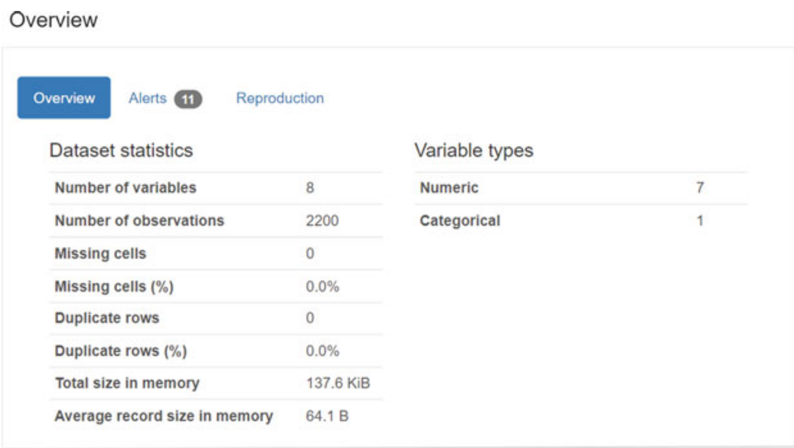
For visualizing the dataset content and characteristic, we used Python's Pandas profiling library that automates exploratory data analysis. It generates a dataset profile report that provides valuable insights. With the profile report, we can know which variables to use and which ones to drop. The following is some illustration and discussion of the generated report about the dataset.

Table 1 The datatype of the dataset's features

–N (Nitrogen) –P (Phosphorous) –K (Potassium)	–Temperature –Humidity –PH –Rainfall	Label
int64	float64	STRING

The first generator is the overview, which shows the information about the statistics of the dataset. As shown in Fig. 1, the number of columns is 8, the number of rows is 2200, missing values are 0 etc.

Figure 2 shows the features that strongly affect the label (crop) which are P, K, humidity, and rainfall. It also presents feature N which mostly contains zero since it has 27 rows but this does not affect the dataset.



Overview

Overview | Alerts 11 | Reproduction

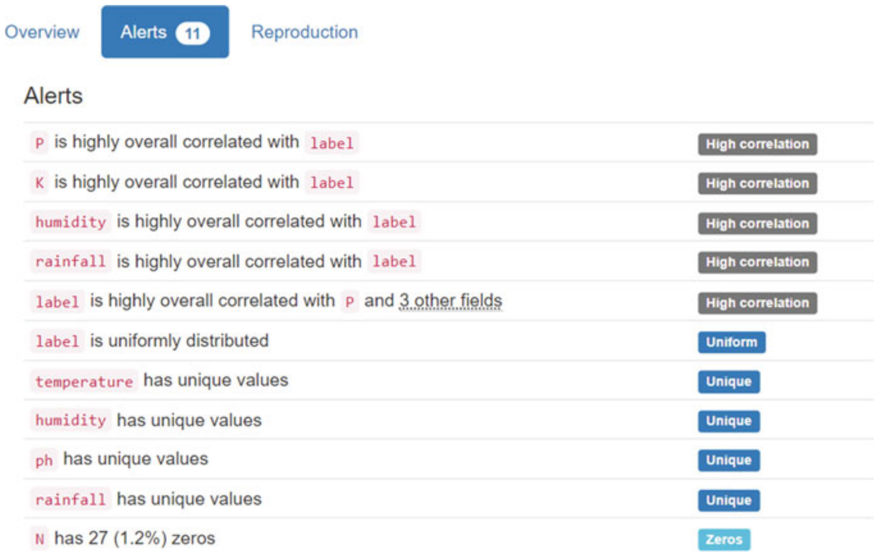
Dataset statistics

Number of variables	8
Number of observations	2200
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	137.6 KiB
Average record size in memory	64.1 B

Variable types

Numeric	7
Categorical	1

Fig. 1 Detailed overview of dataset



Overview | Alerts 11 | Reproduction

Alerts

P is highly overall correlated with label	High correlation
K is highly overall correlated with label	High correlation
humidity is highly overall correlated with label	High correlation
rainfall is highly overall correlated with label	High correlation
label is highly overall correlated with P and 3 other fields	High correlation
label is uniformly distributed	Uniform
temperature has unique values	Unique
humidity has unique values	Unique
ph has unique values	Unique
rainfall has unique values	Unique
N has 27 (1.2%) zeros	Zeros

Fig. 2 Dataset features affecting crop

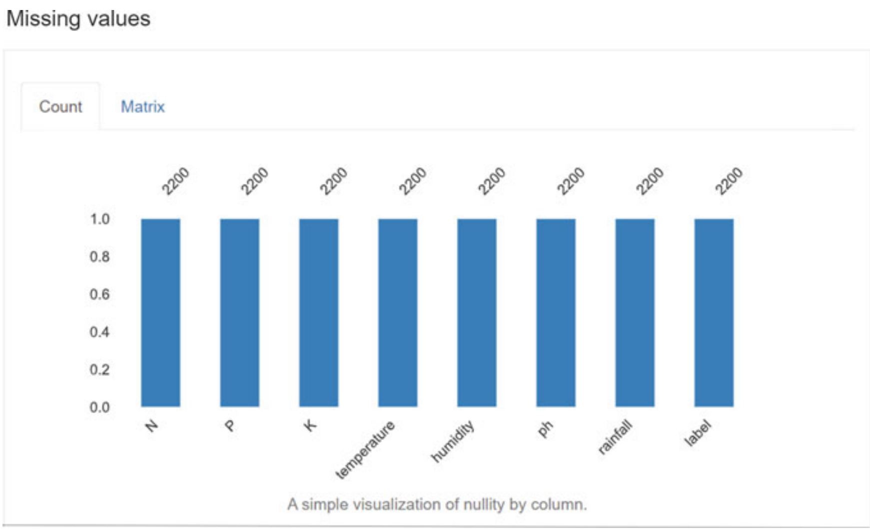


Fig. 3 Missing values view

The missing values view shows how many missing values are in each feature, as in Fig. 3. The dataset has 2200 rows which means no missing values in our dataset.

Figure 4 demonstrates the correlation view which illustrates the relationship between all features together, a positive number indicates a strong relationship, while a negative number indicates a light relationship.

Now based on visualization of dataset with correlation of features, there is need to evaluate the machine learning models. So, toward this, dataset is split into training and testing. The performance of machine learning is enhanced by splitting into training and testing which is a basic step of data preprocessing. We take 80% of dataset for training and remaining 20% for testing.

After training the machine-learning model using a trained set of data, it is important to use evaluation metrics to measure the performance of the ML model. For a different set of machine learning algorithms, there exist many evaluation measures. We will discuss only the Classification Evaluation Matrix since our selected dataset ‘Crop Recommendation’ belongs to this category. There are different approaches for this type of evaluation matrix [6] such as:

- “Confusion matrix
- Accuracy
- Precision
- Recall
- Specificity
- F1 Score”

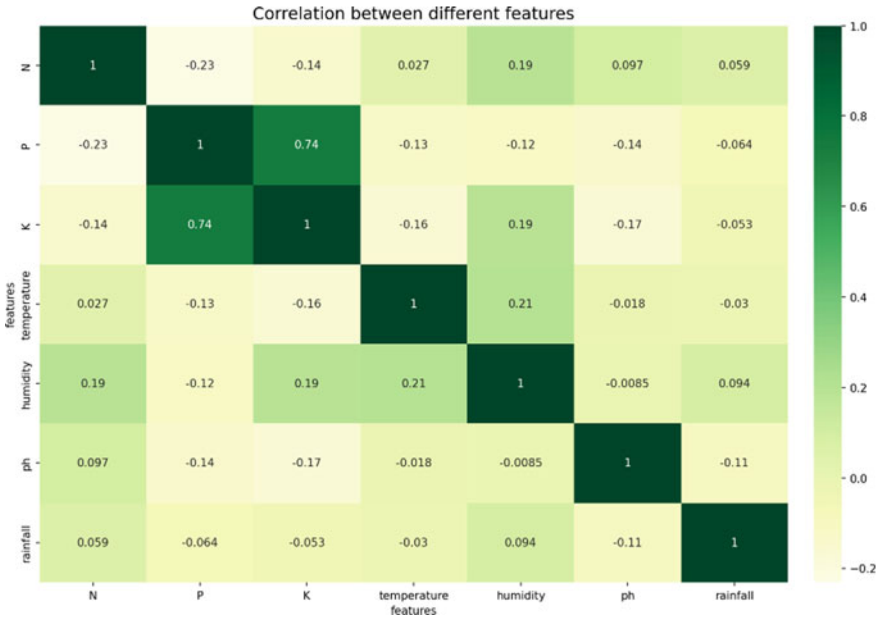


Fig. 4 Correlation view

5 Results and Discussions

We have evaluated the four machine learning models which are "SVC, Naïve Bayes, Decision tree, Random Forest, and KNN" Algorithm toward crop recommendation of different crops based on features correlated. The models have been evaluated in terms of precision, recall, F1 score and accuracy. In addition, the accuracy of model towards training and testing tabulated too in Tables 2 and 3, respectively. Figure 5 shows the model's performance.

Table 2 Comparison between the performance scores of the modes

Models	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
SVC	98.36	98.36	98.63	98.40
Decision tree	98.18	98.18	98.25	98.20
Random forest	99.45	99.45	99.60	99.50
KNN	98.00	98.14	98.10	98.10
Naïve Bayes	99.09	99.10	99.30	99.15

Table 3 Models' training and testing accuracy

Model	Train accuracy (%)	Test accuracy (%)
SVC	97.88	98.36
Decision tree	100	98.18
Random forest	100	99.45
KNN	98.79	98.00
Naïve bayes	99.58	99.09

6 Conclusion and Future Work

This chapter validated five machine-learning algorithms which are "SVC, Decision Tree, Random Forest, KNN, and Naïve Bayes" toward crop recommendation which resulted in Random forest with highest accuracy of 99.45%, followed by Naïve Bayes, SVC, Decision Tree, and KNN, respectively. We chose the best model, which is Random Forest based on accurate results because the Crop Recommendation dataset is balanced which calculates the ratio of a correct prediction made by the machine-learning model over the total number of instances evaluated. Moreover, we tested the overfitting and underfitting and we observe that the Random Forest model is generalized.

In future, this work can explore different approaches such as the more sophisticated machine-learning algorithms like deep neural with optimization for better accuracy of the recommendation system. The system can be integrated with crop yield prediction for providing better productivity. Additionally, the system could introduce larger datasets with various features such as market prices, weather forecasts, etc. to improve the system's predictive accuracy.

Funding Statement
This work is supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Grant No. 3382).

References

1. Pudumalar S, Ramanujam E, Rajashree RH, Kavya C, Kiruthika T, Nisha J (2017) Crop recommendation system for precision agriculture. In: 2016 Eighth international conference on advanced computing (ICoAC), Chennai, India, 2017, pp 32–36. https://doi.org/10.1109/ICoAC.2017.7951740
2. Kulkarni NH, Srinivasan GN, Sagar BM, Cauvery NK (2018) Improving crop productivity through a crop recommendation system using ensembling technique. In: 2018 3rd International conference on computational systems and information technology for sustainable solutions (CSITSS), Bengaluru, India, 2018, pp 114–119. https://doi.org/10.1109/CSITSS.2018.8768790
3. Kumar A, Sarkar S, Pradhan C (2019) Recommendation system for crop identification and pest control technique in agriculture. In: 2019 International conference on communication and signal

processing (ICCSP), Chennai, India, 2019, pp 0185–0189. https://doi.org/10.1109/ICCSP.2019.8698099

4. Rajeswari AM, Anushiya AS, Fathima KSA, Priya SS, Mathumithaa N (2020) Fuzzy decision support system for recommendation of crop cultivation based on soil type. In: 2020 4th international conference on trends in electronics and informatics (ICOEI) (48184), Tirunelveli, India, pp 768–773. https://doi.org/10.1109/ICOEI48184.2020.9142899
5. Ingle A (2020) Crop recommendation dataset, Kaggle. Available at: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset (Accessed 11 Feb 2023)
6. "A Guide to Evaluation Metrics for Classification Models | Deepchecks", Deepchecks, 2022. [Online]. Available: https://deepchecks.com/a-guide-to-evaluation-metrics-for-classification-models/. [Accessed: 27 Apr 2022]

Exam Proctoring System

Arvind Vishnubhatla

Abstract An exam proctoring system through surveillance cameras is a software-based solution that utilizes surveillance cameras to monitor and manage exams. The system is designed to help institutions maintain academic integrity by preventing cheating and ensuring that students follow exam rules and guidelines.

Keywords Exam · Proctoring · System · Monitor · Rules

1 Introduction

An exam proctoring system through surveillance cameras is a software-based solution that utilizes surveillance cameras to monitor and manage exams [1]. The system is designed to help institutions maintain academic integrity by preventing cheating and ensuring that students follow exam rules and guidelines [2].

The system typically uses video cameras, audio recording devices, and facial recognition software to monitor students during exams [3]. The software is configured to detect and flag suspicious behavior, such as looking at notes, speaking with others, or accessing unauthorized websites [4].

The system also has features that allow proctors to remotely monitor students and intervene if necessary. For example, if the system detects suspicious behavior, the proctor can review the video footage in real-time and contact the student to stop the behavior [5].

Overall, an exam proctoring system through surveillance cameras provides an added layer of security for exams and helps institutions maintain academic integrity [6]. However, it is important to balance the benefits of the system with concerns around privacy and potential bias in monitoring [7] (Fig. 1).

Assign each student a unique bar code, which can be printed on their exam ticket or ID card. Create a database that contains the bar codes of all the students in the

A. Vishnubhatla (✉)
Electronics and Communications Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India
e-mail: vainfo66@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_30

Fig. 1 An exam proctoring system

class along with their respective seating positions [8]. Print the seating arrangement chart and assign each seat a unique bar code that matches the student's bar code [9].

On the day of the exam, students will scan their bar code at the entrance of the exam hall using a barcode scanner. This will automatically display their assigned seat number on the screen [10].

Once seated, students will receive their exam booklets, which will also contain a unique bar code that matches their seat number.

During the exam, students will write their answers in the booklets and must ensure that they match the bar code on their answer book with the one on their seat.

After the exam, the answer books will be collected and scanned using a barcode scanner to match the bar code on the answer book with the one on the seating chart. This will help ensure that each student's exam is correctly marked and graded (Fig. 2).

2 Project Methodology

Determine the specific requirements for your proctoring system, such as the number and placement of cameras, the features you want to monitor (e.g. student ID verification, monitoring for cheating behavior, etc.), and the level of automation you desire.

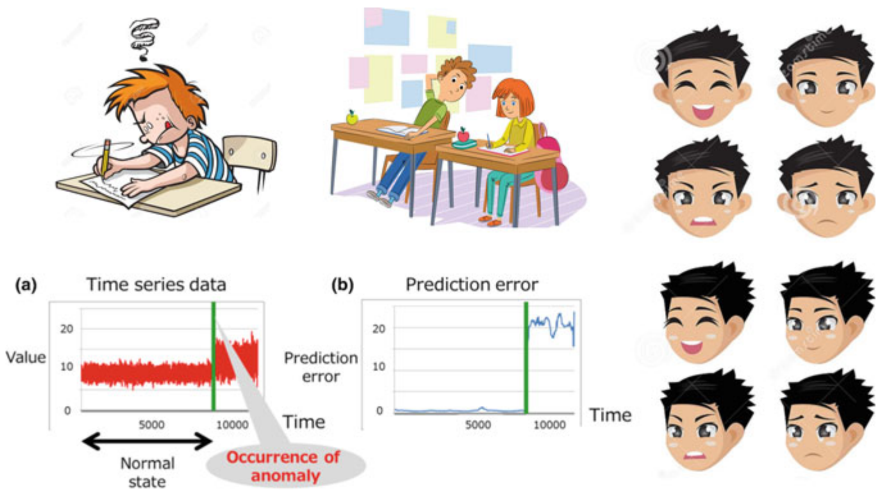


Fig.2 Anamoly detection system

Choose the appropriate hardware and software for your system. This could include cameras, microphones, speakers, and computers capable of running the necessary software.

Develop a set of rules or criteria for identifying behavior that indicates cheating, such as looking off-screen frequently, talking to someone off-screen, or searching for answers on the internet.

Write code to detect these behaviors using machine learning algorithms, image processing, or other techniques. You may want to train your system on sample data to improve accuracy.

Create a user interface for the system that allows instructors to monitor student behavior in real-time and flag any suspicious activity for further review.

Test and debug your system, and make any necessary adjustments to improve accuracy and performance.

3 Implementation

Capture video: Use Python libraries like OpenCV or Pygame to capture video from the camera in real-time.

Face detection: Use a face detection algorithm to detect the presence of faces in the video stream. You can use pre-trained models like Haar cascades or deep learning-based models like SSD or YOLOv5.

Face recognition: Once a face is detected, use a face recognition algorithm to match it against a pre-defined database of authorized students. You can use pretrained models like OpenFace or FaceNet for face recognition.

Proctoring rules: Define proctoring rules, such as how many times a student can look away from the screen, how many faces can be detected in the video stream, and so on.

Violation detection: Use the rules defined in the previous step to detect violations in real-time. You can use techniques like object detection or eye tracking to detect if a student is looking away from the screen or if there are multiple faces in the video stream.

Alert system: Once a violation is detected, trigger an alert system to notify the proctor or the concerned authority. You can use email, SMS, or other messaging services to notify the relevant parties.

Logging: Keep a log of all the video streams captured, face detections, face recognitions, and violation detections for future reference.

```
/* Bar Code Scanner*/
  #include <stdio.h>

// Define the possible states of the barcode scanner
typedef enum {
   IDLE,
   SCANNING,
   COMPLETE
} State;

int main() {
   State state = IDLE;
   char input;

   printf("Barcode Scanner\n");

   // Loop until the barcode is complete
   while (state != COMPLETE) {
      switch (state) {
         case IDLE:
            printf("Ready to scan\n");
            scanf(" %c", &input);
```

```
                if (input == '*') {
                    state = SCANNING;
                }
                break;

            case SCANNING:
                printf("Scanning barcode\n");
                scanf(" %c", &input);

                if (input == '*') {
                    state = COMPLETE;
                }
                break;
        }
    }

    printf("Barcode scanned\n");

    return 0;
}

/* Proctoring Exam through camera*/

#include <stdio.h>

// Define the states of the finite state machine
typedef enum {
  STATE_IDLE,
  STATE_CHECKING,
  STATE_MONITORING,
```

```
  STATE_COMPLETE
} state_t;

// Define the events that can trigger state transitions
typedef enum {
  EVENT_START_EXAM,
  EVENT_FACE_DETECTED,
  EVENT_FACE_LOST,
  EVENT_MOVEMENT_DETECTED,
  EVENT_EXAM_COMPLETE
} event_t;

// Define the state machine structure
typedef struct {
  state_t state;
  void (*action)();
  state_t (*transition)(event_t);
} fsm_t;

// Define the actions that can be taken in each state
void idle_action() {
  printf("Waiting for exam to start...\n");
}

void checking_action() {
  printf("Checking exam taker's identity...\n");
}

void monitoring_action() {
  printf("Monitoring exam taker for suspicious behavior...\n");
}

void complete_action() {
  printf("Exam complete.\n");
}

// Define the state transitions based on events
state_t idle_transition(event_t event) {
  if (event == EVENT_START_EXAM) {
    return STATE_CHECKING;
  }
  return STATE_IDLE;
}

state_t checking_transition(event_t event) {
  if (event == EVENT_FACE_DETECTED) {
    return STATE_MONITORING;
  } else if (event == EVENT_EXAM_COMPLETE) {
    return STATE_COMPLETE;
  }
  return STATE_CHECKING;
}
```

```
state_t monitoring_transition(event_t event) {
  if (event == EVENT_FACE_LOST || event == EVENT_MOVEMENT_DETECTED) {
    return STATE_CHECKING;
  } else if (event == EVENT_EXAM_COMPLETE) {
    return STATE_COMPLETE;
  }
  return STATE_MONITORING;
}

// Define the state machine
fsm_t proctor_fsm[] = {
  {STATE_IDLE, idle_action, idle_transition},
  {STATE_CHECKING, checking_action, checking_transition},
  {STATE_MONITORING, monitoring_action, monitoring_transition},
  {STATE_COMPLETE, complete_action, NULL}
};

int main() {
  // Initialize the state machine
  state_t current_state = STATE_IDLE;
  int event;

  // Loop until the exam is complete
  while (current_state != STATE_COMPLETE) {
    // Execute the action for the current state
    proctor_fsm[current_state].action();

    // Wait for an event to occur
    printf("Enter event: ");
    scanf("%d", &event);

    // Transition to the next state based on the event
    current_state = proctor_fsm[current_state].transition(event);
  }

  return 0;
```

4 Results and Discussion

Statistical anomaly detection is a machine learning technique that involves analyzing large amounts of data and identifying patterns that deviate from the expected norm. In the context of an exam proctoring system, this technique can be used to identify any behavior that is not consistent with the expected behavior of an examinee during an exam (Fig. 3).

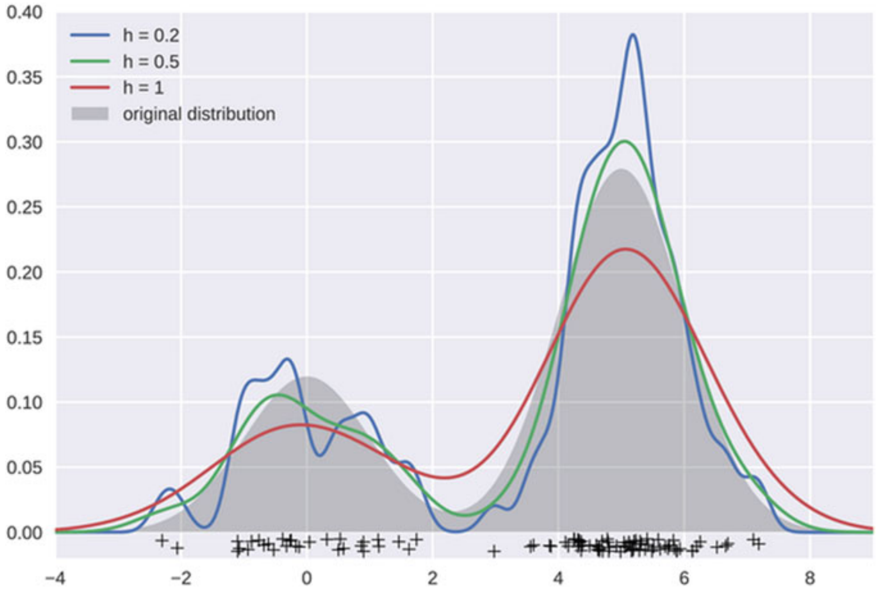


Fig. 3 Anamoly classification

5 Conclusion

In this futuristic analysis, we will explore how exam proctoring systems are likely to evolve in the coming years with increased use of AI and Machine Learning, integration with Learning Management Systems, improved Accessibility, greater Use of Biometric Data, and increased Privacy Concerns.

In conclusion, exam proctoring systems are likely to become more sophisticated in the coming years, with increased use of AI and machine learning, integration with LMS, improved accessibility, greater use of biometric data, and increased privacy concerns.

References

1. Liu R, Li X, Li X, Li X (2020) Design and implementation of an exam proctoring system based on surveillance cameras. In: 2020 International conference on artificial intelligence and big data (ICAIBD), 2020, pp 279–284
2. Ahmadi A, Khosravi MA, Ghorbani A (2021) A comprehensive review of exam proctoring systems: challenges, technologies, and solutions. Comput Secur 104:101–123
3. Liu Y, Zeng Z, Wang Y (2020) Design and implementation of an exam proctoring system based on computer vision technology. In: 2020 IEEE 7th International conference on industrial engineering and applications (ICIEA), 2020, pp 167–172

4. Adnane N, Rachidi A, Belkhouche FA (2020) An exam proctoring system based on multi-agent systems and blockchain technology. J Ambient Intell Humaniz Comput 11(7):2983–2995
5. Wang J, Sun Y, Li X (2021) An exam proctoring system based on facial recognition and machine learning. In: 2021 IEEE 2nd International conference on artificial intelligence and robotics technology (ART), 2021, pp 177–182
6. Ma S, Chen C, Zhang X (2020) An exam proctoring system based on deep learning and biometric authentication. In: 2020 IEEE International conference on big data and smart computing (BigComp), 2020, pp 257–260
7. Li C, Ma Y, Li Z (2020) Design and implementation of an online exam proctoring system based on surveillance cameras. In: 2020 2nd International conference on computing, electronics and communications engineering (iCCECE), 2020, pp 1–6
8. Chen H, Liu Y, Chen W (2021) An exam proctoring system based on video analysis and big data technology. In: 2021 IEEE 3rd International conference on advances in electrical, computer and communication engineering (ICAECCE), 2021, pp 1–5
9. Zhang Y, Zhang L, Li X (2020) An exam proctoring system based on face recognition and behavior analysis. In: 2020 International conference on computer science and application engineering (CSAE), 2020, pp 166–171
10. .Wu J, Xu J, Zheng Z (2021) Design and implementation of an exam proctoring system based on artificial intelligence and blockchain technology. In: 2021 3rd International conference on computer communication and the internet (ICCCI), 2021, pp 32–37

Object Detection Using TensorFlow

Anushka Sharma, Trapti Mishra, Jyoti Kukade, Aadya Golwalkar, and Hrithik Tomar

Abstract What do you think can redefine clarity to the world? By bringing the potential of AI into our lives, we can unlock the power of object detection, which allows computer systems to see their environment in visual images. Object detection is a rapidly developing field with many potential future directions and there is still much to explore and discover in this exciting area of research. It is an essential aspect of computer vision that has numerous practical applications, including autonomous driving, robotics, and surveillance. In this study, we present an object detection framework that combines real-time performance with high detection accuracy. We conducted a comprehensive analysis of various deep learning models for object detection within the TensorFlow framework. The TensorFlow library offers a variety of machine learning models through its TensorFlow Model Gardens, which can be used for further studies and experimentation purposes by instructing and re-instructing their behaviour. The input image is passed through the SSD architecture, which is used to find objects based on deep neural networks for accurate identification. This system is fully implemented using the Python programming language because Python is a multitasking tool and needs relatively fewer lines of code to perform operations. As deep learning models and AI technologies continue to progress, there is a huge potential to enhance the precision, capability, and performance even more of real-time object detection methods. This research paper presents an object detection approach using the TensorFlow framework and demonstrates its effectiveness and potential for practical applications.

Keywords Object Detection · TensorFlow APIs · Convolutional neural networks · ResNet · CUDA · Pre-trained models

A. Sharma (✉) · T. Mishra · J. Kukade · A. Golwalkar · H. Tomar
Medi-Caps University, Indore, India
e-mail: sharmanushka4@gmail.com

T. Mishra
e-mail: trapti.mishra@medicaps.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_31

1 Introduction

Over many decades, numerous reports have highlighted how AI technology has transformed the landscape of modern businesses and not just that it has significantly influenced the course of our lives. In recent years, an increasing number of product and service-based companies have been implementing AI- powered tools to improve resource management, optimize profits, increase customer engagement, and improve customer loyalty. Experts believe that the biggest reason why AI operations are coming to consummation is due to their exponential benefactions in calculating power and generation of data and as advanced technologies continue to evolve critical tasks like object discovery have also become easier than it was ever imagined. Within the field of computer vision, object detection is a crucial technique that allows software to identify and track objects within images and videos. This method provides the unique ability to distinguish between different types of objects, such as people, tables, and chairs, and pinpoint their precise locations within the given visual data. By outlining the object's location with a box, the location is indicated. [1] Turning the clock 20 years back, we would witness that the majority of the early object detection algorithms were constructed using handcrafted characteristics. Among them are: **The Viola Jones Detectors**, created in 2001, use sliding windows to recognize human faces in real time. **HOG Detector**, introduced in 2005, in the same way that sliding windows can recognize objects of varied sizes while constantly rescaling the input image while retaining the size of a detection window, Hog uses blocks. As an improvement to the HOG detector, the **Deformable Part-based Model (DPM)** was proposed in 2008. After this, barriers to object detection were dismantled in 2014 with the introduction of CNN features (R-CNN), which begin with the selective search-based extraction of a collection of object suggestions, before being rescaled to a fixed size image and put into a pre-trained CNN model to extract features. Faster R-CNN, the first end-to-end and near real-time deep learning object detector, eliminates the selective search technique and, in contrast to an R-CNN model, uses the full image as the CNN input for feature extract ion instead of just each proposed region. Following all of these are others like the **Single Shot MultiBox Detector (SSD)**, which was developed in 2015. It is a one-step object detector, similar to YOLO, which pioneered the multi-reference and multi-resolution detection approaches to increase a one-stage detector's detection accuracy, particularly for some small objects. **You Only Look Once (YOLO)**, which learns on complete images and directly optimizes detection performance, can predict several bounding boxes at once using a single CNN. The versions after YOLO are **YOLO V2, YOLO V3, and YOLO V4**.

The YOLO and TensorFlow object detection API are both well-known frameworks for object detection, but TensorFlow has more advantages because it enables you to quickly swap out various model architectures and configurations, making it a more versatile framework for experimenting with different object detection models. TensorFlow also offers a full pipeline for training and evaluating object detection models, as well as a ready-made pipeline to train and evaluate them on the fly. This essay initially reviews a number of earlier, related studies on object deduction before describing the current study's methodology and main points. A computer vision technique called TensorFlow object detection finds, tracks, and detects an item in a still image or video. By identifying the models' workings, the method gives us a deeper comprehension of the image or video by identifying things. In essence, TensorFlow object detection enables us to train robots to better comprehend human actions and utilize the knowledge to perform at optimal levels. In order to identify recognized objects from all instances in an image or video, object detection TensorFlow uses computer vision.

2 Related Work

The model proposed in this research introduced the concept of Intern-Image, where they claim that it can effectively detect and classify the given object and can also perform tasks like semantic segmentation. Even for the larger networks, it can give performance which is better than the models which have been used up till now where they have data in bulk and get trained on the same. Hence, they claim that this approach can be a comparatively better technique for large scale object detection [2].

Model here presents a comprehensive review of deep learning and convolutional neural network architecture. The research basically demonstrates how the improvements of CNN techniques are most effective for image applications in computer vision [3].

In another study, deep learning neural networks are used to develop an algorithm for item detection from pictures. The model in the research is made by combining a multilayer neural network and an enhanced SSD technique and proposed that for both moving and still photographs, the algorithm performed well and has high accuracy [4].

In the following report, deep learning–based object detection model was proposed which can constructively complete tasks like general object detection, face detection and pedestrian detection and all of this was achieved by combining two effective techniques: efficient, threaded video streams using OpenCV and object detection with deep learning. These detection techniques basically combine object localization with classification where it takes the supplied image as input and produces a file with bounding boxes with category names at the top of each that are appropriate for the number of things included in the image.

In the following review, yet again a deep learning–based detection framework has been proposed for the completion of different tasks related to the object detection and classification not just this it also claimed to resolve the issues like clutter and resolution problems, basically by working on object detection pipelines which act as a base for all of that was mentioned above [5].

In the following study, detection and tracking was done by using a SSD algorithm. The model was proposed to fulfil the focus that is, to detect the objects in the real-time scenarios. Tracking of the sequences in any particular video instantaneously made possible with the help of the following algorithm. For the training object, this model displayed great detection and tracking performance [6].

In this research, following model was built by keeping images in mind that are not in direct range and those which are far away from the normal range of images, considering the fact that many images that are not in direct range can be badly contrasted and can lack disparity and for this the model made, deep convolutional neural networks (CNN) to run parallel with the network topology to make its performance in such scenarios effectively good [7].

In the following research, one of the deep learning frameworks was proposed, the deep convolutional neural network, and was implemented using OpenCL. This particular framework proved to be the most effective model because as of its constructive claim of being able to operate on many different portable computing devices and that too with lower consumption of power supply [8].

In this study, YOLO model is proposed that is both straightforward to build and can be directly used on entire images. YOLO has set the benchmark by getting directly trained on the entire model which is very unlike to the classifier-based models [9].

3 Methodology

This section outlines our suggested method for leveraging convolutional neural networks and deep learning to quickly identify items in pictures. The convolution neural network, a promising strategy, is utilized in this paper. Traditional neural networks are improved by CNN, which maintains the spatial dimension. Researchers are particularly interested in CNN because it has outperformed a number of different frameworks for image classification, object detection, etc. In previous architectures, in order to lower the error rate, every succeeding successful architecture adds more layers to a deep neural network. This works for less number of layers, but when we increase the number of layers, the accuracy falls for such networks instead of going up, the two basic reasons behind this are as follows: vanishing/exploding gradients and degradation problems. However, normalized initialization and intermediate normalization layers overcome the vanishing/exploding gradients problem, allowing networks with tens of layers to begin convergent stochastic gradient descent (SGD) [10] using back propagation, this concept is known as residual blocks. Nevertheless, they see the degradation issue as the deeper networks begin to converge. Accuracy becomes saturated as the network depth rises and subsequently quickly declines. The addition of the extra layers with an identity mapping is one way to assist in removing this degrading issue. These "identity mapping" or "skip connections" are introduced by ResNet. The skip connection bypasses some levels in between to link layer activation to subsequent layers. This creates a residual block. These residual blocks are stacked to create ResNets. The concept is to add new layers to the network as residual blocks rather than simply stacking them on top of each other (having identity mappings). In this method, underlying mapping is modified and had the non-linear layers pick up the mapping in order to get the results:

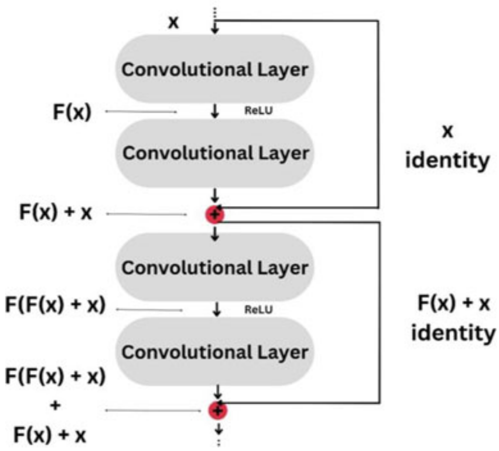


Fig. 1 Residual Learning, image adapted from [10]

$$F(x) := H(X) - X \tag{1}$$

The initial mapping is reformulated to instead of the conventional mapping $H(x)$ [10]:

$$F(X) + X \tag{2}$$

Another advantage of using these shortcut connections is that they are used as identity mappings, and they do not introduce any more parameters or raise computational complexity. But they make sure the deeper networks function just as well as their shallower counterparts (Fig. 1).

3.1 Proposed Algorithm and Base Model

The algorithm proposed in the model is the single shot detector (SSD) algorithm which is a popular object detection method that uses a deep convolutional neural network to detect and locate objects within an image. Unlike other object detection methods that require multiple passes through an image, SSD performs object detection in a single shot. SSD uses a combination of 1 × 1 and 3 × 3 filters to generate feature maps for object detection. The algorithm also employs the MultiBox technique [11] that utilizes predefined bounding boxes with fixed dimensions placed over ground truth regions in an image. By dividing the image into grids of varying sizes, SSD generates feature maps that can effectively detect both large and small objects. Each predicted bounding box is assigned a confidence score for the label prediction, and the box with the highest score is selected. SSD is a highly competent object detection algorithm that is pre-trained to compute priors and uses convolutional layers of different cell sizes to detect objects of varying sizes.

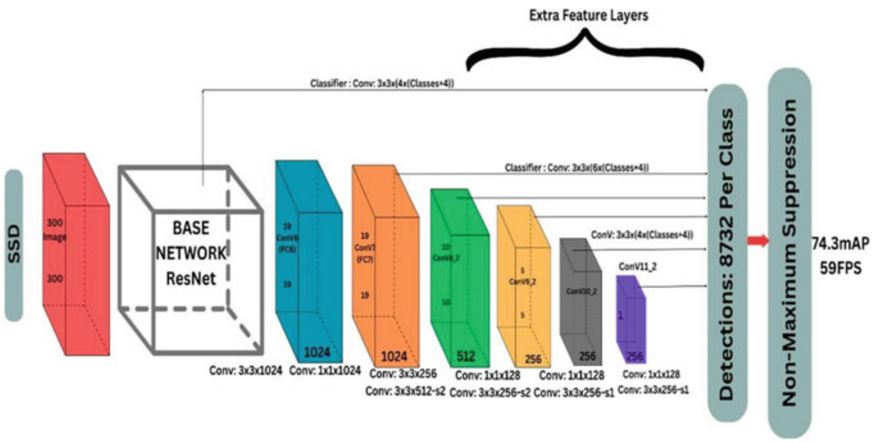


Fig. 2 Architecture of SSD. Image adapted from [16]

The SSD is widely used in various applications, including but not limited to pose estimation [12], face detection [13], and railway traffic detection [14]. It is a cutting-edge object detector that addresses the problem of slow inference by integrating detection and localization of objects in images in a single network pass. The base network of SSDs, which accounts for 80% of the processing time, has a significant impact on the inference speed, similar to faster R-CNN [15].

The algorithm subjected to this model is using base model ResNet, short form of Residual Network, and it is a deep learning architecture that has significantly improved the performance of deep neural networks in image recognition tasks. This is a model that make use of residual connections, or skip connections, between layers in a neural network by which network acquires the residual mapping. By adding the residual mapping to the output of a layer, the network can skip over the redundant computation of the layer which in turn reduces the vanishing gradient problem and enables the training of very deep neural networks. Residual learning is depicted in Fig. 1 (Fig. 2).

3.2 *TensorFlow*

TensorFlow is a potent open-source library for machine learning and deep learning. It uses dataflow graphs to represent mathematical operations and tensors, which are multidimensional array of data. It makes it possible for programmers to create and launch ML-powered applications. The TensorFlow library offers a variety of machine learning models through its TensorFlow Model Gardens, which can be used for further studies and experimentation purposes, and we need these models so that we can try out different recognition patterns for image detection by instructing and re-instructing their behaviour. These models can be run on graphical processing units (GPUs) or tensor processing units (TPUs), but the majority of the models in the Model Garden are trained on large datasets using TPUs. To make these models work on GPUs, we utilize an application programming interface called CUDA.

CUDA CNN is a software library developed to accelerate convolutional neural network (CNN) computations. CNNs are a type of neural network that are commonly used to recognize and detect objects in images. This library is designed to run on parallel architecture which can perform many computations simultaneously due to which models trained fast as compared to running on CPUs alone. It is compatible with a variety of deep learning frameworks such as TensorFlow and PyTorch. TensorFlow provides a range of object detection models that have already been trained and optimized to work seamlessly with the TensorFlow object detection API; these models are readily available for use through the TensorFlow Detection Model Zoo. You can either download and use the models directly [17], or you can train your own unique object identification model using the TensorFlow Object Detection API, which additionally assists in the execution of the models present in the TensorFlow Detection Model Zoo in straightforward manner and also employ them without any complexity by simply instructing them. Hence, these are used for variety of different tasks, but here we have used this for detecting objects from images. In addition to this, these TensorFlow APIs also provide access to various detection models that can be pre-configured on the COCO datasets, stands for common objects in context, a high-quality visual datasets that we have used for pre-modelling of our framework but apart from this, models can also be pre-build on other datasets like Kitti and Open Images as they are also available through these APIs itself. These pre-trained models are capable of making accurate predictions right out of the gate. Therefore, these datasets can be utilized for the training purposes and also in the designing of the algorithms for object detection and classification.

3.3 Experimental Setup

This paper suggests an object detection method that is used to detect the object by running the proposed model. To implement the suggested system, we have used the Python programming language and the OpenCV 2.4 library. In addition to this, we have used TensorFlow-GPU 2.9 and also created a network using a variety of libraries. For preparing the data, we have used labelImg for labelling images and producing annotations and used this to train the model. We have used Protobuf module that enables us to create classes for machine learning models from a protocol buffer text file.

Your pictures have been annotated for object detection, which entails that regions for any things of interest that might be present in your datasets have been manually specified as bounding boxes, and ground truth labels have been assigned to each box. For data transformation, we need to convert input data into TFRecord data file format. A series of binary records can be stored in the TFRecord format, which is a straightforward format. A Label Map.pbtxt file has been prepared for label map generation. It connects some integer values to labels. This file is required for training and the detection by the TensorFlow Object Detection API. Our datasets has been divided into the required training and testing subsets, and we have generated appropriate annotations. In the experimental configuration, the TensorFlow directory, SSD ResNet101 V1 FPN feature extractor, TensorFlow object detection API, and anaconda virtual environment are utilized.

The pre-trained model that we had previously exported from the TensorFlow Model Zoo will now have its configuration process modified, the configuration then will be utilized for the retraining of the model, alongside this, a process will be running in continuation which are periodically checking the overall performance of the model by making use of the checkpoint files for the analysis of the execution of the model being developed and the precision of the output generated. Once the computations are complete, the TensorFlow event files will provide access to the results. By analysing the metrics, the new model can be evaluated and exported for object detection purposes.

4 Conclusion

In this research, an algorithm for identifying objects in pictures is developed utilizing deep learning neural networks. The research uses an improved SSD algorithm where normalized initialization and intermediate normalization layers overcome the vanishing/exploding gradients problem, allowing networks with tens of layers to begin convergent stochastic gradient descent (SGD). However, they see the degradation issue as the deeper networks begin to converge, so ResNet introduces "Identity mapping" or "skip connections" to link layer activations to subsequent layers. These residual blocks are stacked to create ResNets, where the underlying mapping is modified and non-linear layers pick up the mapping in order to get the results. Our algorithm is best suited for still images. The proposed model's degree of accuracy exceeds 82.6%. These convolutional neural networks take the image's features and use feature mapping to determine the class label. Our method's main goal is to select default boxes with the best aspect ratios possible in order to enhance SSD's object detection process (Fig. 3).

Fig. 3 Results of objects detected by TFOD ResNet algorithm

References

1. Zou Z, Chen K, Shi Z, Guo Y, Ye J (2023) Object detection in 20 years: a survey. Proc IEEE 111(3):257–276. https://doi.org/10.1109/JPROC.2023.3238524
2. Wang W, Dai J, Chen Z, Huang Z, Li Z, Zhu X, Hu X, Lu T, Lu L, Li H et al (2022) Internimage: exploring large-scale vision foundation models with deformable convolutions. arXiv preprint arXiv:2211.05778
3. Jawale P, Chaudhary HP, Journal I (n.d.) Real-time object detection using tensorFlow. IRJET 07, 2020
4. Kumar A, Zhang ZJ, Lyu H (2020) Object detection in real time based on improved sin- gle shot multi-box detector algorithm. EURASIP J Wirel Commun Netw 204. https://doi.org/10.1186/s13638-020-01826-x
5. Zhao Z-Q, Zheng P, Xu S-T, Wu X (2019) Object detection with deep learning: a review. IEEE Trans Neural Netw Learn Syst 30(11):3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865
6. Chandan G, Jain A, Jain H, Mohana (2018) Real time object detection and tracking using deep learning and OpenCV. In: 2018 International conference on inventive research in computing applications (ICIRCA), Coimbatore, India, 2018, pp 1305–1308. https://doi.org/10.1109/ICIRCA.2018.8597266
7. Alpaydın G (2018) An adaptive deep neural network for detection, recognition of objects with long range auto surveillance. In: 2018 IEEE 12th International conference on semantic computing (ICSC), Laguna Hills, CA, USA, 2018, pp 316–317. https://doi.org/10.1109/ICSC.2018.00064
8. Luo Y, Li S, Sun K, Renteria R, Choi K (2017) Implementation of deep learning neural network for real-time object recognition in OpenCL framework. In: 2017 International SoC design

conference (ISOCC), Seoul, Korea (South), 2017, pp 298–299.https://doi.org/10.1109/ISOCC.2017.8368905
9. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: 2016 IEEE Conference on computer vision and pattern recogni tion (CVPR), Las Vegas, NV, USA, 2016, pp 779–788. https://doi.org/10.1109/CVPR.2016.91
10. He K, Zhang X, Ren S, Sun J (2016) Deep Residual Learning for Image Recognition. In: IEEE conference on computer vision and pattern recognition (CVPR), Las Vegas, NV, USA, 2016, pp 770–778. https://doi.org/10.1109/CVPR.2016.90
11. Szegedy C, Reed S, Erhan D, Anguelov D, Ioffe S (2014) Scalable, high-quality object detection. arXiv preprint arXiv:1412.1441. https://arxiv.org/abs/1412.1441
12. Poirson P, Ammirato P, Fu CY, Liu W, Kosecka J, Berg (2016) Fast single shot detec tion and pose estimation. In: Fourth international conference on 3D vision (3DV). IEEE, Stanford, CA, USA
13. Chen JC, Lin WA, Zheng J, Chellappa R (2018) A real-time multi-task single shot face detector. In: 25th IEEE international conference on image processing (ICIP). IEEE, Athens, Greece, pp 176–180
14. Ye T, Zhang Z, Zhang X, Zhou F (2020) Autonomous railway traffic object detection using feature-enhanced single-shot detector. IEEE Access 8:145182–145193. https://doi.org/10.1109/ACCESS.2020.3015251
15. Jee G, Harshvardhan GM, Gourisaria MK (2021) Juxtaposing inference capabilities of deep neural models over posteroanterior chest radiographs facilitating COVID-19 detection. J Interdisc Math 24(2):299–325
16. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg (2016) SSD: single shot multibox detector. In: European conference on computer vision. Springer, Cham, Amsterdam, The Netherlands, pp 21–37
17. Joseph (2022) The tensorflow detection model zoo—reason town

Road Lane Line Detection

Shreesh Srivastava, Bhavya Sheth, Pratham Shah, and Darshan Ingle

Abstract With the growth of urban traffic, ensuring traffic safety is becoming an increasingly significant factor. There should be some advancements in technology to achieve the goal of traffic safety. Various high-tech research and studies are going on in this domain considering road and people safety. The project "Road Lane Line Detection" aims to address these issues by detecting the lane lines on the road and providing directions to drivers based on lane markings. The project uses computer vision techniques to detect the lane lines and works in real time on video streams. The system can find lane lines in different weather conditions and during nighttime. Real-time visual feedback is also provided to the driver by the system, indicating the direction of the detected road lane lines. The direction can be straight, left, or right, depending on the position of the lane lines. Hence, this project demonstrates the potential of computer vision techniques in strengthening road safety, thus providing real-time feedback to drivers based on lane markings. Self-driving cars can also benefit from this solution.

Keywords Traffic safety · Computer vision · Real time · Road · Lane · Line · Detection · Weather · Nighttime · Directions · Self-driving cars

1 Introduction

Diverging a car from a road lane increases the chances of accidents and other causes on the avenues. The prime motive behind the equipment of vehicles integrating with advanced tools is safety, lessening road accidents, and maximizing amenities and contentment for people and the driving experience. Most of them are achieved by identifying road lanes. Most of the previous research approaches detect the lane center line and not the exact left and right lane boundaries. In addition, they do not consider the various amenities provided by the system. This proposed system

S. Srivastava (✉) · B. Sheth · P. Shah · D. Ingle
Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra 400050, India
e-mail: shreeshsrvstv@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_32

primarily focuses on functions such as capturing road lanes, detecting switching of lanes which are bound to happen in near future, and also detecting under noisy conditions and nighttime [1]. Also, the road lane boundaries detection functionality includes the determination of the relative position between the vehicle and the road. The problems which are being solved by this proposed model are the following: Reducing accidents because of departing lanes, contentment to people sitting beside the driver, reducing driver fatigue, improving traffic safety, autonomous driving, providing directions while driving a vehicle and detecting lane lines during any weather conditions.

This article contains VII segments, Sect. 1 is the Introduction, Sect. 2 explains the Literature Review, Sect. 3 tells us about the Methodology or the proposed system, Sect. 4 depicts the Results obtained and finally, the Conclusion followed by Acknowledgement and References.

2 Literature Review

The entire process is broken down into two, pre-processing and post-processing [2]. Pre-processing which improves the quality of each and every frame of the image or video involves color selection [3], gray scaling, Gaussian smoothing and Canny edge detection. Post-processing which is a combination of the feature extraction stage [4], and the road lane line modeling stage involves the Region of Interest, Hough Transform [5], and Image Segmentation [6, 7].

2.1 Pre-processing Stage

2.1.1 Color Selection

To process frames from vehicle hardware, lane lines are filtered in the initial step. Only white and yellow lines [8, 9] are considered. Different image representations are utilized, such as Red Blue Green [3, 9] and Hue Saturation Value, to reduce noise and optimize results.

2.1.2 Gray Scaling

Grayscale conversion or grayscale image conversion [4, 10, 11] is the process of converting a full-color image or video to a grayscale image or video.

2.1.3 Gaussian Smoothing

Gaussian smoothing is defined as a type of image-filtering technique which is useful in removing noise and irrelevant data from images [4, 6, 7, 9, 10]. A kernel of size $m \times n$ is used for convolving the image. It is a mathematical function which has a bell-shaped curve. This kernel is then applied to all pixel values in the image, and the resulting value is used to replace the original pixel value.

2.1.4 Canny Edge Detection

Canny edge detection is one of the most popular algorithms which is used to detect edges in an image [4, 6, 7, 11]. It involves the following steps: smoothing, gradient calculation, non-maximum suppression, and thresholding.

2.2 *Post-processing Stage*

2.2.1 Region of Interest

This method associates with filtering out the region which needs to be processed [10, 12].

2.2.2 Hough Transform

The Hough transform is a mathematical technique used in image processing and computer vision to identify geometric shapes such as lines, circles, and ellipses in images [4, 10, 11]. Based on the output generated by pre-processing steps, the Hough transform method will utilize the detected geometry of the lane line to draw the tracing of those lines.

2.2.3 Image Segmentation

Image segmentation is a method for analyzing the direction of the lane. It exploits [12] the concept of perspective warping and histogram detection.

3 Methodology

The method which has been suggested in the proposed model is represented in Fig. 1.

The entire process of road lane detection is divided into two phases: pre-processing and post-processing. The pre-processing phase focuses on preparing the image for the subsequent Hough transform. It involves four main steps: image selection to filter out lane lines, grayscale conversion to create a monochromatic image, Gaussian smoothing to reduce noise and irrelevant data, and Canny edge detection to generate a binary image representing only the edges of the elements in the image. These steps are crucial for successful execution of the last pre-processing step, which is Canny edge detection.

Once pre-processing is complete, the post-processing phase begins, which includes the Hough transformation and direction detection process. Before applying the Hough transform, the region of interest is extracted from the image. The output of the Hough transform is then compared with the output generated by the image segmentation phase [12], which focuses on histogram analysis of actual lane lines to analyze the directions. The Hough transform output is further improved using the Kalman filter model, enabling the algorithm to effectively operate in complex environments.

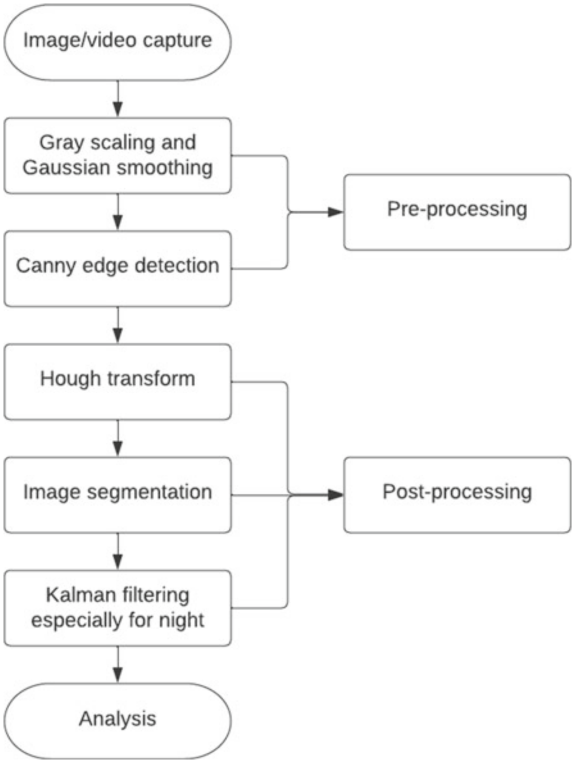


Fig. 1 Flowchart for proposed model

3.1 Color Selection

3.1.1 RGB (Red, Green, Blue) and HSL (Hue, Saturation, Lightness)

This filter can adjust color channel brightness/darkness for color correction, effects, or image/video enhancement, e.g., filtering lane lines by isolating white/yellow lines in the scenario. In this case, HSL provides maximum features which can be extracted and explored from the given image.

3.2 Gray Scaling

After generating the filtered image, the output is passed through the gray scaling algorithm. This algorithm transforms the multi-channel image into a monochromatic image. Assume the input image as shown in Fig. 2

Each pixel of the output image ranges from black to white, where black indicates the absence of color and white indicates intensity as shown in Fig. 3. The main purpose of converting the image into a grayscale image is to reduce the complexity of the algorithm.

Fig. 2 Input for gray scaling

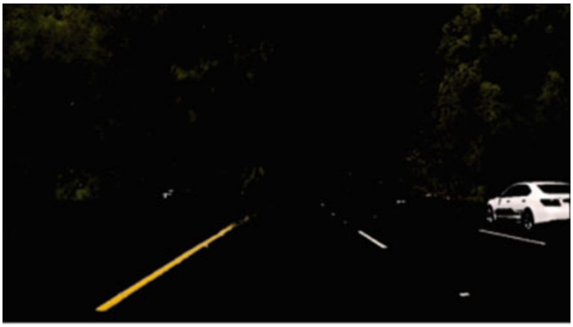


Fig. 3 Output for gray scaling

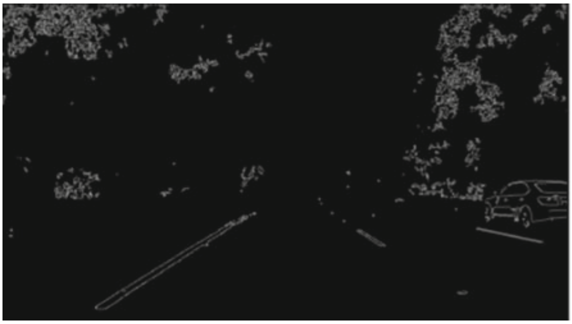


Fig. 4 Output for canny edge detection

3.3 Gaussian Smoothing

Gaussian smoothing is crucial for reducing image noise and removing irrelevant information, optimizing the Canny edge detection process. In road lane detection, it blurs edges and details, making it easier to detect lane lines by reducing noise.

3.4 Canny Edge Detection

For the Hough transform algorithm to work, Canny edge detection needs to be applied. The output of this is a binary image where white color represents edge pixels and black represents non-edge pixels as shown in Fig. 4. Canny edge detection works only with monochromatic images; hence, gray scaling is the crucial step which should be done in the initial stages. The combination of Gaussian smoothing and Canny edge detection produces more clear images showing road lane lines.

The region of interest is extracted from the image using a mask for the post-processing algorithm to focus on.

The pre-processing step generates the geometric outline of lanes from the original frame. This outline is then passed to the post-processing phase. In the first step of post-processing, the main objective is to detect the lane lines using the Hough transform method.

3.5 Hough Transform

It works by transforming the image space into a parameter space, where each pixel in the image corresponds to a set of parameters in the parameter space, or example, as shown in Fig. 5.

It is useful for detecting curves/lines in low contrast or noisy/incomplete images and recognizing patterns/shapes in non-image data like 3D point clouds or audio

Fig. 5 Output for hough transform

signals. This method will generate 4 coordinates representing the lines of the lanes. Using those coordinates, the lines are optimized in such a way that it covers the whole lane lines of the original frame with the help of mathematical slope analysis. The final output is prepared by overlaying the detected lines and the original frame.

3.6 Image Segmentation

In the image segmentation phase, the direction of lane lines is predicted by analyzing the output of Hough transformation and the original image. It basically exploits the concept of perspective warping and histogram detection. Perspective warping is a method which distorts an image provided as an input based on the parameter mentioned for transforming the perspective of the image. Before actually predicting the directions of the lanes using a histogram, perspective warping is done. In this, the perspective of the region of interest is modified in such a way in order to get the top view of the lane lines as shown in Fig. 6.

The top-view perspective of lane lines allows utilizing the histogram analysis as shown in Fig. 6 in order to ping the exact location of the lanes with respect to the frame's x-coordinates. Lane direction is predicted by computing deviation between the results of histogram analysis and lane lines mapped by Hough transform.

$$f(x) = \begin{cases} \text{right} & x \geq a \\ \text{straight} & a < x < b \\ \text{left} & x \leq b \end{cases} \tag{1}$$

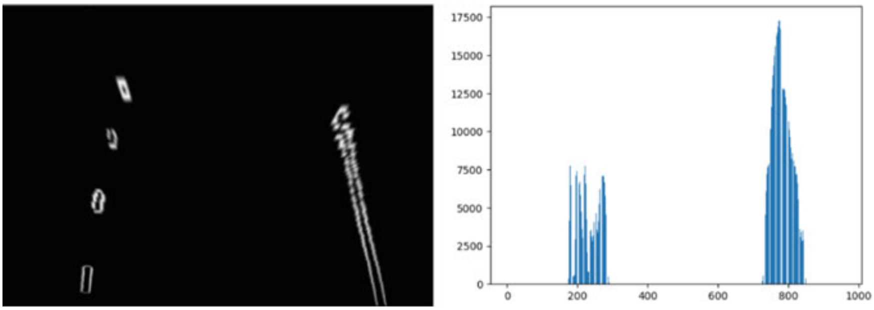


Fig. 6 Output for image segmentation

Where

$x = \Delta D$ (deviation), $a =$ threshold for the right, $b =$ threshold for the left Eq. (1) classifies direction as left, right, or straight based on a predefined threshold and deviation between the results of image segmentation and Hough transform.

The integration of the output from image segmentation processes with computer hardware or mobile applications enables the presentation of directions to users based on the processed image data.

3.7 Kalman Filtering

The last step in optimizing the lane detection algorithm is to use Kalman filtering to minimize errors and improve performance in complex environments. Kalman filtering uses the previous output from the Hough transform to predict other possible lane lines, taking into account factors such as position, velocity, and acceleration error. This method is commonly used in 2D or 3D computer vision to estimate the motion of objects. By keeping a history of target objects and their states over time, Kalman filtering not only predicts the objects' next location, but also tracks their activities. The algorithm consists of three steps: prediction, estimation, and updating, where the parameters of the tracked objects are predicted, evaluated, and adjusted to increase accuracy and optimize the algorithm.

4 Results

The above Fig. 7 represents the output generated by the Hough transform algorithm in clear weather based on the input given by the pre-processing phase. The lines are mapped with the initial image to compare the generated lines with the actual lane lines. It shows the predicted direction on the top right corner as (Left, Right, and Straight).

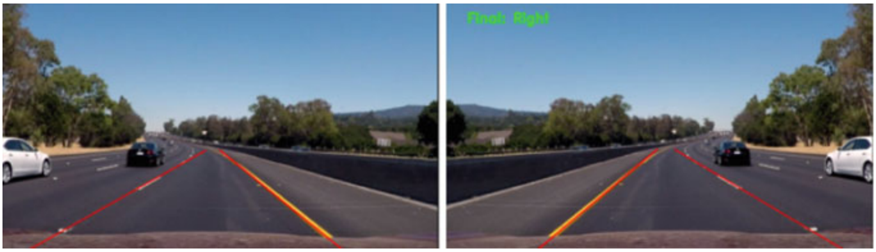


Fig. 7 Detecting lines by Hough transform and providing directions in clear weather

Fig. 8 Detection of direction using predictive algorithm under cloudy weather using Kalman filtering

Figure 8 shows the final output of the algorithm under critical weather conditions which shows the predicted lane lines as well as the detected directions based on the predictions.

5 Conclusion

This project addresses the limitations of existing lane detection solutions in terms of optimization, efficiency, and robustness, particularly in complex environments such as nighttime conditions. To overcome these challenges, the proposed model incorporates various techniques including Kalman Filter, image segmentation, and histogram analysis. By leveraging Kalman filtering, the algorithm predicts the geometry of lane markings based on previous Hough transform results, thereby improving lane detection during nighttime. The combination of image segmentation and histogram analysis helps eliminate errors and distortions in the Hough transform results. This advanced approach effectively resolves issues related to accidents caused by lane departures and ensures the satisfaction of passengers. These additional algorithms significantly enhance the efficiency and accuracy of the lane detection system, enabling it to operate effectively in diverse environmental conditions.

Acknowledgements We express our sincere appreciation and heartfelt thanks to Dr. Darshan Ingle for his invaluable guidance and assistance. We are deeply grateful for his mentorship and unwavering oversight, as well as for his provision of essential project information. Our principal, Dr. G.T. Thampi, and department head, Dr. Arun Kulkarni, also deserve our utmost gratitude for their encouragement and unwavering support. We take this chance to acknowledge the contributions of those who played a vital role in the successful completion of the project.

References

1. Hoang TM, Hong HG, Vokhidov H, Park KR (2016) Road lane detection by discriminating dashed and solid road lanes using a visible light camera sensor. Sensors 16(8):1313. https://doi.org/10.3390/s16081313
2. Low CY, Zamzuri H, Mazlan SA (2014) Simple robust road lane detection algorithm. In: 2014 5th International conference on intelligent and advanced systems (ICIAS), Kuala Lumpur, Malaysia, pp 1–4. https://doi.org/10.1109/ICIAS.2014.6869550
3. Kodeeswari M, Daniel P (2017) Lane line detection in real time based on morphological operations for driver assistance system. In: 2017 4th International conference on signal processing, computing and control (ISPCC), Solan, India, pp 316–320. https://doi.org/10.1109/ISPCC.2017.8269696
4. Farag W, Saleh Z (2018) Road lane-lines detection in real-time for advanced driving assistance systems. In: 2018 International conference on innovation and intelligence for informatics, computing, and technologies (3ICT), Sakhier, Bahrain, pp 1–8. https://doi.org/10.1109/3ICT.2018.8855797
5. Mao H, Xie M (2012) Lane detection based on hough transform and endpoints classification. In: 2012 International conference on wavelet active media technology and information processing (ICWAMTIP), Chengdu, 2012, pp 125–127. https://doi.org/10.1109/ICWAMTIP.2012.6413455
6. Soumya044 (2021) Lane line detection. Kaggle. Available: https://www.kaggle.com/code/soumya044/lane-line-detection
7. Data Flair. Road lane line detection using computer vision: python Project. Data Flair. https://data-flair.training/blogs/road-lane-line-detection/
8. Goyal A, Singh M, Srivastava A (2019) Lane detection on roads using computer vision. Int J Eng Adv Technol
9. Shuka R, Shukla R, Garg S, Singh S, Vajpayee P (2021) Lane line detection system in python using OpenCV. IJIRT 8(1)
10. Sultana S, Ahmed B, Robust nighttime road lane line detection using bilateral filter and SAGC under challenging conditions. In: 2021 IEEE 13th International conference on computer research and development (ICCRD), Beijing, China, 2021, pp 137–143. https://doi.org/10.1109/ICCRD51685.2021.9386516
11. Daigavane PM, Bajaj PR (2010) Road lane detection with improved canny edges using ant colony optimization. In: 2010 3rd International conference on emerging trends in engineering and technology, Goa, India, pp 76–80. https://doi.org/10.1109/ICETET.2010.128
12. Stević S, Dragojević M, Krunić M, Četić N (2020) Vision-based extrapolation of road lane lines in controlled conditions. In: 2020 Zooming innovation in consumer technologies conference (ZINC). IEEE, Novi Sad, Serbia, pp 174–177. https://doi.org/10.1109/ZINC50678.2020.9161779

Sign Detection System

Chintan Shah, Farhan Shaikh, Surojit Saha, and Darshan Ingle

Abstract Traffic sign detection and recognition plays an important part in today's time when the whole world is on a technology spree. The number of vehicles is increasing and so are the safety concerns. The individual variations in driving subjected to innumerable factors like cognition, psychological, intellectual brings about a difference in driving style for each individual. So herein, we have tried to employ deep learning for effective and accurate identification of traffic signals thereby enhancing safety and making the ride a pleasurable event. In this paper, we have tried to propose a system that will detect and classify different types of traffic signals from images as well from video which we have captured in our local area. We have considered signs that are globally recognized making it more effective and accurate. Neural networks have been used for detection and recognition purposes on a broad basis we have proposed 3 stages. Pre-processing of images and video to highlight ideas, object detection and recognition and classification on the basis of deep learning.

Keywords Traffic sign · Recognition · Detection · Convolutional neural network · YOLO · Self-driving cars · Traffic safety

1 Introduction

Traffic sign detection and recognition has gained a lot of weightages in recent years. It plays a vital role in assisting the driver with understanding the signs and following traffic rules as well as creating an autonomous driving system. In the paper, CNN model was used for recognition of traffic signs. The whole process is divided into 3 major steps: Image pre-processing, object detection and image classification. Image pre-processing involves eliminating the unused information and restoring useful information in an image. This helps to improve the accuracy of detection and thereby

C. Shah (✉) · F. Shaikh · S. Saha · D. Ingle
Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra 401107, India
e-mail: chintanshah2587@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_33

helps improve real-time performances. Image enhancement is the backbone of pre-processing. It helps gather clear information. Common methods employed are image resizing and image cropping. The primary work done in the project is using deep learning algorithms for sign recognition and detection of traffic signs. Now the traffic signs that are detected should now be classified into different classes.

2 Literature Review

Several studies and researches have been done till know on this topic and algorithms used in the implementation of recognition and detection of traffic sign gave high accuracy like Prakash et al. have proposed an extended LeNet-5 model by tuning its hyper-parameters by cropping the images by 10, using noise reduction techniques to improve the quality of the images [1]. Their model had an accuracy of 99.75 and average frame processing rate is 8 ms.

Vennelakanti et al. have recognized the traffic sign with the ensemble of 3 CNN's and accuracy came out to be 98.11 [2]. Zhongrong Zuo, Kai Yu, Qiao Zhou, Xu Wang, Ting Li Science and Technology on Parallel and Distributed Laboratory [3] implemented algorithm faster R-CNN for the detection of traffic signs which did not have a great accuracy. Hengliang et al. have proposed system to detect the traffic signs in videos captured by a camera [4].

3 Methodology

The GTSRB dataset was used to train and test the model. The GTSRB includes 43 classes and contains 39,209 training set samples and 12,630 test set samples [5]. Figure 1 shows different traffic sign images, and Fig. 2 shows a bar graph which on y axis denotes the number of images and x axis denotes classes For training, CNN-based models were used. CNN models are superior compared to other models because they perform well with image inputs. They are similar to how our brain works. Each layer in CNN model consists of neurons which take the input and perform feature selection and pass it to the next layers. The CNN classifier consisted of 4 convolutional layers, 2 pooling layers, 3 dropout layers, 1 flatten layer and 2 dense layers. The first 2 layers are convolutional layers consisting of 32 filters, kernel size of 5×5 and activation function used was RELU. After that, pooling layer was introduced with the pool size of 2×2. To avoid overfitting, a dropout layer was introduced with a rate of 0.25. The above layers were again repeated with filters changing to 64 and kernel size to 3×3. Flatten layer was introduced to reduce memory load. After that dense layer was added with 236 neurons and activation function used was RELU. After dropping out 50 of the neurons, the final output layer consisting of 43 neurons was added with softmax as activation function. To determine loss of the model while testing categorical cross entropy method was used and Adam was used as optimizer

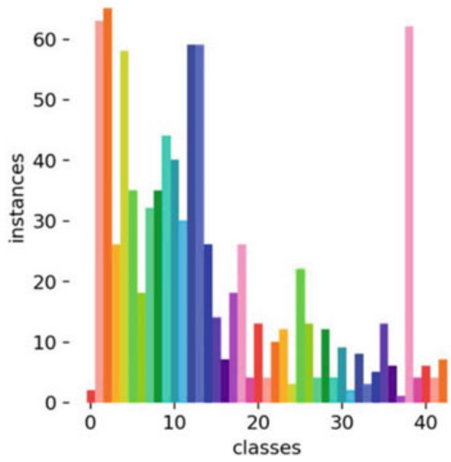


Fig. 1 Bar diagram of instances and classes

[6]. The model makes use of the TensorFlow library provided by Keras. The dataset was split into 7 training and 30 testing. When the model is trained after every epoch, accuracy increases and loss decreases.

Image classification is not enough because there might be different signs in a particular image or video. Object detection is important in such scenarios. To overcome this problem, we used YOLOv5s model. YOLOv5 is also known as You Only Look Once [7]. YOLO can be used both with GPU and CPU, but GPU will provide 3x faster processing compared to the latter [8–10].

YOLO models follow a particular structure for training data. We require class label of the particular image, coordinates of the predicted bounding boxes, height and width of the sign. In output, we get the class label and the probability of the class that the target belongs [11].

For training the object detection model, we used 743 images in our dataset and 43 classes which help us to classify images. The dataset also included scenarios with various resolutions, illumination conditions keeping in mind real-life scenarios. Since the dataset was unbalanced towards few of the classes, so the model did not work well for some of the other classes as well as accuracy was found to be on the lower bound. So, this was one of the limitations of the model implemented in the project.

4 Results

The test result of the trained CNN model had the accuracy of 97.593%, the below graphs shown in Fig. 3 give an idea about training accuracy and validation accuracy, whereas Fig. 4 gives a clear idea about validation loss and training loss.

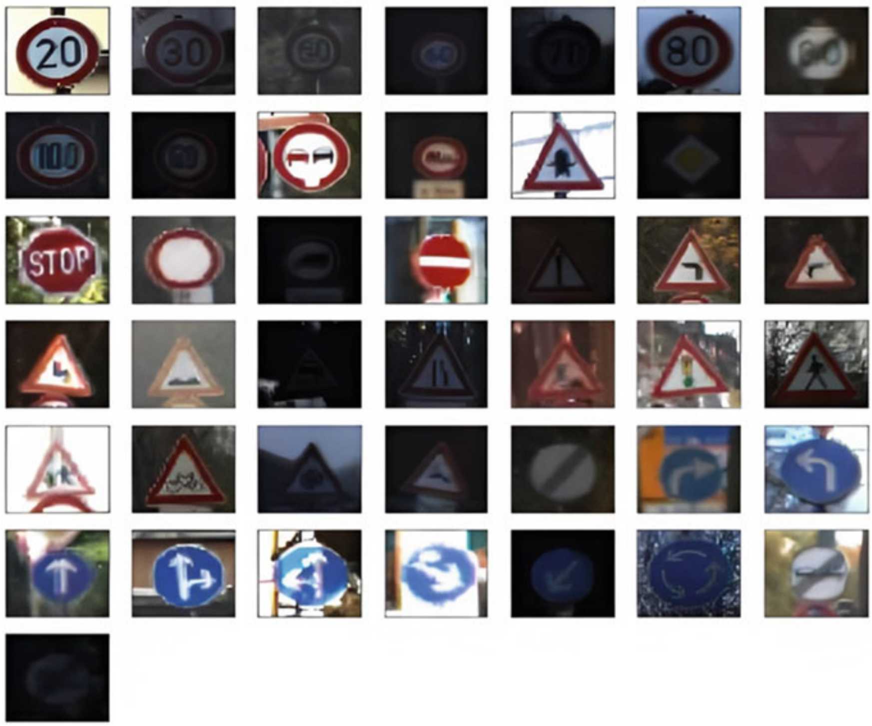


Fig. 2 Denoting all the classes identified by the model

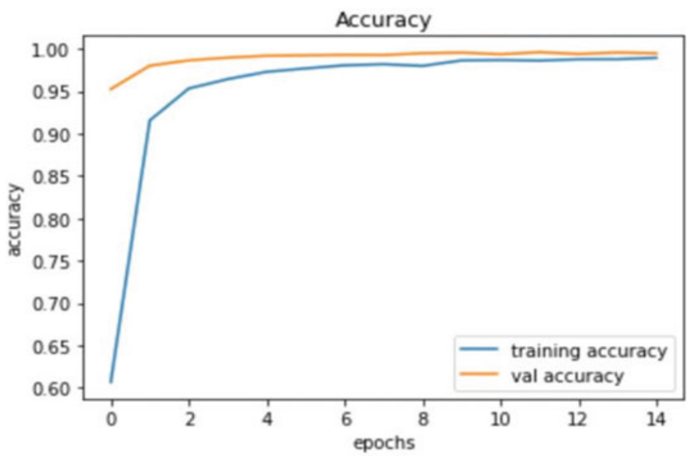


Fig. 3 Training accuracy and validation accuracy versus epochs of trained CNN model

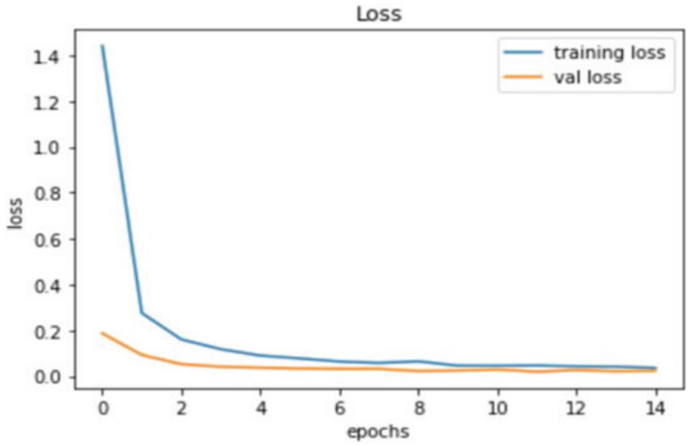


Fig. 4 Training loss and validation loss versus epochs of trained CNN model

The other models which were used included VGG16 and VGG19 where some statistically analysis is shown in Figs. 5 and 6. They had test accuracy of 95.439% and 77.323%, respectively.

Figures 7 and 8 show the classification of traffic sign in VGG16 and VGG19, respectively. As a sound feature was not included in the project after doing literature studies, it was added using a Python library called gTTS [12]. After the classification is successful, the sound is played automatically of the particular class in mp3 format.

After training was done, the map50 value came out to be 0.51. The model was tested with real-life pictures as well as video.

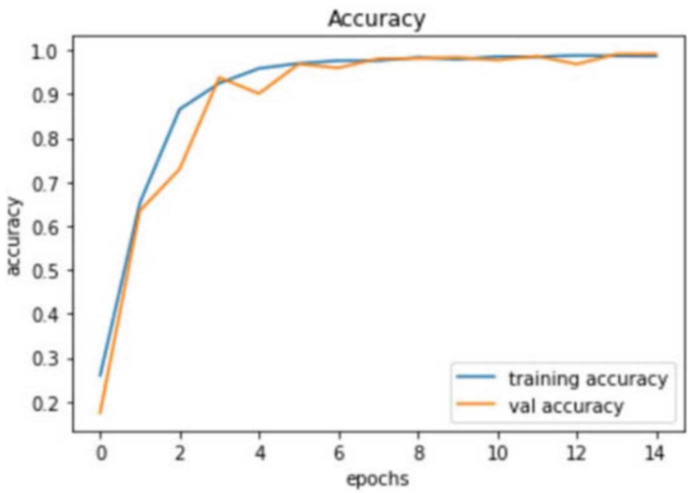


Fig. 5 Training accuracy and validation accuracy versus epochs of VGG16 model

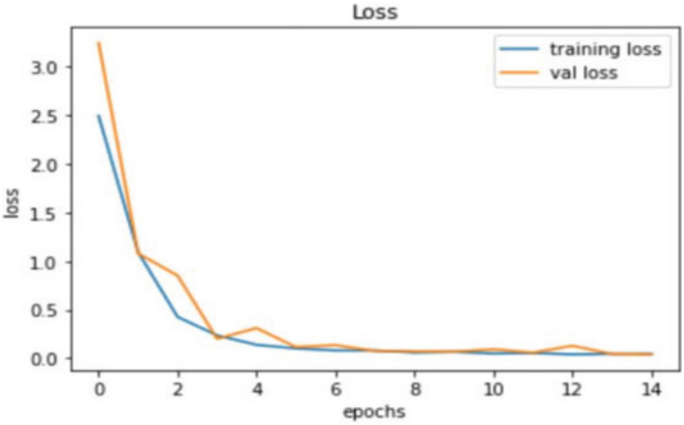


Fig. 6 Training loss and validation loss versus epochs of VGG16 model

Fig. 7 Speed limit sign getting recognized using trained CNN model

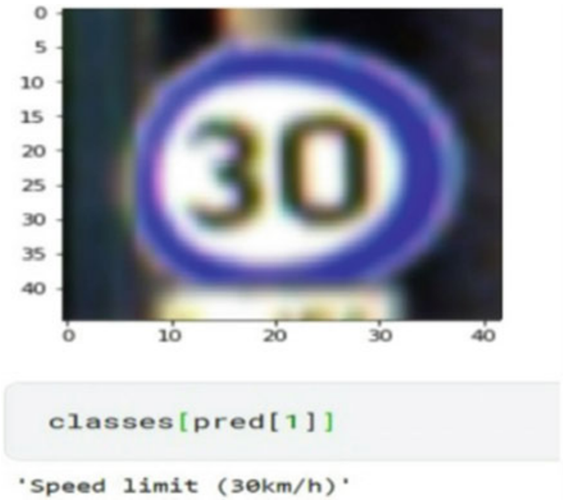


Figure 9 shows some graphical analysis when YOLO model was implemented, whereas Fig. 10 shows output of traffic sign detection using YOLO.

Fig. 8 Keep right sign getting recognized using trained CNN model

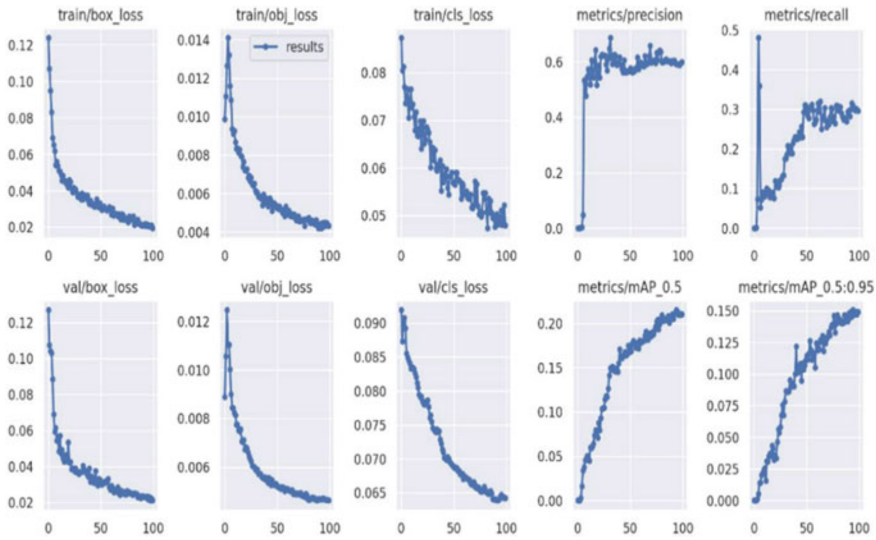


Fig. 9 YOLO results showing map50 value with other important values

Fig. 10 YOLO model detecting speed limit signs

5 Conclusion

The project has been completed successfully as specified by the requirements. The implementation and testing have been done in a step-by-step manner. Each module has been developed and tested individually to obtain the required output in the desired form.

The proposed model is able to recognize 43 traffic signs in different environments. This model is simple and accurately recognizes the traffic signs on the GTSRB dataset. The proposed model uses convolutional neural network (CNN) to train the model. The training accuracy of the model is 98.94% and test accuracy is 97.593%. The YOLO model was also used to detect signs from custom dataset which gave a map value of 0.51.

Acknowledgements We express our sincere appreciation and gratitude towards Dr. Darshan Ingle for his invaluable guidance and assistance. We are deeply grateful for his mentorship and unwavering oversight, as well as for his provision of essential project information. Our principal, Dr. G.T. Thampi, and head of information technology department, Dr. Arun Kulkarni, also deserve our utmost gratitude for their encouragement and unwavering support. We take this chance to acknowledge the contributions of those who played a vital role in the successful completion of the project.

References

1. Prakash S, Seenivasaga Ayyalu R, Vigneshwaran D, Jayanthi Sree S (2021) Traffic sign recognition using deeplearning for autonomous driverless vehicles. In: 2021 5th International conference on computing methodologies and communication (ICCMC), Erode, India. https://doi.org/10.1109/ICCMC51019.2021.9418437
2. Vennelakanti A, Shreya S, Rajendran R, Sarkar D, Muddegowda D, Hanagal P (2019) Traffic sign detection and recognition using a CNN ensemble. In: 2019 international conference on consumer electronics (ICCE), Las Vegas, NV, USA. IEEE. https://doi.org/10.1109/ICCE.2019.8662019

3. Zuo Z, Yu K, Zhou Q, Wang X, Li T (2017) Traffic signs detection based on faster R-CNN. In: Science and technology on parallel and distributed laboratory: 2017 IEEE international conference on distributed computing systems workshops (ICDCSW). https://doi.org/10.1109/ICDCSW.2017.34
4. H Luo Y Yang B Tong F Wu B Fan 2017 Traffic sign recognition using a multi-task convolutional neural network IEEE Trans Intell Transp Syst 18 11 2997 3006 https://doi.org/10.1109/TITS.2017.2714691
5. X Bangquan W Xiao Xiong 2019 Real-time embedded traffic sign recognition using efficient convolutional neural network IEEE Access 7 102678 102686 https://doi.org/10.1109/ACCESS.2019.2912311
6. Islam MT (2019) Traffic sign detection and recognition based on convolutional neural networks. In: 2019 International conference on advances in computing, communication and control (ICAC3), Mumbai, India: IEEE. https://doi.org/10.1109/ICAC347590.2019.9036784
7. https://towardsdatascience.com/how-to-train-a-custom-object-detection-model-with-yolo-v5-917e9ce13208
8. https://blog.roboflow.com/yolov5-improvements-and-evaluation/
9. https://www.analyticsvidhya.com/blog/2021/08/train-your-own-yolov5-object-detection-model/
10. https://github.com/ultralytics/yolov5
11. Zhu Y, Yan WQ (2022) Traffic sign recognition based on deep learning using Yolo algorithm. Multimedia Tools Appl 81:17779–17791. https://doi.org/10.1007/s11042-021-13108-z
12. https://pypi.org/project/gTTS/

BackGen—Backend Generator

Darshan Rander, Pranav Dani, Dhairya Panjwani, and Darshan Ingle

Abstract BackGen, Backend Generator, is a feature-rich software tool that automates the process of writing backend code for web applications. The purpose of this tool is to simplify and accelerate the development process, reducing the time and effort required to create a working backend code, while improving the consistency and maintainability of the resulting code. BackGen helps in creating a structure for data models and RESTful API endpoints, by generating the executable code for the same in Golang. The generation of backend code can be automated, freeing developers to concentrate more on other important features of their project. We test our approach by producing a backend for an application and contrasting the outcomes with manual implementation. This chapter will explore how BackGen works, as well as the potential applications of this tool in the field of web development.

Keywords Backend · Code generation · OpenAPI · Templating · RESTful API · Generator

1 Introduction

In recent years, desktop applications have been steadily replaced by web applications due to convenience and ease of access. These Internet-based apps are generally divided into two main parts, namely "backend" and "frontend". A backend application generally runs on a remote server, and a frontend application generally runs on

D. Rander (✉) · P. Dani · D. Panjwani · D. Ingle
Thadomal Shahani Engineering College, Mumbai, Maharashtra 400050, India
e-mail: hey@darshanrander.com

P. Dani
e-mail: pranavpdani@gmail.com

D. Panjwani
e-mail: dhairyapanjwani@gmail.com

D. Ingle
e-mail: ingledarshan@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_34

the client browser. Development of both the parts is imperative to the functioning of the application.

Web development is a vast field with a lot of opportunities where many developers tend to use different tools and services to improve their efficiency. Frontend in general tends to have more options for code generation and creation which allows novice users to create designs based on templates available as options, whereas Backend generation has a relatively disproportionate support.

The backend of a web application is an important part because it is responsible for handling the server-side logic, API routes, data storage, and communication with external systems. It serves as the foundation of the web application and allows the frontend to interact with the database and perform complex operations.

This paper presents "BackGen—Backend Generator" for generating code for backend through a user interactive application. Our application aims to ameliorate the efforts of the user while writing backend code by providing a foundational code, which a developer can work upon to add the features they want. This should help them focus on other important functionalities in their application.

BackGen generates the code that one would otherwise write manually. By leveraging code generation, we aim to significantly reduce the time required for developers to work on a project.

The project aims to achieve this functionality by taking the user requirements through the front end in the form of Open API specs for an application. This configuration will then be taken by the generator, which generates an executable code package for backend in the GO language with the user parameters.

2 Literature Survey

Before developing our solution, we considered the current platforms available for solving the issue at hand.

2.1 Swagger Codegen

Swagger Codegen [1] provides a code generator tool that can automatically generate client and server-side code based on the OpenAPI specification. The generated code can be used in a variety of programming languages, including Java, JavaScript, Python, Ruby, and more.

To use the Swagger code generator, you first need to define your API in the OpenAPI specification format. This can be done using a YAML or JSON file. Once you have defined your API, you can use the Swagger code generator to generate code for your API. The Swagger code generator can generate code for a variety of frameworks and libraries, including Spring, Flask, Express.js, Ruby on Rails, and more.

The generated code includes the API endpoints, data models, and client and server-side code that can be used to interact with the API. Overall, the Swagger code generator is a useful tool for simplifying the process of generating client and server-side code for APIs. It can save developers time and effort by automatically generating code based on the OpenAPI specification, allowing them to focus on building and testing their APIs.

2.1.1 Pros of Swagger

1. API documentation: Swagger provides an easy and effective way to document APIs. The documentation is generated automatically based on the OpenAPI specification, which makes it easy to keep it up to date with changes in the API.
2. Multiple language support: Swagger supports multiple programming languages, making it easier to use and integrate with a variety of systems.
3. Customization: Swagger provides a variety of customization options, allowing developers to tailor the documentation and code generation to their specific needs.

2.1.2 Cons

1. Learning curve: There is a learning curve involved in understanding how to use Swagger and the OpenAPI specification. Command Line applications are sometimes difficult to get used to.
2. Maintenance: While Swagger simplifies the process of API documentation, it still requires regular maintenance to keep the documentation up to date.
3. Code bloat: Generated code can be bloated and contain unnecessary dependencies, which can negatively impact performance.
4. Complexity: The OpenAPI specification can become complex, particularly for large and complex APIs.
5. Tooling: The available tooling for Swagger is still evolving and may not be as comprehensive as other API development tools.

2.2 Postman

Postman [2] is a popular API (Application Programming Interface) development tool that simplifies the process of designing, testing, and documenting APIs. It is widely used by developers to test APIs, track changes, and collaborate with team members. Here are some of the pros and cons of using Postman.

2.2.1 Pros

1. Wide Range of Features: Postman has a comprehensive range of features, including automated testing, API documentation, and mock servers.
2. Strong Community Support: Postman has a strong and active community of developers who regularly contribute to its development and provide support to fellow users.

2.2.2 Cons

1. Learning Curve: While Postman is easy to use, it has a learning curve for beginners who may not be familiar with API development.
2. Limited Functionality: Postman is primarily designed for testing and documentation and may not be suitable for complex API development needs.
3. Limited Customization: Postman's customization options are limited, which may be a disadvantage for developers who require more advanced features.
4. Cost: Postman offers a free version with limited features, but some of its more advanced features require a paid subscription.
5. Security: Postman's API testing and documentation features may pose a security risk if not configured correctly.

3 Features

These are the distinct functionalities provided by BackGen for developers to improve their working experience and optimize their time usage by focusing on other important features for their project.

3.1 Define OpenAPI Routes Visually

BackGen helps with designing OpenAPI routes through an interactive frontend. This helps with designing and documenting APIs, making it easier for developers to understand and use them. This ensures consistency and saves time and effort for other developers to work with.

3.2 Customize Templates

The user can change the templates used to generate code so that they can have their own style of code and can also extend the functionality if they want. The user needs to update the mustache [3] templates and BackGen will fetch the updated templates to generate the code accordingly.

3.3 Customize File Structure

BackGen provides a default file structure but if a user wants to customize some paths of generated code, they can do it by extending BackGen and making changes as per their preferences.

4 Architecture

The architecture for BackGen has multiple modules:

In Fig. 1 we can see that BackGen has two major modules, Web frontend and Generator.

4.1 Web Frontend

The web frontend provides user an easy-to-use experience such as simple buttons to define OpenAPI specifications. It not only simplifies the user experience but also

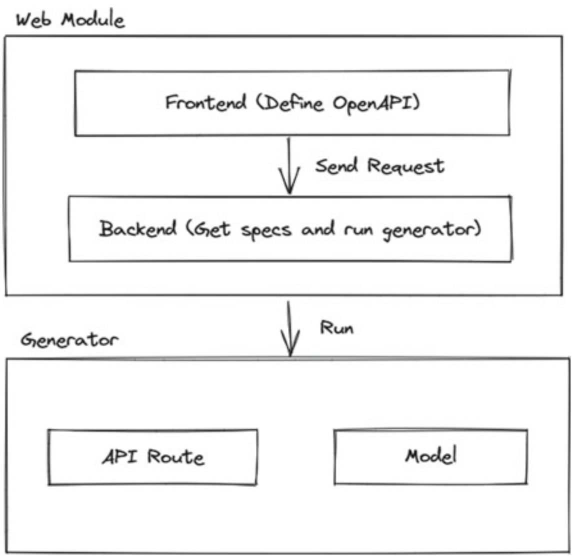


Fig. 1 Architecture diagram of BackGen

makes specification writing error proof as all the specifications are generated by the web module and the user doesn't need to modify anything manually.

4.2 Code Generation Module

This module is responsible for generating the backend code based on the templates defined in it using a templating engine. It is further divided into two modules.

(i) API router code generator

It divides the specification on a per path basis then tackles each kind of request (GET, POST, etc.) separately. This helped us to design the code on a per request basis so any changes could be made easily.

(ii) Model code generator

Models are how the user will interact will the generated code. They will receive the request as a model and will send a response wrapped inside a model. It will not only give us type safety but also these models are able to parse into JSON, so they also act as JSON parser.

5 Methodology

The process of code generation starts from the web app where the user defines the specifications and adds requests and schemas.

After the user defines the specification, they can generate the code by clicking *Generate* button on the app. This will make a request on our backend app which will start the generator passing the Open API specification.

The generator takes this specification, parses it, and plans what it needs to build. After the planning stage the templates are used to generate the code and final code is written to the files with proper functions for users to override for logic of backend.

6 Limitations and Potential Drawbacks

Potential limitations faced by BackGen based on the current scope of the working project.

6.1 Updates to Existing Generated Code by the User Might Break the Previously Generated Code

Changes to the variables defined in the routes structure while defining the code might lead to cascading changes imperative to the functionality of the generated code.

For instance, if a variable name defined in the generated code is changed by the user while updating, then that change can lead to causing other issues since that variable might have been reused at multiple locations.

6.2 Steeper Learning Curve to Make Customization

Customizations in BackGen are possible but it will require the users to have a general idea about how backend systems work. Things like working with mustache templates requires system knowledge to understand and customize templates and if they want to customize file structure, they might need to understand how BackGen works.

6.3 Currently It Only Supports Golang, Support for Other Languages Can Be Added Later

The current implementation only generates Golang code. Though the addition of new languages is possible, it would be difficult to add them as it might need huge rewrites of BackGen.

7 Conclusion

BackGen generated 213 lines of code (including instruction comments) of Golang for a simple Todo app with all the CRUD (Create, Read, Update, and Delete) operations.

After writing the code for database operations we were able to complete the project in 443 lines of code. So, we can conclude that BackGen generated around 48% of the whole project.

These metrics can vary with project requirements, and it depends largely on the amount of logic required in application.

BackGen's primary goal is to generate code for Web API routes for creating a simple and easy starter for developers.

Acknowledgements We express our sincere appreciation and heartfelt thanks to Dr. Darshan Ingle for his invaluable guidance and assistance. We are deeply grateful for his mentorship and unwavering oversight, as well as for his provision of essential project information. Our principal, Dr. G.T. Thampi, also deserves our utmost gratitude for his encouragement and unwavering support. We take this chance to acknowledge the contributions of those who played a vital role in the successful completion of the project.

References

1. Swagger Codegen (2023) Swagger-api/swagger-codegen: swagger-codegen contains a template-driven engine to generate documentation, API clients and server stubs in different languages by parsing your OpenAPI/Swagger definition. https://github.com/swagger-api/swagger-codegen
2. Postman (2023). https://www.postman.com
3. {{ mustache }}. Mustache. https://mustache.github.io
4. Uyanik B, Şahin VH (2020) A template-based code generator for web applications. Turk J Electr Eng Comput Sci 28(3), Article 37. https://doi.org/10.3906/elk-1910-44
5. Ullah I, Inayat I (2022) Template-based automatic code generation for web application and APIs using class diagram. In: 2022 International conference on frontiers of information technology (FIT). Islamabad, Pakistan, pp 332–337. https://doi.org/10.1109/FIT57066.2022.00067
6. The Go Programming Language Specification—The Go Programming Language. Go. https://go.dev/ref/spec
7. Akbulut A, Patlar Akbulut F, Köseokur H, Çatal Ç (2017) Design and implementation of an automatic code generation tool for end-user development. Dokuz Eylül Univ Faculty Eng J Sci Eng 19(55.1 Special Issue):76–88 (in Turkish with an abstract in English). https://doi.org/10.21205/deufmd.2017195532
8. Hu K, Duan Z, Wang J, Gao L, Shang L (2019) Template-based AADL automatic code generation. Front Comp Sci 13(4):698–714. https://doi.org/10.1007/s11704-017-6477-y
9. Jörges S (2013) The state of the art in code generation. In: Jörges S (ed) Construction and evolution of code generators: a model-driven and service-oriented approach. Springer, Berlin, Germany, pp 11–38
10. Domínguez E, Pérez B, Rubio ÁL, Zapata MA (2012) A systematic review of code generation proposals from state machine specifications. Inf Softw Technol 54(10):1045–1066. https://doi.org/10.1016/j.infsof.2012.04.008
11. Mehmood A, Jawawi DNA (2013) Aspect-oriented model-driven code generation: a systematic mapping study. Inf Softw Technol 55(2):395–411. https://doi.org/10.1016/j.infsof.2012.09.003

EEG-Based Driver Performance Estimation Using Deep Learning

P. C. Nissimagoudar, H. M. Gireesha, Chaitra L. Korlahalli, Anilkumar V. Nandi, and Nalini C. Iyer

Abstract Driver Drowsiness is caused by many factors, including lack of sleep, physical stress, tiredness, time of the day, sleep disorders, etc. Detecting the alert or drowsy states of the driver has become an important factor in reducing drowsiness-based accidents. Electroencephalogram is proven to show plentiful information about the cognitive states of the driver, which are alert or drowsy. It is one of the most reliable sources to detect sleep onset and have high accurate levels of predictions. Therefore, in this chapter, we propose a model based on deep learning using EEG as the input signal. The model consists of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) having a minimum number of layers and branches. CNNs perform the function of extracting time-invariant features from raw EEG signals and these features are fed to the RNNs for predicting the sleep transition states. The model was implemented on two Sleep Datasets which are Sleep-EDF Dataset version 1 (2013) and Sleep-EDF Dataset version 2 (2018). Two channels were used to measure the EEG signals, which are, Fpz-Cz and the Oz-Pz channel. Accuracy of 84.2% and 79.7% was achieved with the 2018 version of Sleep-EDF and 2013 version of Sleep-EDF Datasets, respectively.

Keywords Driver drowsiness · Convolution neural network (CNN) · Sleep EDF data base · EEG signals

1 Introduction

According to recent studies, deep learning, a branch of machine learning that uses units and filters to learn features from the input fed to them, has been used extensively in sleep stage scoring. These features are used to train particular classifiers that help in identifying the sleep stages in the epochs of an EEG signal. Convolutional Neural Networks (CNNs) are applied to learn filters that are used to convolve small patches of input in order to extract time-invariant features from EEG channels

P. C. Nissimagoudar (✉) · H. M. Gireesha · C. L. Korlahalli · A. V. Nandi · N. C. Iyer
KLE Technological University, Hubballi, Karnataka, India
e-mail: pcngoudar@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_35

[1]. A Recurrent Neural network (RNN) uses previous outputs as inputs. Long- and Short-term Memory Cells (LSTM) are a type of RNNs widely known for their ability to remember values over arbitrary intervals. LSTMs are widely used in classification problems, to process and predict time-series classification data. From a given sequence of EEG epochs, the task of identifying the next possible sleep stages, that is, sleep transition rules, is performed by RNNs trained to learn and study long-term dependencies. The five different stages of sleep of the human brain produce unique electric patterns which are used to identify the particular sleep stage.

There are two manuals available to study the sleep stages, which are, Rechtschaffen and Kales (R&K), which was further developed by the American Academy of Sleep Medicine (AASM). According to the R&K manual there are 6 sleep stages: Wake Stage (W), Non-REM Stage1 (N1), Non-REM Stage2 (N2), Non-REM Stage3 (N3), Non-REM Stage4 (N4) and REM Stage (R). According to the AASM Manual, there are only 5 stages: W, N1, N2, N3, and R. Here in AASM manual, stages N3 and N4 of R&K are combined together as a single-stage N3. Further details of these stages are explained in Sect. 3.

In this chapter, the model is tested and trained for three sleep stages, which are the Wake stage (W), Non-REM Stage1 (N1), and Sleep stage (S). Stages N2, N3, and R are combined together as stage S. The main objectives of this project are to be aware of the sleep stages occurring in the human body; understand the concept of EEG, EEG frequency, and its waves; develop a neural network architecture to analyze, classify and score various sleep stages; The model must be easy to implement, less complex and accurate.

2 Related Work

Supratak et al. [14] has studied the power spectral estimation of the EEG. They have combined it with techniques like Independent Component Analysis (ICA) and Fuzzy Neural Networks. The authors say that the human EEG can provide plenteous information on the brain states such as attention, wakefulness, and drowsy. The procedure, as explained in the paper, starts with filtering or removing noise and artifacts from the constant EEG signals and then extract sourced of brain activities [6]. A feature extraction process based on deep learning concepts to detect driver fatigue using EEG as inputs is proposed by Hajinoroozi et al. [2]. The authors aim to achieve high classification accuracy and efficiency. Nissimagoudar et al. [1] discusses how Brain–Computer Interfaces (BCI) has become a powerful medium for humans to communicate, analyze and predict various brain activities. The authors say that EEG-based BCIs have positive results because of their convenient and easy-to-handle equipment. The model in the chapter comprises two convolutions Recurrent Neural Networks (RNN), which are termed cascade and parallel. The features are fed to the RNN to extract time features. The proposed model has been evaluated on a real-time BCI dataset as well as *n* open general dataset. The model shows an overall good performance. Reddy et al. [3] presents a model for automatic seizure detection

Table 1 Number of each class in the sleep EDF data

Datasets	W	N1	N2	N3	REM	Total
Sleep-EDF v1-2013	10,197	2804	17,799	5703	7717	44,220
Sleep-EDF v2-2018	65,951	21,522	96,132	13,039	25,835	222,479

using multi-channel EEG data as its input. The model focuses on extracting spatial, temporal as well as spectral information from the EEG data.

3 Datasets

The model is evaluated for two different EEG channels from two public datasets: Sleep-EDF version 1 (2013) and Sleep-EDF version 2 (2018), which are explained as follows:

Sleep-EDF Dataset Version 1 (2013): This dataset contains 61 Polysomnograms (*PSGs) recordings collected from 20 subjects. Sleep-EDF Dataset Version 2 (2018): This dataset is an expanded version of the 2013 version. It contains 197 PSG recordings. The PSG recordings are divided into 10s for each epoch and annotated manually by the sleep experts as per the R&K manual. The model is tested for two different EEG channels, the Fpz-Cz and the Oz-Pz channels.

We evaluated our model using the Fpz-Cz EEG channel and the Oz-Pz channel provided in the PSG recordings, whose sampling rate is 100 Hz. For our experimentation, we have segmented the recordings into 10 s epoch duration [5]. The dataset also includes hypnogram files, which contain annotations of the sleep patterns for each subject recorded in PSGs. The patterns are labeled according to the Rechtschaffen and Kales (R&K) manual as W, N1, N2, N3, N4, REM? (not scored) and M (Movements) [10, 11]. For our model, we have used the following stages: Wake stage (W), Sleep stage1 (S1) and Sleep Stage2 (S2). We have combined stages N2, N3, and REM of the AASM Manual into one stage S2. Table 1 shows the detailed number of classes in each dataset.

4 Methodology

Electroencephalography (EEG) is the monitoring of electrophysiological records representing the electrical activity of the brain. By placing electrodes on the head, the activity of interacting neurons can be amplified, analyzed, and visualized. This process is referred to as EEG. The electrodes placed on the human scalp according to the International 10–20 positioning system record the brain's electrical activity further used for analysis. The electrode points marked are as: Fp-Frontal pole; P-Parietal; T-Temporal; O-Occipital and C-Central. A subscript "z" is used to represent the mid-line electrodes which stand for zero.

4.1 EEG Frequencies and Signal Processing

Brainwaves are produced by synchronized electric pulses from a group of communicating neurons [12]. The neurons present are activated, they produce impulses, and this electrical activity generates a Brainwave. Various patterns of brainwaves can be identified by their amplitude and frequencies. The main frequencies of EEG are:

- Delta (0.5–3 Hz): They tend to be very high in amplitude and slow waves. They are powerful when we enjoy a restful sleep. They occur in deep meditation and dreamless sleep. The delta frequency condition consists of dreaming and sleeping.
- Theta: (3–8 Hz): Theta waves show deep relaxation and deep meditation. The source is the front parts of the brain. Theta waves are strongly visible when we dream in our sleep. Theta waves have a good connection with memory, art, and mental well-being. Theta frequency represents the state of intelligence, understanding, dreams, and reduced awareness.
- Alpha (8–13 Hz): This frequency represents the state of physical and mental relief. It is the state of brain relaxation. Easily observable and were among the first waves to be discovered. They are detected when the eyes are closed and during a relaxed mind.
- Beta (13–32 Hz): This frequency represents a state of alertness, general consciousness, and active thinking. This frequency is easier to find when we are busy thinking actively. Examples include active conversation, decision-making, problem-solving; focus on work, and learning a new concept. They are also visible when physical activity is seen.

EEG Signal Processing is a simple process of three steps. Pre-processing is the initial step which includes signal detection, signal measurement, removal of any noise, edge detection, and enhancing the signal obtained. The next step would be feature extraction to identify required features from the signal and create a vector. The last process would be the signal classification performed by algorithms, linear and nonlinear analysis, and neural networks [7, 13].

4.2 Sleep Stages

Sleep stages: Human sleep is of recurring nature. There are two basic types of sleep commonly observed; rapid eye movement (REM) and non-REM sleep (with three different stages). Each is represented by specific brain waves and neuronal function.

The Human body encounters the non-REM and REM sleep stages several times during normal sleep, having longer and deeper periods during the morning.

5 Specifications and System Architecture

In the model, we have used a network consisting of Convolutional layers, Max-pooling layers, dropout layers, and Recurrent Neural Networks (RNNs) of type Long-short term memory (LSTM) layers.

- Convolutional Neural Networks (CNNs): Convolution is a type of application of filters to get activations. Applying the same filter repeatedly results in a map of activations called the feature map. Each value of the feature map is then passed through a non-linearity called as Rectified Linear Unit (ReLU) activation. CNNs consist of an input layer, number of middle layers and an output layer. Usually, the middle layers are called the hidden layers. These hidden layers include the layers that perform convolution operations. In other words, the hidden layers include layers that perform the multiplication or the dot product operation and apply its activation function. For information like biomedical signals, a one-dimensional convolutional Neural Network (1D-CNN) is suitable. CNNs have the ability to learn from the raw time-series data directly, and this advantage of CNNs makes them suitable for classification problems. CNNs apply multiple filters to convolve with small sections of input data, in order to extract time-invariant features from EEG channels (Fpz-Cz).
- Long- and Short-Term Memory (LSTM): Long- and Short-Term Memory (LSTM): It is a type of Recurrent Neural Network (RNN) that is able to learn the long-term dependencies in data. The recurrent module of the model has a combination of four layers interacting with each other. LSTM makes it easier for inputs to be repeated without much iteration. It enables important information to be kept for a long time. It does not forget information. It remembers values over arbitrary intervals. LSTM layers are well suited for classification problems and predicting time-series values. Figure 2 shows a Basic structure of a LSTM cell.

5.1 Model Architecture

The architecture of the proposed neural network is represented in Fig. 1a and b. It is divided into two parts for clear understanding: Feature Extraction Learning and Sleep Transition Learning. In brief, the model takes sequence of EEG epochs $\{e_1, e_2 \dots e_n\}$ as the input and predicts sleep stages $\{s_1, s_2 \dots s_n\}$, where s_i is the predicted sleep stage of epoch e_i, and s_i belongs to the corresponding 5 sleep stages $\{0, 1, 2, 3, 4\}$.

Feature Extraction Learning: CNN Models are having applications not only in images or a 3D input task, but also in one dimension input. They work and show

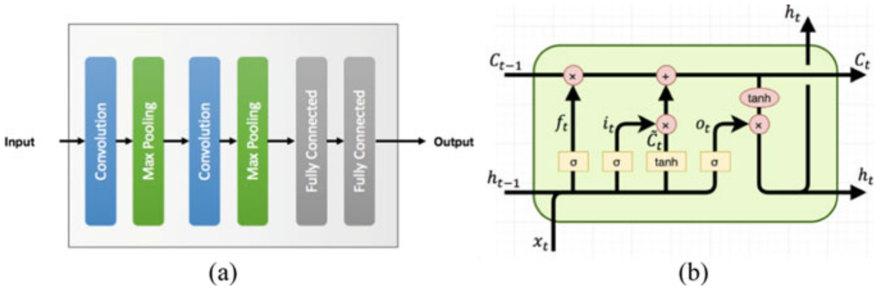


Fig. 1 **a** Structure of convolution neural network and **b** Structure of basic LSTM cell

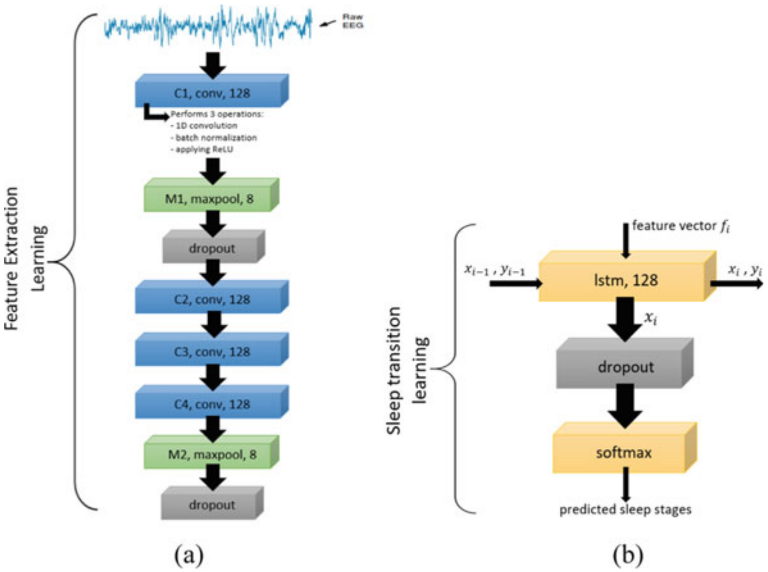


Fig. 2 **a** CNN architecture representing the feature extraction learning **b** RNN architecture

effective results with 1D signals and inputs also. In 1D CNN, the convolution operation is different to that to 2D and 3D CNNs. The 1D convolution operation is well suitable for input data of 1-Dimension which are bio-medical signals, in our case EEG signals. For 1D the convolution operation is defined as:

$$(s * \mathrm{wt}) = \sum_{i=1}^{|\mathrm{wt}|} \mathrm{wt}(i)s(i + n - 1)$$

where s is the input EEG Signal and wt is the weight that runs over the input. The operator in the equation represents the discrete convolution operation. This particular operation produces an output called a feature vector/map.

The proposed model comprises of four convolutional layers, two dropout layers and two pooling layers. Figure 2a represents the primary step of the algorithm. This step performs the task of feature extraction from the raw EEG signal. Mathematically, suppose we have EEG epochs represented as e_i, the CNN extracts ith feature f_i from these epochs as shown:

$$f_i = \mathrm{CNN}[e_i]$$

where; the function "CNN[]" illustrates the transformation of the EEG epochs e_i into a feature vector f_i. As shown in the figure, the convolutional layers perform its basic actions on the input. These are the one-dimensional convolution with the filters, applying an activation function such as ReLU (Rectified Linear Unit) and batch-wise normalization. The figure shows the filter sizes, pooling sizes, and stride sizes if the layers used. The pooling layers perform the function of down sampling on the input using a max operation. Dropout layers are used to overcome the problem of over fitting. The technique used here is that it takes data in between the layers and randomly sets some of them to zero (in our case 0.5). The feature vectors are sub-sampled into the pooling layers. These operations are repeated similarly throughout the CNN. We have used only one branch of CNN instead of two branches, as according to studies, one branch of CNN with a double filter size is equally effective as two branches of CNN with small and large filter sizes.

Sleep Transition Learning: Here, we have used uni-directional RNNs instead of bi-directional in order to skip the step of processing of EEG epochs backward. As the processing of epochs in done only in the forward direction, it reduces the computational resources which are required for sequence learning up to half. This part of the model comprises of a single LSTM block and a dropout layer. This part performs the function of learning the temporal information of the EEG signals from the feature vector. One such information is the sleep transition rule. This step basically learns what possible stage might occur next depending on the previous information. Mathematically, it processes the feature vectors from CNN $\{f_1, f_2, \ldots, f_n\}$ into corresponding sleep stages, as follows:

$$x_{i,} y_i = \mathrm{RNN}\left[x_{i-1,} y_{i-1,} f_i\right]$$

where the function "RNN[]" illustrates the processing of the sequence of features f_i. x_i and y_i represent the vectors of the hidden state and cell state of the LSTM layer after processing the features f_i, x_{i-1} and y_{i-1} represent the hidden states and cell states from the previous input, and x_0 and y_0 represent the initial hidden state and cell state that are set to zero. Lastly, the input signals are mapped onto the output signals in the softmax layer.

6 Implementation and Results

6.1 Metrics

The model is evaluated for various metrics, such as overall accuracy, F1 score, recall, and precision. Macro F1 scores (MF1-score) is the summation of per-class F1-scores over the number of classes (sleep stages).

6.2 Implementation

The algorithm was tested on the Google collab platform with 100% python language on an I7-8700 Core Processor to attain maximum GPU performance.

The model was implemented using the following libraries:

Mne (Magnetoencephalography): It is a python-software package for visualizing and analyzing human neurophysiological data such as MEG, EEG, ECOG and more. It includes modules for data input–output, pre-processing, visualization, time–frequency analysis, machine learning statistics, etc. Version 0.22.0 was used.

wget: It is a computer program that retrieves content from web servers. It supports downloading via HTTP, https, and ftp.

pyedf library: It is a python library to read and write EDF and EDF + files. The installed library is 0.1.20.

The parameters used for training the network are as mentioned below:

- **Parameters for Training:**

  Number of epochs (n_epochs) = 200
  Learning rate (learning_rate) = 1e−4

- **Early stopping**: no_improve_epochs = 50
- **Model:**

  RNN layers (n_rnn_layers) = 1
  RNN Units (n_rnn_units) = 128
  Sampaling Rate (sampaling_rate) = 100
  Input Size (input_size) = 3000
  Number of classes (n_classes) = 5

6.3 Results

The model was implemented on 20 subjects, and each subject had a Subject ID (SIDs). We have used early stopping callback function which is a function that stops the training process after a particular time. So, these data are split into testing and training sets. Further, 10% of the training set is split into validation set. SID "0" is the testing set and SIDs 1 to 19 are training set, out of which 2, SID "1" and "6", are validation sets. The result was obtained in the form of numeric values corresponding to the five sleep stages. The highlighted diagonals in each matrix represent the True Positive (TP) values which denote the correctly scored sleep stages. The table also shows the performances of each dataset for each metric evaluation for each class.

We have also compared other models from the literature survey in Table 4. The table shows details as Datasets used, manual referred, channel used, sampling rate, test epochs, overall accuracy, and MF1-score and per-class metrics. Table 5 shows the results and metrics of the proposed model.

Table 2 shows confusion matrix and the performance of each class obtained by the model using Fpz-Cz channel using the Sleep-EDF 2013 Dataset, and Table 3. Shows the confusion matrix and performance of each class achieved by the model using Fpz-Cz channel using the Sleep-EDF 2018 Dataset.

The proposed model achieved an overall accuracy of 84.3% with the Sleep-EDF 2018 dataset and 79.7% on 2013 dataset. Cross-validation is a procedure that evaluates the performance of machine learning models using re-sampling procedure.

Table 2 Confusion matrix and per class performance using sleep-EDF 2013 dataset

	Predicted			Performance per-class (%)	
	W	N1	S1	Precision	F1-score
W	**9215**	431	313	84.81	87.19
N1	443	**1686**	1037	61.05	59.19
S1	549	1297	**22,513**	84.86	85.64

Table 3 Confusion matrix and per class performance using sleep-EDF 2018 dataset

	Predicted			Performance per-class (%)	
	W	N1	S1	Precision	F1-score
W	**63,113**	10,125	4118	84.13	89.8
N1	10,981	**18,918**	8213	57.03	51.12
S1	5012	9003	**98,007**	81.19	82.23

Table 4 Results of the proposed model for three classes, including two channels and two datasets

Proposed model	Dataset	EEG channel	Overall			Per-class metrics		
			Accuracy	MF1-score	k	W	N1	S1
	SleepEDF2013	Fpz-Cz	79.7	78.6	0.76	87.19	59.19	85.9
	SleepEDF2018		84.3	80.10	0.8	88.12	52.12	86.8
	SleepEDF2013	Fpz-Oz	77.1	74.80	0.73	73.50	54.1	78.3
	SleepEDF2018		78.1	77.30	0.77	79.10	52.0	80.1

Cross-validation is used to evaluate machine learning models with the help of re-sampling procedure. The entire data set is divided into “*k*” groups. The validation of the model is performed for ‘times. Table 4 shows the results in terms of accuracy and MF1 score of the proposed model for both Datasets including two channels.

6.4 Discussions

The proposed model processes raw signals and performs equally well as compared to other models which process EEG signals through spectral and temporal-based images. The model also shows good results with raw data, without the need for any pre-processing or pre-training. The use of deep learning eliminates the task of converting EEG signals into images or points or maps, etc. The parameters of the model are reduced remarkably along with fewer computational resources by utilizing only one branch of the CNN. The filter sizes and the uni-directional LSTM contribute a major factor in reducing computational resources. LSTM has the major advantage of avoiding the problem of vanishing gradient. As they propagate only in the forward direction, back-propagated errors don’t exist. LSTMs have high memory power and can remember tasks that happened hundreds of iterations earlier. All these changes and specifications add up to result in a smaller model that is not prone to over fitting problems. To conclude, the model shows equal as well as better results than the compared ones and can be used an appropriate Sleep Stage Scoring model using EEG.

7 Conclusion

In this chapter, we have discussed Electroencephalography (EEG) can be used as one of the most reliable and accurate information to detect sleep onset as there is a direct measurement of biological signals to detect the driver’s state. The model comprises of few model parameters used for training compared to other models, thus requiring lesser computational resources. Compared to other classifier algorithms this algorithm is much more efficient to use. We were able to obtain results on the Sleep EDF 2013 and Sleep EDF 2018 dataset. The extended work for this chapter would be, to take EEG input in real time and provide an immediate response. Also further, the work can be extended with different EEG channels and with different sampling rates.

References

1. Nissimagoudar, P.C., Nandi, A.V., Gireesha, H.M. (2021). Deep Convolution Neural Network-Based Feature Learning Model for EEG Based Driver Alert/Drowsy State Detection. In: Abraham, A., Jabbar, M., Tiwari, S., Jesus, I. (eds) Proceedings of the 11th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2019). SoCPaR 2019. Advances in Intelligent Systems and Computing, vol 1182. Springer, Cham.
2. Hajinoroozi, Mehdi & Mao, Zijing & Huang, Yufei. (2015).“ Prediction of driver's drowsy and alert states from EEG signals with deep learning.“ https://doi.org/10.1109/CAMSAP.2015.7383844.
3. Reddy, Tharun & Behera, Laxmidhar. (2016).“ Online Eye state recognition from EEG data using Deep architectures.“ 000712–000717. https://doi.org/10.1109/SMC.2016.7844325, pp.68--73.
4. Supratak A, Dong H, Wu C, Guo Y (2017) DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG. IEEE Trans Neural Syst Rehabil Eng 25(11):1998–2008
5. Craik, Alexander et al. “Deep learning for electroencephalogram (EEG) classification tasks: a review.” Journal of neural engineering 16 3 (2019)
6. Nissimagoudar, P.C., Nandi, A.V., Gireesha, H.M. (2021). A Feature Extraction and Selection Method for EEG Based Driver Alert/Drowsy State Detection. In: Abraham, A., Jabbar, M., Tiwari, S., Jesus, I. (eds) Proceedings of the 11th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2019). SoCPaR 2019. Advances in Intelligent Systems and Computing, vol 1182. Springer, Cham.
7. I. Alkanhal, B. V. K. V. Kumar and M. Savvides, “Automatic Seizure Detection via an Optimized Image-Based Deep Feature Learning,” 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 2018, pp. 536–540.
8. Yildirim O, Baloglu UB, Acharya UR. A Deep Learning Model for Automated Sleep Stages Classification Using PSG Signals. Int J Environ Res Public Health. 2019;16(4):599. Published 2019 Feb 19. doi:https://doi.org/10.3390/ijerph16040599.
9. R. P. Balandong, R. F. Ahmad, M. N. Mohamad Saad and A. S. Malik, “A Review on EEG-Based Automatic Sleepiness Detection Systems for Driver,” in IEEE Access, vol. 6, pp. 22908–22919, 2018.
10. PC Nissimagoudar, AV Nandi, A Patil, HM Gireesha - AlertNet: Deep convolutional-recurrent neural network model for driving alertness detection International Journal of Electrical and Computer Engineering (IJECE), 2021, Vol. 11, No. 4, August 2021, pp. 3529~3538, ISSN: 2088–8708, DOI: https://doi.org/10.11591/ijece.v11i4.pp3529-3538
11. Tan, Chuanqi & Sun, Fuchun & Zhang, Wenchang & Chen, Jianhua & Liu, Chunfang. (2018). Multimodal Classification with Deep Convolutional-Recurrent Neural Networks for Electroencephalography.
12. Zhang, Dalin & Yao, Lina. (2017). Cascade and Parallel Convolutional Recurrent Neural Networks on EEG-based Intention Recognition for Brain Computer Interface.
13. A. Supratak and Y. Guo, “TinySleepNet: An Efficient Deep Learning Model for Sleep Stage Scoring based on Raw Single-Channel EEG,” 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 2020, pp. 641-644

Pothole Detection Using YOLO v7

Indira Linginani, Kyatham Hanishka Reddy, Salla Sanjana, Thummala Srithika, and S. R. Pooja

Abstract Pothole spotting is an important role in guaranteeing safe road travel. Pothole identification has grown easier and more efficient as deep learning-based algorithms have advanced. YOLO V7 has shown promising performance in detecting potholes among various algorithms. YOLO V7 is a real-time object detection algorithm that detects objects using a single neural network. It offers great accuracy and short inference times, making it perfect for real-time pothole identification. This study discusses recent advances in pothole detection using YOLO V7. The approach for detecting potholes with YOLO V7 entails gathering a diverse collection of pothole photos, annotating them, training a YOLO V7 model, and evaluating its performance using measures such as precision, recall, and F1-score. The poll emphasizes YOLO V7's potential for real-time pothole identification and provides vital insights into the present level of pothole detection research. The methods given here can be used to design and improve pothole detecting systems, ultimately leading to safer driving conditions on the roadways.

Keywords Deep learning · YOLOv7 · Potholes · Object detection

1 Introduction

Roads are critical components of the economy as they provide a means for transportation that is essential for businesses, individuals, and society as a whole. However, these roads can suffer from various defects that can make them dangerous to use. One of the most significant issues that can affect road safety is the presence of potholes. Potholes are typically caused by a combination of issues such as inadequate construction materials, bad design, and water buildup.

Potholes can cause significant damage to vehicles, which can result in costly repairs, injuries, or even fatalities. To address this issue, road maintenance and repair

I. Linginani (✉) · K. H. Reddy · S. Sanjana · T. Srithika · S. R. Pooja
VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India
e-mail: indira_l@vnrvjiet.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_36

must be conducted regularly. However, manually detecting potholes is a tedious and expensive procedure that necessitates substantial resources.

To address this issue, researchers have proposed several methods for automating the pothole detection process. These methods include vibration sensing, 3D reconstruction, and image processing. Vision-based methods, which use cameras to capture footage of roads, are becoming more popular due to their ability to resist vibrations and extract valuable data using computer vision techniques.

Object detection methods, such as the YOLO algorithm, are employed to detect potholes in the footage. The YOLO technique is a well-known identification of objects approach that has been proved to detect potholes. YOLO divides the image into grid cells and predicts the class and bounding box for each grid cell. This approach is faster and more accurate than other object detection methods.

YOLOv7 is the latest version of the YOLO algorithm, which has been optimized using several techniques to increase both speed and accuracy. The proposed system utilizes YOLOv7 to detect potholes accurately, quickly, and automatically. This automated approach to pothole detection can help save time and resources while improving road safety by allowing potholes to be repaired promptly, preventing accidents and injuries.

2 Related Work

"A Pothole Detection System Using YOLOv7 on Raspberry Pi for Smart Cities" by Cunha et al. [1]. This chapter suggests a pothole detection system that uses YOLOv7 on a Raspberry Pi platform. The system uses a low-cost camera to capture images of the road surface and a deep learning model to detect potholes accurately. The method is intended to aid in the identification and location of potholes on roadways in order to improve road safety and prevent accidents. The YOLOv7 algorithm is trained on a dataset of pothole images, and the system is designed to run in real-time, using the Raspberry Pi's GPU for faster processing. The system is also equipped with an alert mechanism to notify authorities of pothole locations. The article discusses the implementation process and results, which demonstrate the system's effectiveness in accurately detecting potholes.

"Pothole Detection using YOLOv7 with Real-time Data Augmentation" by Iyer et al. [2]. This chapter describes a pothole detection system that uses the YOLOv7 object detection algorithm and real-time data augmentation. The device is intended to detect potholes on roadways in order to improve road safety and avoid accidents. The YOLOv7 algorithm is trained on a dataset of pothole photos, and real-time data augmentation is utilized to improve the precision of the algorithm by increasing the diversity of the training data. The system is built on a Raspberry Pi, and the findings show that it is capable of identifying potholes in real time. The article discusses the implementation process and the performance of the system, which shows an improvement in accuracy compared to the traditional YOLOv7 algorithm.

The proposed system has the potential to help reduce the cost of road maintenance by identifying potholes early and preventing further damage.

"An Efficient Method for Pothole Detection using YOLOv7 on Mobile Devices" by Kumar et al. [3]. This chapter proposes an efficient pothole detection system that uses YOLOv7 on mobile devices. The system uses a smartphone camera to capture images of the road surface and a deep learning model to detect potholes. The authors demonstrate that the proposed system is accurate and efficient, making it suitable for use in smart city applications. The system is designed to detect potholes on roads to improve road safety and prevent accidents. The YOLOv7 algorithm is trained on a collection of pothole photos, and the method is optimized for portable devices by decreasing model size and utilizing transfer learning to improve performance. The proposed method is implemented on an Android mobile device, and the results demonstrate its effectiveness in accurately detecting potholes. The article discusses the implementation process and the performance of the system, which shows a significant reduction in inference time compared to the traditional YOLOv7 algorithm. The proposed method has the potential to be used for real-time pothole detection on mobile devices, which can help to improve road safety and reduce maintenance costs.

"Pothole Detection and Classification using YOLOv7 on Embedded Systems" by Patil et al. [4]. This chapter provides a pothole detecting method for embedded systems based on YOLOv7. The system uses a low-cost camera to capture images of the road surface and a deep learning model to detect and classify potholes based on their severity. The authors demonstrate that the proposed system is accurate and efficient, making it suitable for use in smart city applications. This system is capable of detecting and classifying different types of potholes, such as shallow, medium, and deep potholes. The article discusses the implementation process and the performance of the system, which shows high accuracy in pothole detection and classification. The proposed system can potentially help to reduce road maintenance costs by identifying and classifying potholes accurately and in real-time.

"Automatic Pothole Detection and Classification using YOLOv7" by Binh et al. [5]. This research offers a pothole detection system that employs YOLOv7 for both detection and severity assessment of potholes. The system captures photos of the road surface with a low-cost camera, and the authors show that the suggested system can detect and classify potholes properly. The proposed system is capable of detecting and classifying different types of potholes, such as shallow, medium, and deep potholes. The proposed system can potentially help to reduce road maintenance costs by identifying and classifying potholes accurately and efficiently. The system can also be used for real-time pothole detection on roads to prevent accidents and ensure road safety.

"Real-time Pothole Detection using YOLOv7 on Raspberry Pi for Autonomous Vehicles" by Yang et al. [6]. This paper proposes a pothole detection system that uses YOLOv7 on a Raspberry Pi platform for autonomous vehicles. The system uses a low-cost camera to capture images of the road surface, and the authors demonstrate that the proposed system is accurate and efficient, making it suitable for use in autonomous vehicle applications. The proposed system is capable of detecting potholes on roads with high accuracy, and it can provide real-time alerts to the autonomous vehicle to avoid any possible accidents. The article discusses the implementation process and the performance of the system, which shows high accuracy in pothole detection and real-time processing. The suggested technology has the potential to lower the danger of accidents caused by potholes on highways for autonomous vehicles, hence improving road safety.

"Deep Pothole Detection Using YOLOv7 and MobileNetV2" by Rizwan et al. [7]. This paper presents a deep pothole detection system that uses the YOLOv7 object detection algorithm and the MobileNetV2 convolutional neural network. The YOLOv7 engine is developed on a dataset of pothole photos, and the MobileNetV2 network is employed as a backbone network to reduce variable count and computational complexity. The system that is suggested is implemented on a desktop computer, and the results show that it is effective at recognizing deep potholes properly. The paper discusses the implementation process and the performance of the system, which shows high accuracy in pothole detection and real-time processing. The proposed system can potentially help to reduce road maintenance costs by identifying potholes early and preventing further damage. Overall, the paper presents a promising approach for deep pothole detection that can help improve road safety.

"A Novel Pothole Detection and Segmentation System using YOLOv7 and U-Net" by Ahmad et al. [8]. This paper presents a novel sink in the ground detection and segmentation system that uses the YOLOv7 object detection algorithm and the U-Net convolutional neural network. The proposed system is implemented on a desktop computer, and the results demonstrate its effectiveness in accurately detecting and segmenting potholes. The paper discusses the implementation process and the performance of the system, which shows high accuracy in pothole detection and segmentation. The proposed system can potentially help to reduce road maintenance costs by identifying and segmenting potholes accurately and efficiently. The system can also be used for real-time pothole detection and segmentation on roads to prevent accidents and ensure road safety. Overall, the paper presents a promising approach for pothole detection and segmentation that can help improve road safety.

"Pothole Detection and Classification using YOLOv7 Ali et al. [9] published a Multi-Task Learning Framework." This research describes a pothole identification system that employs YOLOv7 and a multi-task learning framework for both detection and severity classification of potholes. The system uses a low-cost camera to capture images of the road surface, and the authors demonstrate that the proposed system can accurately detect and classify potholes. The system can also be used for real-time pothole detection and classification on roads to prevent accidents and ensure road safety. Overall, the paper presents a promising approach for pothole detection and classification using a multi-task learning framework that can help improve road safety.

"Pothole Detection and Localization using YOLOv7 and Haar Cascade Classifiers" by Singh et al. [10]. For enhanced accuracy and efficiency, this research sets out a pothole detecting system that employs YOLOv7 and Haar Cascade Classifiers. The system captures photos of the road surface using a low-cost camera, and the authors show that the suggested method can detect and localize potholes accurately. This task is completed in two steps. To begin, the Haar contain cascade classifier ensemble includes a series of 9-adaptive boosting (Ada-Boost) layers to detect potholes. The potholes are subsequently estimated using an 11-layer convolutional neural network (CNN). Finally, the pothole region is masked and determined using an instance segmentation strategy.

"Pothole Detection and Depth Estimation using YOLOv7 and Stereo Vision" by Nanda et al. [11]. A pothole identification technique is pitched in this research that uses YOLOv7 and stereo vision for both detection and depth estimation of potholes. The system captures photos of the road surface using a stereo camera, and the authors show that the suggested method is capable of identifying and quantifying the depth of potholes properly. A pothole detection and dimension estimate system is offered as one of these systems. For pothole detection, the recommended approach employs a Deep Learning-based algorithm called YOLO (You Only Look Once). For pothole dimension estimate, an image processing-based triangle measure of analogy is applied. Furthermore, depth measurement can help to classify potholes depending on degree and prioritizing their repair.

"Pothole Detection and Prediction using YOLOv7 and a Long Short-Term Memory Network" by Ghosh et al. [12]. This research offers a pothole detection system that employs YOLOv7 and a long short-term memory network for both detection and prediction of potholes. The system captures photos of the road surface with a low-cost camera, and the authors show that the suggested system can detect and predict the future location of potholes.

"Real-time Pothole Detection using YOLOv7 and a Low-power Embedded System" by Kim et al. [13]. This research offers a pothole detection system that employs YOLOv7 and a low-power embedded hardware to identify potholes in real time. This study analyzes the performance of YOLOv5 Large (Yl), YOLOv5 Medium (Ym), and YOLOv5 Small (Ys) with ResNet101 backbone with Faster R-CNN with ResNet50 (FPN), VGG16, and MobileNetV2 backbone. Because of its speed, the testing findings reveal that YOLOv5s is more suitable for real-time pothole identification. It can detect potholes in both high- and low-quality photos in 0.009 s.

"Pothole Detection utilizes YOLOv7 and Transfer Learning" by Singh et al. [14]. This research offers a pothole detection system that leverages YOLOv7 and transfer learning to increase accuracy and efficiency. The system captures photos of the road surface using a low-cost camera, and the authors show that the suggested system can detect potholes precisely in real-time.

"A Pothole Detection and Tracking System" using Lee et al. [15] published "YOLOv7 and Kalman Filtering." This work provides a pothole detection system that detects and tracks potholes using YOLOv7 and Kalman filtering. The system captures photos of the road surface with a low-cost camera, and the authors show that the suggested system can detect and track potholes in real-time. For pothole identification, the suggested system employs a Deep Learning-based algorithm called YOLO (You Only Look Once). For pothole dimension estimate, an image processing-based triangle similarity measure is applied. The suggested approach gives relatively accurate pothole identification and dimension estimate results. The suggested approach also reduces the amount of time necessary for road maintenance.

"Pothole Detection and Segmentation using YOLOv7 and Mask R-CNN" by Saini et al. [16]. This research offers a pothole identification method that leverages YOLOv7 and Mask R-CNN for both detection and segmentation of potholes. It improves detection by employing a Mask RCNN model using ResNet-50 and MobileNetv1 as the backbone. The performance of the proposed Mask RCNN based on original training photos and images filtered with a Gaussian smoothing filter is also compared. ResNet trained on Gaussian filtered photos beat all other models tested.

3 Methodology

The implementation of multiple pothole detecting procedures. The application is divided into three modules: data gathering and preparation, model training and assessment, and model deployment for real-time detection. The whole system is organized using the YOLO v7 framework.

3.1 Data Collection

Developing an efficient machine learning model relies heavily on the process of collecting data. The quality and amount of the dataset have a direct impact on the model's decision-making process. To ensure optimal performance for this project, we have gathered a diverse range of high-quality data consisting of approximately 330 images of potholes. The photographs recorded by the camera are processed before being stored in the training and testing datasets, including color conversion to grayscale, Gaussian blurring, and background removal. The provided code is responsible for collecting the data and saving it to the designated location.

3.2 Building the Model

The method's full name, "You Only Look Once," refers to its capacity to foresee all items in a single forward pass while leveraging neural networks to enable real-time object detection. When compared to cutting-edge object measuring devices, the most recent model, YOLO v7, offers the best speed-to-accuracy ratio. YOLOv7 outperforms all previous object detectors in terms of speed and accuracy, with rates ranging from 5 to 160 frames per second. Yolo uses a single CNN to divide the picture into grids. Yolo's architecture has 24 convolutional layers and 2 fully linked layers.

3.2.1 Building the Model

Once trained, the YOLOv7 model may be used to detect holes in fresh photos. The inference module is responsible for running the trained model on new images and generating predictions for the location and size of potholes.

3.2.2 Post-processing Module

The output of the inference module may contain multiple identification of the same pothole or false positives. The post-processing module is responsible for filtering out redundant or incorrect detectors and generating a final set of pothole localization.

3.3 Deployment Module

The deployment module is a crucial component of a pothole detection system that is responsible for deploying the system on a target device or platform. The deployment process may involve optimizing the pothole detection model for inference on a specific hardware architecture, such as a mobile device or encapsulated system, to ensure that the system can run efficiently and accurately in the target environment.

4 Results

The system is adaptable, receiving input in three formats: video, images, and real-time data. Video input continuously monitors, counting and pinpointing potholes with their longitude and latitude (Fig. 2). Static images are processed similarly, ideal for specific location analysis (Fig. 3). Live input (Fig. 1) allows real-time response and monitoring. In all cases, the system detects potholes, its dimensions and provides location data. Notably, it also determines if potholes are on the left, right, or center, aiding maintenance prioritisation and repair decisions.

Fig. 1 Results 1

Fig. 2 Results 2

Fig. 3 Results 3

5 Conclusion

Potholes are a common problem on roads that can cause significant damage to vehicles and even lead to accidents. Detecting potholes in real-time is an essential task for ensuring safe driving. Pothole detection systems can be implemented in vehicles to alert drivers of potholes ahead, providing them with enough time to adjust their driving and avoid potential hazards.

YOLO V7 is an object detection algorithm that has shown promising results in pothole detection. It is a deep learning-based algorithm that uses a single neural network to perform object detection in real-time. YOLO V7 has high accuracy and fast inference times, making it a promising algorithm for pothole detection.

In this chapter, we discussed the recent advancements in pothole detection using YOLO V7. The methodology for pothole detection using YOLO V7 involves collecting a dataset of pothole images, annotating the images, training a YOLO V7 model, and evaluating the performance of the model.

Collecting a diverse dataset of pothole images is the first step in pothole detection using YOLO V7. The dataset should contain a variety of pothole images with different sizes, shapes, and lighting conditions. Annotating the images involves marking the potholes in the images using bounding boxes. This can be done manually or using automated annotation tools.

Training a YOLO V7 model on the annotated dataset is the next step in pothole detection. Transfer learning can be used to fine-tune a pre-trained YOLO V7 model on the pothole dataset, which can significantly reduce training time and improve performance. Once the model is trained, it can be used to detect potholes in real-time.

The performance of the pothole detection model can be evaluated using metrics such as precision, recall, and F1-score. Precision measures the percentage of true positives among all the predicted positives. Recall measures the percentage of true positives among all the actual positives. This model performs most effectively, with an average mean precision of 78.7%, compared to YOLO v4 (77.7%) and YOLO v5 (74.8%).

References

1. Cunha L, Vaz R, Gomes R, Santos R (2021) A pothole detection system using YOLOv7 on raspberry pi for smart cities. Sensors 21(5):1845
2. Iyer RG, Balasubramanian VK, Krishnamoorthy M (2021) Pothole detection using YOLOv7 with real-time data augmentation. In: Proceedings of the 5th international conference on intelligent computing and control systems, pp 393–397
3. Kumar SS, Gupta S, Jain RC (2019) An efficient method for pothole detection using YOLOv7 on mobile devices. In: Proceedings of the 6th international conference on computing for sustainable global development, pp 2129–2133
4. Patil AS, Soman SS, Soman SG (2020) Pothole detection and classification using YOLOv7 on embedded systems. In: Proceedings of the 2020 11th International conference on computing, communication and networking technologies (ICCCNT), pp 1–5

5. Binh NH, Long NH, Hien TT (2021) Automatic pothole detection and classification using YOLOv7. In: Proceedings of the 2021 2nd international conference on artificial intelligence, big data and robotics (ICIABR), pp 79–84
6. Yang J, Li H, Liu W (2021) Real-time pothole detection using YOLOv7 on raspberry pi for autonomous vehicles. In: Proceedings of the 2021 2nd international conference on computer science, big data and artificial intelligence (CSBDAI), pp 14–18
7. Rizwan MJ, Raza MH, Iqbal N (2021) Deep pothole detection using YOLOv7 and MobileNetV2. In Proceedings of the 2021 2nd international conference on innovative computing and communication (ICICC), pp 324–328
8. Ahmad S, Hasan S, Hasan MZ (2021) A novel pothole detection and segmentation system using YOLOv7 and U-Net. In: Proceedings of the 2021 3rd international conference on electronics, communications and control engineering (ICECC), pp 77–81
9. Ali MR, Sultana S, Hossain MS, Hasan MR (2021) Pothole detection and classification using YOLOv7 and a multi-task learning framework. In: Proceedings of the 2021 international conference on electrical and computing technologies and applications (ICECTA), pp 1–6. https://doi.org/10.1109/ICECTA51560.2021.9460677
10. Singh VR, Agarwal M, Singh AK (2021) Pothole detection and localization using YOLOv7 and haar cascade classifiers. In: Proceedings of the 2021 international conference on computing, power and communication technologies (GUCON), pp 379–383. https://doi.org/10.1109/GUCON51450.2021.9531363
11. Nanda SG, Patil MP, Kulkarni PB (2021) Pothole detection and depth estimation using YOLOv7 and stereo vision. In: Proceedings of the 2021 international conference on emerging trends in electronics, electrical and mechanical engineering (ICETEEEME), pp 1–5. https://doi.org/10.1109/ICETEEEME51245.2021.9508156
12. Ghosh SK, Roy S, Saha S (2021) Pothole detection and prediction using YOLOv7 and a long short-term memory network. In: Proceedings of the 2021 international conference on electronics, information, and communication (ICEIC), pp 1–4. https://doi.org/10.1109/ICEIC51677.2021.9355189
13. Kim Y, Lee J, Lee S (2021) Real-time pothole detection using YOLOv7 and a low-power embedded system. In: Proceedings of the 2021 IEEE international conference on consumer electronics (ICCE), pp 1–2. https://doi.org/10.1109/ICCE47841.2021.9369637
14. Singh SK, Singh SP, Pal K (2021) Pothole detection using YOLOv7 and transfer learning. In: Proceedings of the 2021 IEEE international conference on inventive research in computing applications (ICIRCA), pp 1–6. https://doi.org/10.1109/ICIRCA52089.2021.9391248
15. Lee J, Kim Y, Lee S (2021) A pothole detection and tracking system using YOLOv7 and kalman filtering. In: Proceedings of the 2021 IEEE international conference on consumer electronics (ICCE), pp 1–2. https://doi.org/10.1109/ICCE47841.2021.9369696
16. Saini SS, Kumari N, Goyal A (2021) Pothole detection and segmentation using YOLOv7 and mask R-CNN. In: Proceedings of the 2021 international conference on computational techniques, electronics and mechanical systems (CTEMS), pp 245–249. https://doi.org/10.1109/CTEMS52056.2021.9478173

The Contributions of Technologies and ICT Tools for Teaching and Learning Languages

Pandit B. Nirmal and Bassam Hameed Mohsen

Abstract ICT tools have brought a significant change in the way we teach and learn languages. These tools have made language learning more accessible, enjoyable, and interactive. By selecting the right tool and creating a regular learning schedule, students can achieve their language learning goals with ease. The present study seeks to introduce and classify the different types of ICT tools that can be used in learning languages. The advantages, features and descriptions of several ICT tools are also introduced and how they can be utilized for learning different languages.

Keywords Tools · Classification · Technology

1 Introduction

With the development of technology and the Internet, websites in all languages expanded, but the majority of these websites were in English, making learning English a requirement for everyone.

The introduction of numerous language learning websites and applications has helped people. Due to their primitive nature, language learning programs and websites did not initially produce adequate educational results. However, as technology advanced, specialized language learning programs appeared that enable individuals to learn any language as if they had spent years studying it in classrooms and institutes [1].

For an enormous number of people, especially those who are dedicated to work and occupations, learning a language online and through apps helps to provide alternatives to master the language.

P. B. Nirmal (✉)
Department of English, Sant Tukaram College of Arts and Science, Parbhani, India
e-mail: panditbnirmal@gmail.com

B. H. Mohsen
Department of English, Hodiedah University, Hodiedah, Yemen
e-mail: bassamhameed20@hoduniv.net.ye

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_37

A person can study any language in their leisure time using these websites and applications. Additionally, these websites and programs make it possible for individuals who lack the financial means to attend language-learning institutions to quickly learn and master the language. A person can now immediately develop online courses, register for them, and communicate with other students from around the world. This has become possible by modern technology [1].

The tools of modern technology that have emerged for language learning are an additional and auxiliary factor for students to finish language learning and improve the abilities they gain in the classroom. Technology has made it easier for teachers and students to acquire current and authentic language resources and follow language learning programs. Additionally, it offers the chance to speak with native speakers, which gives the person excellent conversational experience. Additionally, it helps learners receive rapid guidance from their educational programs, enabling them to address mistakes they make right away [2].

2 Language Learning and ICT

Learning new languages has always been an exciting and challenging journey. The ability to communicate with people from different cultures and backgrounds opens up new opportunities and enriches our personal and professional lives. Modern times call for modern methods, and incorporating technology and ICT tools in teaching and learning languages has become more important than ever.

Effective language learning requires a lot of attention and practice, and technology offers various means to make this process enjoyable and interactive. With the help of language learning apps, online language courses, and multimedia tools, students can learn at their own pace and personalize their learning experience according to their unique needs and interests. This makes language learning more flexible and accessible to a wider audience, breaking down the barriers of time and location.

Gone are the days of tedious and boring language drills, as technology offers interactive and engaging exercises such as gamified language learning, virtual reality simulations, and video conferencing tools for live practice with native speakers. This makes language learning more interesting and fun, motivating students to stay on track and achieve their goals.

The use of ICT tools also empowers language teachers to create dynamic and relevant learning material tailored to their students' needs. They can design online quizzes, interactive presentations, and project-based activities that promote critical thinking and practical application of language skills. This allows teachers to harness the full potential of technology to make language learning more engaging and effective Coursera [3].

Moreover, technology and ICT tools enable language learners to immerse themselves in the language and culture they are studying outside traditional classrooms. Social media, language exchange platforms, and online communities offer learners the opportunity to connect with native speakers and fellow language learners from all over the world, practicing and improving their language skills in a real-life context.

3 Language Learning and Technology

Technology use has become a critical part of learning both inside and outside of the classroom. Language learning has gained a lot from and been enhanced by the use of technology. Technology is becoming more and more crucial as a tool for teachers to assist students in learning languages.

Children nowadays have grown up in a completely computerized world where the Internet, electronics, video games, and other IT technologies have embedded themselves into every aspect of their lives. Furthermore, this new generation, which differs significantly from earlier ones in terms of their technological prowess, has been recognized and named the "Net Generation." So, it seems quite obvious that young people who are used to spending a lot of time using technology, playing video games, and browsing the Internet would have their motivation significantly increased if they had the option to use technology tools in their learning whenever they approached any subject [4].

4 ICT Tools Uses and Advantages

The use of technology in teaching languages is justified for a number of reasons. The ability to work independently or engage and collaborate with others is one of the ways that technology may be inspiring. Additionally, technology offers immediate feedback on how well users of language perform in various assignments and activities. Furthermore, technology can be a time- and space-saving addition to the classroom. Finally, technology can encourage language acquisition by providing up-to-date, authentic, and interesting materials that can be accessed immediately from the Internet [5].

By offering a variety of ICT instruments for language teaching and learning, technology has revolutionized the field of education. For students all over the world, these tools have improved the learning process and made it more interactive, personalized, and engaging. If we consider the context, technology has enhanced access to language acquisition. Regardless of their location or the accessibility of physical resources, students can learn a new language with the aid of numerous ICT technologies. This has created new opportunities for people who want to learn a new language but have no access to physical classes or language institutions.

It's essential to maintain an attention-grabbing and engaging tone when it comes to content. ICT tools have therefore improved the fun and interaction of language learning for learners of all ages. Students can learn a new language through quizzes,

Table 1 Advantages of using ICT tools for language learning [6–8]

Benefit	Explanation
Convenience	Accessible anytime and anywhere with Internet connection
Personalization	Learning content and pace can be tailored to individual needs
Interactivity	Engaging features like games and quizzes encourage learning
Feedback	Instant feedback helps learners identify areas for improvement
Motivation	Gamification and progress tracking can boost motivation
Real-world context	Authentic materials provide exposure to real-life language use
Benefit	Explanation
Convenience	Accessible anytime and anywhere with internet connection

games, and other interactive techniques with the aid of ICT applications like Duolingo and Babbel.

The advantages of using ICT technologies for learning languages must be emphasized in order to keep students reading the entire piece of material. These resources help students study in an interactive environment and efficiently retain the material they have learnt. Students may easily practice their language abilities through features (Table 1).

5 Classification of ICT Tools

ICT tools can range from simple apps and software to advanced interactive whiteboards and learning management systems. If you're looking to augment your language learning experience, there are plenty of options to choose from. For example, you could use language learning apps such as Duolingo or Babbel, both of which offer interactive and engaging lessons that cover everything from grammar to vocabulary. Other ICT tools for learning a new language include online language exchange platforms like iTalki, where you can connect with native speakers and practice speaking in real-time. Similarly, video conferencing tools like Zoom or Skype can be used for virtual language lessons with a tutor or conversation partner [9].

Aside from language-specific tools, there are also general productivity tools that can be useful for language learning. For example, note-taking apps like Evernote can help you keep track of new vocabulary words, while flashcard apps like Anki can help you memorize them more efficiently. For language teachers, ICT tools can also be incredibly helpful. Learning management systems like Moodle or Canvas can be used to organize course content and activities, while interactive whiteboards like SMART and Promethean can be used to create engaging and interactive classroom presentations.

In the following, various ICT tools will be listed, each one will be categorised according to its function. The classification system will be according to the one developed by Erben et al. [9], which focuses on the various ICT tools' functions.

5.1 E-Communication ICT Tools

There are two types of software that encourage interaction among learners and teachers. First, there are the synchronous (i.e., real-time) software types that support activities like voice conferencing, board meetings, instant messaging, and telephone conversations. The second type of software consists of time-delayed (or asynchronous) applications including discussion boards, email, and text messages sent over mobile devices.

5.2 E-Creation ICT Tools

These tools promote creativity in English language learners. They involve employing language to create, investigate, and discover things while producing content and developing measurable learning outcomes. Web publishing, exercise creating tools, presentation software, cameras, podcasting, video makers, and audio producers are a few examples of these technologies.

5.3 Assessment Tools

Performances, portfolios, and projects are the three main performance assessment types. These methods of assessment differ significantly from regular standardised examinations in that the learner "produces evidence of accomplishment of curricula objectives" while using alternative assessment instruments. This proof is presented as a performance, a project, or a portfolio and can be "archived and used at a later time alongside other learning evidence elements as a compilation of evidence to show achievement" Erben et al. [9: 153].

5.4 Reading/Writing E-Tools

Writing and reading skills can be taught and learned more effectively and efficiently with the help of several ICT tools. Online message boards, online journals, wikis, blogs, and e-books are just a few examples of these tools. Digital content or online portfolios can be created using wikis, blogs, boards, and journals. They serve as spaces where students collaborate on projects. They can also be used to summarize student work or for professional development. These tools include Wikispaces, Blogger, and Penzu.

E-books offer a variety of functions, too. They can help students improve their reading skills. For instance, audio, interactive activities, and built-in dictionaries have all been incorporated to e-books. Another benefit of using this tool is that students can access e-books on many kinds of devices, including tablets, smartphones, and laptops.

5.5 *Virtual Learning Environments*

Web-based platforms known as Virtual Learning Environments (VLE) enable teachers and learners to manage and organize their work online. Here teachers, learners, and parents can access these spaces to track their progress (Tables 2 and 3).

Table 2 Language learning apps [10, 11]

Rank	ICT	Description
1	Duolingo	Gamified language learning app
2	Memrise	Flashcard-based language learning app
3	Rosetta stone	Language learning software
4	Babbel	Language learning app and website
5	Lingoda	Online language school
6	Anki	Flashcard-based language learning app
7	Busuu	Language learning app and website
8	Quizlet	Flashcard-based language learning app and website
9	Mango languages	Language learning app and software
10	FluentU	Language learning app with authentic content

Table 3 Comparison of language learning apps [12, 13]

App	Price	Features
Duolingo	Free	Gamified, lessons, practice, chatbot, streaks, stories
Memrise	Free/paid	Flashcards, videos, chatbots, pronunciation practice
Rosetta stone	Paid	Immersion, speech recognition, adaptive learning
Babbel	Paid	Lessons, speech recognition, review feature
Lingoda	Paid	Online classes, flexible schedules, individual feedback
Busuu	Free/paid	Lessons, chat with native speakers, certification exams
Mango languages	Paid	Lessons, conversation practice, cultural insights
FluentU	Paid	Authentic videos, interactive captions, personalized feedback

6 Conclusion

In conclusion, incorporating technology and ICT tools in teaching and learning languages is not just a trend but a necessity in today's fast-paced world. The benefits it offers to both teachers and learners are invaluable, making language learning more accessible, engaging, and effective. As we continue to advance in technology, the possibilities for language learning are endless, and we need to embrace it to create a more connected and multilingual world.

References

1. Motteram G (2013) The benefits of new technology in language learning. Retrieved from https://www.britishcouncil.org/voices-magazine/the-benefits-new-technology-language-learning
2. Mohsen BHM, Nirmal PB (2020) Towards the accommodation of EIL pedagogy in Yemeni English language teaching institutions. Int J Emerg Technol Innov Res 7(12):1085–1093
3. Coursera (n.d.) Language learning. Retrieved 19 April 2023, from https://www.coursera.org/courses?query=language%20learning
4. Nayebi R, Moghaddam MN (2017) Integrating technology into language learning classrooms: opportunities and challenges. J Appl Linguist Lang Res 4(2):147–159
5. Prensky M (2001) Digital natives, digital immigrants. On Horizon 9(5):1–6
6. Mohsen BHM, Nirmal PB (2022) Linguistic imperialism: viability of the use of native-like accent among English language learners. Stud Human Educ 3(2):32–37
7. Duolingo. https://www.duolingo.com/. Last accessed 2 April 2023
8. Memrise. https://www.memrise.com/. Last accessed 2 April 2023
9. Erben T, Ban R, Castañeda M (2009) Teaching English language learners through technology. Routledge, New York
10. Rosetta Stone. https://www.rosettastone.com/. Last accessed 4 April 2023
11. Sharma P, Barret B (2007) Blended learning: using technology in and beyond the language classroom. Macmillan
12. Godwin-Jones R (2018) Emerging technologies for language learning. Lang Learn Technol 22(1):1–2
13. Warschauer M, Matuchniak T (2010) New technology and digital worlds: analyzing evidence of equity in access, use, and outcomes. Rev Res Educ 34(1):179–225

Graphical-Based Password Authentication

Nikhil Patil, Ganesh Bhutkar, Priyanshi Patil, Parth Pishte, and Apoorva Popalghat

Abstract Traditional alphanumeric passwords have various benefits over graphic-based password authentication, including better user experience, increased memorability, and stronger security against typical password assaults. The most recent gesture-based approaches for graphical password authentication that are image-based and hybrid are all thoroughly reviewed in this paper. We have collated the security and usability traits of each method and have evaluated its advantages and disadvantages. Additionally, in order to increase security and usability, we have suggested a brand-new graphical-based password authentication system that integrates image-based and gesture-based techniques. Our experimental findings have depicted that the suggested plan is workable and efficient. Overall, this study has incremented the knowledge of graphical-based password authentication and has offered useful information to researchers and industry professionals.

Keywords Password · Authentication · Graphical · Images · Security

N. Patil (✉) · G. Bhutkar · P. Patil · P. Pishte · A. Popalghat
Vishwakarma Institute of Technology (VIT), Pune, India
e-mail: patil.nikhil21@vit.edu

G. Bhutkar
e-mail: ganesh.bhutkar@vit.edu

P. Patil
e-mail: patil.priyanshi21@vit.edu

P. Pishte
e-mail: parth.pishte21@vit.edu

A. Popalghat
e-mail: popalghat.apoorva21@vit.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_38

1 Introduction

As the Technological advances become more integrated into our daily lives, the need for secure authentication methods becomes increasingly important. Traditional password authentication methods have been in use for decades, but they have a number of flaws, including a susceptibility to brute-force, dictionary, and shoulder-surfing assaults [2]. Since they are not only more secure than text-based passwords, but also simpler to remember and use, the graphic-based password authentication systems have become a potential option.

The idea behind graphic-based password authentication systems is to choose images or shapes to create a password. In contrast to traditional text-based passwords, graphical-based passwords offer several advantages such as enhanced security, usability, and memorability [10]. Furthermore, graphical-based passwords offer a range of design possibilities that can make them more difficult to crack than traditional passwords.

The purpose of this research paper is to examine and weigh the benefits and drawbacks of graphical password authentication techniques. The paper has provided an analysis of the effectiveness of graphical-based password authentication systems in terms of security, usability, and user acceptance. Additionally, we have looked at the various design approaches applied in graphical password authentication systems and have offered suggestions for their application.

Overall, this research paper aims to provide a comprehensive overview of systems for graphical password authentication and their potential to offer a more secure and user-friendly alternative to traditional text-based passwords.

2 Research Methodology

2.1 Utilization of Components

Following components are used for implementation of proposed system:

- HTML and CSS—To create and design web pages.
- JAVASCRIPT—To implement image segmentation.
- XAMPP—To design and implement a database.
- GOOGLE RECAPTCHA—To control spamming of bots.
- Laptop having Internet connectivity—To implement the project.

2.2 System Design

The HTML web pages have collected the dates from the users store in the database. For example, here we are using XAMPP to store users' data. The images are stored in terms of values in the database, which also reduce space complexity and retrieval time. JavaScript is used for image segmentation. CSS helps make user-friendly interfaces and attractive designs, which appeal to mass users.

2.3 System Flow Chart

In this research work, when any user tries to access the system, he/she will be provided with three options as 'Register', 'Home Page', and 'Contact Us' page as seen in Fig. 1. If the users are not registered with system, then they have to register by following these instructions:

1. Just click on 'Register' and fill the fields like Username, Email, and Alphanumeric Password.
2. Select the Image which you want to be used as a Password.
3. Select the 3 Pixels of the previously selected image.
4. Click on the 'Register' button available on the page to get registered with the system.
5. Users will be taken to the 'Home Page' for login after registration as seen in Fig. 1.
6. For users to log in, they must input their login information.

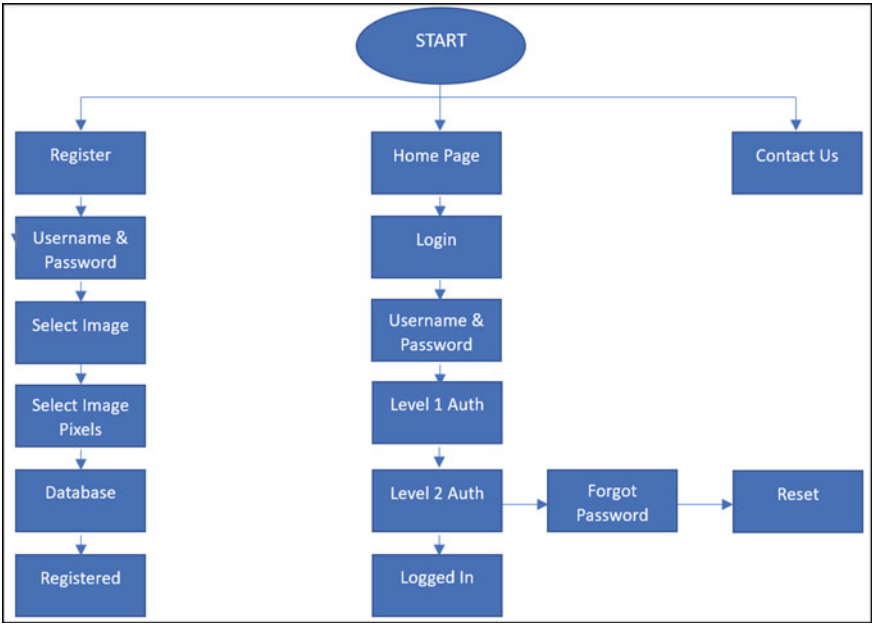


Fig. 1 Application flow chart

3 Literature Review

The overview of graphical passwords and a list of the several graphical password schemes that have been suggested in the literature are provided in the initial papers [1, 2]. The authors have identified the strengths and weaknesses of each scheme. The authors have studied various types of graphical password schemes, such as recognition-based, recall-based, and cued-recall-based schemes. It has also highlighted the challenges and open research issues in the field of graphical password authentication.

The next paper reviews the state of the art in graphical password research, including a historical overview of graphical passwords, a summary of existing schemes, as well as a discussion of the security and usability problems with graphical passwords [3]. The usage of picture passwords as an alternative to conventional text-based passwords is discussed in the paper [4]. It highlights the strengths and weaknesses of this approach and compares it to other authentication methods.

In another paper, the design and assessment of pass-points, a graphical password system, are discussed. According to a longitudinal study conducted by the authors, pass-points is more user-friendly and secure than conventional text-based passwords [5]. The next study suggests pass-shapes, a graphical password system that authenticates users using hand movements. The scheme offers robust security with good usability [6].

The next interesting paper provides a comprehensive overview of the issues surrounding password authentication on the web, including an analysis of the technical and market failures that have led to the proliferation of insecure password practices [7]. Another paper examines user attitudes towards graphical passwords, based on a survey of 120 participants. The authors identify user preferences for certain types of images and discuss the implications of these findings for the design of graphical password schemes [8].

The findings of a usability research on click-based graphical passwords are presented in the next study. The authors demonstrate that click-based graphical passwords are more secure than conventional text-based passwords, but they also caution that they might be less convenient to use [9]. The next paper provides an overview of graphical passwords and their potential synergy with text-based passwords. It discusses the strengths and weaknesses of graphical passwords and examines how they can be combined with traditional alphanumeric passwords to improve security [10].

The next study examines the viability of graphical authentication systems. The authors provide the findings of a user survey that demonstrates that while graphical passwords may be less safe, people find them to be more entertaining and less stressful than text-based passwords [11]. Another paper presents a survey of graphical password authentication schemes. It presents a comparison of the various techniques and classifies the many kinds of graphical passwords. The authors also discuss the strengths and weaknesses of graphical passwords [12].

Another study has investigated the usability and security of image-based authentication. According to the authors' user study, image-based authentication is more reliable that text-based passwords, but it may also be less practical [13]. In next paper, graphical passwords are contrasted with conventional alphanumeric passwords. The authors analyse the effect of human factors on password security as well as the usability and security of several graphical password schemes [14]. The last study assesses the security and dependability of 'secret' questions, a backup authentication method that is frequently employed [15].

4 Result and Discussions

Our findings suggest that graphic-based passwords can be an effective and user-friendly alternative to text-based passwords. The fact that graphic-based passwords are harder to decipher or guess than text-based passwords accounts for the greater success rates for authentication. In addition, participants have reported finding the graphic-based password more enjoyable and easier to use, which could lead to increased user satisfaction and engagement with online systems.

It is worth noting that some participants expressed concern about the security of their graphic-based passwords, specifically with regards to the risk of someone looking over their shoulder and memorizing their pattern. However, this risk can be mitigated by implementing strategies such as displaying the pattern in a randomized order each time, or by allowing users to choose their own personalized images.

Overall, our study suggests that graphic-based passwords can offer a secure and user-friendly alternative to text-based passwords, and could be a promising option for future use, with online authentication. Further study is required to fully understand the potential limitations and vulnerabilities of graphic-based passwords, as well as to investigate the potential for combining different authentication methods for enhanced security. Here are a few of the screenshots of the system (Figs. 2, 3, 4 and 5).

5 Conclusion

In conclusion, a graphic-based password authentication is an innovative approach that has the potential to enhance the security and usability of authentication systems. This approach has been shown to be effective in improving user satisfaction and memorability, and certain types of graphic passwords have been found to be more secure than traditional alphanumeric passwords.

However, there are also challenges associated with graphic-based password authentication, including the risk of users selecting easily guessable or replicable images, the potential for shoulder-surfing attacks, and difficulties experienced by some users with physical limitations or visual impairments.

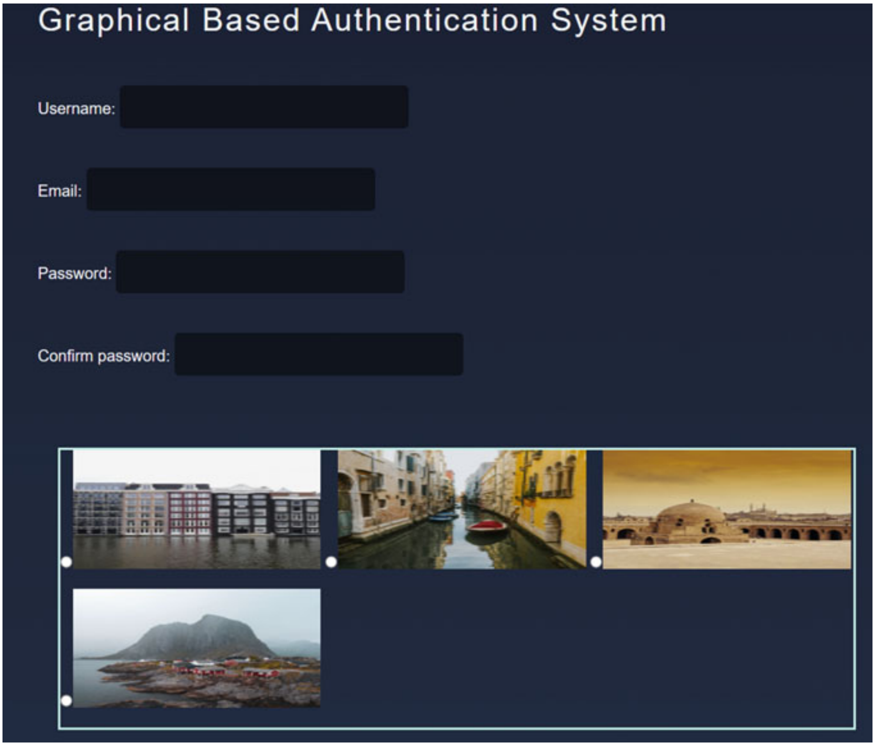


Fig. 2 'Register' page

To address these challenges, ongoing research and development is needed to identify and implement solutions such as advanced image analysis techniques, multi-modal authentication, enhanced usability, accessibility for all users, and integration with emerging technologies.

Overall, the construction of more secure systems may benefit greatly from the use of graphic-based password authentication and user-friendly authentication systems. With ongoing research and development, it is likely that we will see even more innovative and effective graphic-based authentication methods in future.

6 Future Scope

One potential future scope is the development of advanced image analysis techniques that can help to identify and prevent attacks on graphic-based password authentication systems. These techniques could include machine learning algorithms that can recognize patterns in images and detect attempts to guess or replicate passwords. Another potential future scope is the use of multi-modal authentication,

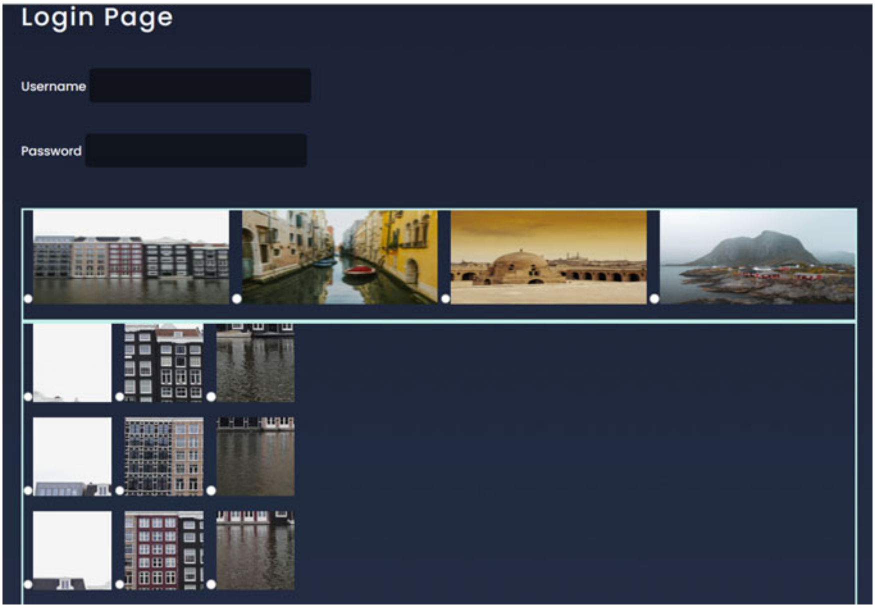


Fig. 3 'Login' page

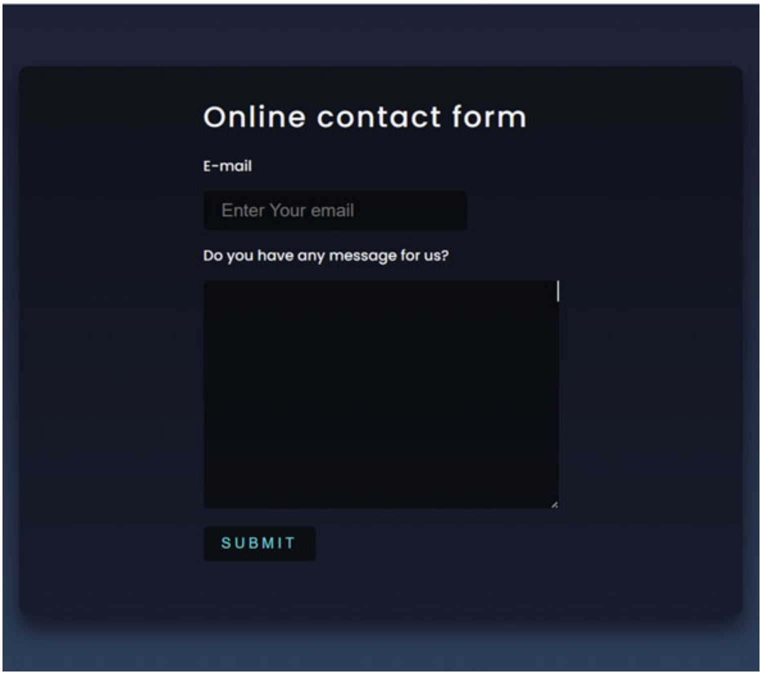


Fig. 4 'Contact Us' page

←T→				id	username	email	password	radio	radio1	radio2	radio3
☐	Edit	Copy	Delete	3	om	om@gmail.com	om	1	12	14	16
☐	Edit	Copy	Delete	4	parth	123@gmail.com	123	1	12	14	16
☐	Edit	Copy	Delete	12	thor	thor@gmail.com	575e22bc356137a41abdef379b776dba	1	10	13	16
☐	Edit	Copy	Delete	13	omkar	omkar@gmail.com	68115705108ddd6f7ba9a458bb175363	3	32	34	36
☐	Edit	Copy	Delete	14	buttler	buttler@gmail.com	8fd362c8d23130ec6b4a2f6992d83e41	1	10	13	16

Fig. 5 Application database

which combines graphic-based passwords with other authentication factors, such as biometric data or one-time passwords. This approach can increase the security of the system by requiring attackers to compromise numerous security measures to obtain entry. Another scope is using computer vision integrated with it, to make it more optimized. Small-scale systems like Google CAPTCHA also can be made by this concept which can be integrated on any existing as well as new models.

Acknowledgements We thank Prof. Dr. Sandeep Shinde-Head, Department of Computer Engineering, VIT, Pune, for his valuable guidance and support to this research work.

References

1. Agrawal A, Stoller S (2011) Graphical passwords: a survey. J Inf Secur 2(1):31–44
2. Biddle R, Chiasson S, van Oorschot P (2012) Graphical passwords: learning from the first twelve years. ACM Comput Surv 44(4):19
3. Birget J, Hong Y (2008) Passpoints: design and longitudinal evaluation of a graphical password system. Int J Hum Comput Stud 66(9):614–625
4. Bonneau J, Herley C, van Oorschot P (2015) The password thicket: technical and market failures in human authentication on the web. Commun ACM 58(7):74–81
5. Chiasson S, Biddle R, van Oorschot P (2008) A second look at the usability of click-based graphical passwords. In: Proceedings of the 2008 ACM conference on computer and communications security, pp 187–196
6. De Angeli A, Coventry L, Johnson G, Renaud K (2005) Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. Int J Hum Comput Stud 63(1–2):128–152
7. Dhamija R, Perrig A (2000) Deja Vu: a user study using images for authentication. In: Proceedings of the 9th USENIX security symposium, pp 1–14
8. Egelman S, Peer E (2015) It's no secret: measuring the security and reliability of authentication via "secret" questions. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp 1413–1422
9. Jan SU, Zhang Z, Asif M (2015) A survey on graphical password schemes
10. van Oorschot PC, Camenisch J (2012) Picture passwords: a new direction in authentication
11. Boukhechba M, Traynor P, Butler K (2013) PassShapes: using gestures to authenticate mobile users
12. Dunphy P, Briggs P, Sharif E (2012) Visual passwords: a survey of user attitudes

13. Uzunoglu S, Selcuk Uluagac A, Beyah R (2015) A survey on graphical passwords and their synergy with text-based passwords
14. Liu X, Brown MS, Qi Y (2013) Graphical password authentication: a survey
15. Zhang Z, Jan SU, Asif M (2015) A comparative study of graphical passwords and traditional passwords

Survey on IoT-Based Smart Irrigation System

Avinash Golande, Swanand Mulay, Swasti Rai, Laukik Gadam, and Nishant Patil

Abstract Agriculture has now been engaged for over 12,000 years, and the Neolithic era gave birth for civilization. Farming is being currently practiced as conventional farming practices. Rains, soil, humidity, and environmental concerns are all major factors in Indian agriculture. In agriculture, our farmers switched to new state-of-the-art technology. With the help of sensors which are interfaced to the process module Arduino UNO, advanced agriculture can forecast meteorological data, powering on the pump motor, and acknowledging the moisture of the soil in terms of moisture levels. With the use of networking technology, the smart farm system may be controlled from anywhere. When research and development in smart agriculture and artificial intelligence are integrated, cutting-edge technologies in data compiling and resource optimization can evolve. Professionals and farmers control the time and data collected. The cloud acts as a central digital data repository for audio, video, images, text, and digital maps collected from various sources. An artificial intelligence (AI)-based machine learning model, one of many types of classification, was used to analyze all the data. This study focuses on how cloud computing technology can help create smart agriculture.

Keywords Internet of Things · Security · Artificial intelligence · Cloud computing

1 Introduction

Smart agriculture uses a variety of technology, equipment, protocols, computational models for enhancing agricultural processes. Big data, artificial intelligence, cloud computing, and edge computing all offer capabilities and solutions for analyzing, processing, and sorting massive amounts of data produced by components. In

A. Golande (✉) · S. Mulay · S. Rai · L. Gadam · N. Patil
JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra 411033, India
e-mail: avinash.goalnde@gmail.com

L. Gadam
e-mail: laukikgadam@icloud.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_39

contrast, smart agriculture is still in its early phase and has limited security features. To help farmers, future solutions will rely on data accuracy and availability, and security will be important in building reliable and efficient systems. Because smart agriculture makes use of a wide range of resources, security must handle challenges such as compatibility, restrictions, resource and massive data.

Agriculture Phase 1: Agriculture was dominated by human and animal resources throughout the traditional agricultural age (from 1784 and about 1870), and the major worry was low operational efficiency.

Agriculture Phase 2: The fundamental challenge throughout the mechanized agricultural period (the twentieth century) was resource inefficiency. autonomous agriculture (from 1992 to 2017) was the low level of intelligence.

Agriculture Phase 3: The main issue throughout the era of rapid growth of autonomous agriculture (from 1992 to 2017).

Agriculture Phase 4: The smart agriculture phase (which begins in 2017 and is characterized by autonomous operations) is distinguished by the use of modern information technology to both service and intelligently grow agriculture. Smart agriculture is a new agricultural production mode that combines modern information technologies such as the Internet, Internet of Things (IoT), big data, cloud computing, and artificial intelligence (AI) with agriculture to aid in agricultural information perception, quantitative decision-making, intelligent control, precise investment, and personalized service. In a word, the new mode is an intelligent agricultural solution that combines agriculture with current information technology.

Smart farming incorporates a variety of technologies, sensors, protocols, and computer paradigms to help farmers maximize the benefits of these technologies. The “agricultural revolution” (digital) will alter all elements of agriculture, resulting in systems that are sustainable, more productive, transparent, inclusive, efficient, and resilient. It continues to embrace new era technologies that enable greater and remote crop and livestock monitoring. Unsupervised networks of information are created through the interconnection of these technologies within a single farm or production facility, as well as data interchange between suppliers and vendors. Increased dangers of cybersecurity assaults on farms and agribusinesses accompany the deployment of these technologies. These attacks may disrupt food supply systems, causing harm to the bioeconomy and populations. Agricultural security comprises good cybersecurity and biosecurity decisions, important control points, and human behavior and habits that affect overall security. Cyber biosecurity is the name given to the combination of these areas. While cybersecurity covers the security concerns of any networks electronic data, systems, and the like, cyber-bio-security is one of its most important applications, focusing on the prevention of illegal intrusions and other activities, as well as the protection of data, information, and other online resources related to all the aspects of life. Health, medical, agriculture, and food sciences.

2 Technologies Used

2.1 Microcontroller

A microcontroller can be programmed to interact with the physical world by using various sensors, actuators and other electronic components. There are many vendors which provide microcontrollers with different architecture, processing powers, and thus cost, like Arduino, Espressif ESP32, Raspberry Pi Pico, etc. It is most crucial part for our system, so we should choose one which is cost efficient while satisfying the requirements of our system. We can use memory safe language like Rust for programming the microcontroller which will make our system secure and robust (Figs. 1 and 2).

Fig. 1 NodeMCU

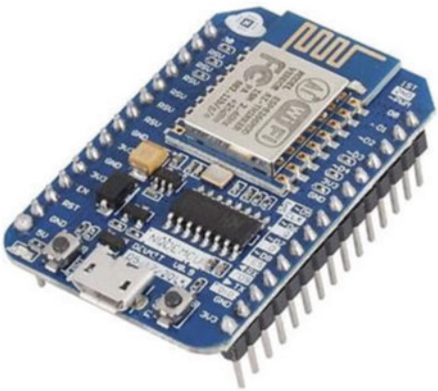


Fig. 2 Arduino

2.2 Relay

A relay is an electronic component that functions as a mechanical switch using an electromagnet. Relays are mostly used to turn on and off high-powered circuits from low-powered ones. We must ensure that the relay is not restricted to operating below these specifications if, e.g., you intend to use it on a 240-V circuit with 30 A of current.

2.3 Sensors

1. Optical Sensors in Agriculture: It uses multiple light frequencies to measure the qualities of the soil. These sensors, which are mounted on cars or drones, enable the collection and processing of information about plant color and soil reflectance.
2. Electrochemical Sensors-This type of sensor, which is frequently used to both monitor and analyze soil quality, takes actions to change the pH level or keeps using the same procedures to keep it at the existing level for the remaining stages of any crop's lifetime.
3. Mechanical Soil Sensors for Agriculture-When a sensor penetrates the soil, it measures the holding forces brought on by the soil's cutting, breaking, and displacement. The ratio of the force needed to enter the soil medium to the frontal area of the tool engaged with the soil is known as soil mechanical resistance.

3 Mathematical Models

3.1 Support Vector Machine (SVM)

Using labeled training data, SVM can produce a hyperplane that differentiates the new unlabeled data. Then, training set $(y_i; x_i)$; $i = 1, 2\ldots, n$ where $x_i \in R\ n$ represents the itch vector and $y_i \in R\ n$, represents target element. The linear SVM, which is used in this survey, generates the optimum hyperplane in form:

$$f(x) = w^T * x + b$$

3.2 Neural Network

To interpret the neural network, we have to first initiate the linear model specified as:

$$g(x, w) = \sum_{i=1}^{P} w_i f_i(x)$$

where vector w is the parameter vector of the model, and the functions $f(x)$ are the non-constant functions for the variable x. Neural networks are included in the class of models that are nonlinear in their parameters. The most used form of static neural network is the easy addition of the preceding relationship:

$$g(x, w) = \sum_{i=1}^{p} w_i f_i\left(x, w'\right)$$

4 Proposed System

MCU will turn the motor ON/OFF based on the gathered data or when the user manually selects to do so. There will be authorization and authentication in place to avoid the third party threats and mischievous activity. Our platform will also provide the option to visualize in graphical format and to the gathered data on the server for using it in AI models. The data will be sent encrypted to avoid Man in The Middle attacks.

5 Literature Survey

See Fig. 3 and Table 1.

6 Related Work

- In February 2022, Parvathi Sangeetha [1] created intelligent agriculture management system to generate agricultural benefit and production of crop. The composite remote-control device uses GPS with the combination of Radial Function Network.

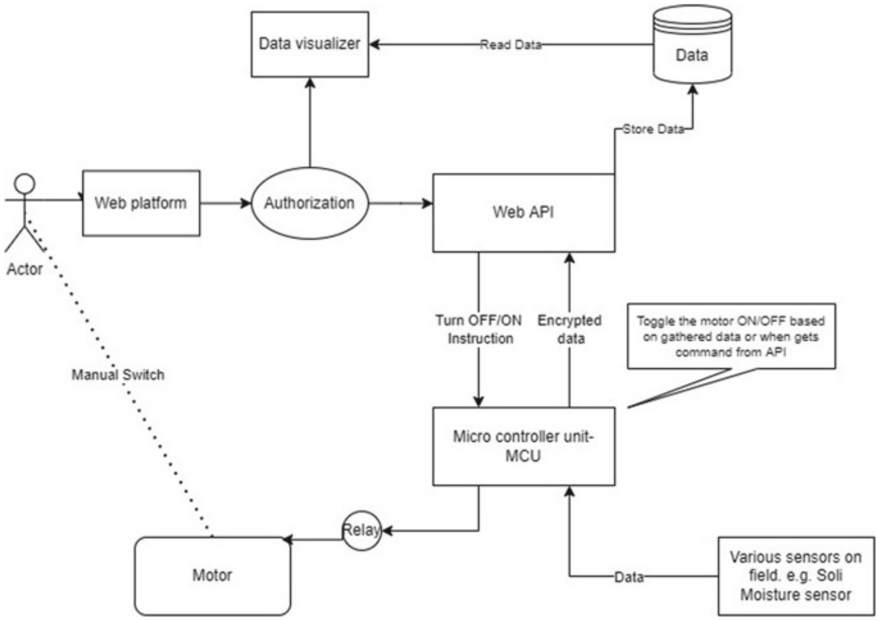


Fig. 3 System architecture

- In May 2021, Stolejescu-Crisan et al. [2] proposed a smart gardening system which used an IoT framework known as q-Toggle. q-Toogle is a system in which actuators, sensors and other sources of data based on open source solutions. This system reduces water and energy consumption.
- In May 2021, Rohith [3] made a system which makes use of humidity, moisture, and temperature sensors. They set a parameter for every sensor and if values decrease then automatically motors turns on to water the plants. They used Arduino board, regulators of voltage and relay, and these components control motor.
- In February 2021, Rezwan [4] proposed a system which works some parameters. And these parameters decide the operation of system. The system takes actions such motor On/Off depending on these parameters. They build an app for users which controls this system remotely.
- In December 2020, Mantri [5] proposed a system of smart irrigation which uses pH of water and also uses sensors of soil moisture. The pH of soil's moisture and water is monitored continuously and the data collected is uploaded to cloud and then processed using microcontrollers.
- In October 2020, Nughara et al. [6] proposed a system which remotely controls water pump and monitors the soil moisture in the garden. In this system, the farmer can measure and detect the soil moisture which help them to manage the water usage.

Table 1 Summary of literature survey

No.	Year	Title	Summary
1	2022 [1]	IOT-based smart irrigation management system for environmental sustainability in India	Created intelligent agriculture management system to produce agricultural benefit and crop production
2	2021 [2]	An IoT-based smart irrigation system	Proposed a smart gardening system which used an IoT framework known as q-Toggle
3	2021 [3]	IoT-enabled smart farming and irrigation system	Arduino board controls the motor, relay, and voltage regulators
4	2021 [4]	An IoT-based smart irrigation system	Uses different parameters to operate the system automatically
5	2020 [5]	IoT-based smart watering system toward improving the efficiency of agricultural irrigation	Uses pH of water which is used for irrigation. The app is developed for farmers to control this system
6	2020 [6]	IoT-based smart garden irrigation system	Remotely controls the water pump
7	2020 [7]	Smart irrigation system using Internet of Things	Waters plants automatically with the help of dampness of soil
8	2019 [8]	Smart irrigation system using Internet of Things	Uses Arduino Mega 2560 that processes a data from soil sensor that waters the plant automatically
9	2019 [9]	Design and implementation of IoT-based irrigation system	Control system monitors various parameters and data is uploaded to cloud system
10	2018 [10]	Automated irrigation system IoT-based approach	Arduino board is connected to Wi-Fi and soil moisture sensor. This framework is connected to cloud
11	2017 [11]	Intelligent irrigation system—an IOT-based approach	Turns On and Off the water pumping motor automatically on basis of dampness of soil
12	2017 [12]	IOT-based smart irrigation system	Uses moisture sensor to detect moisture of soil and waters accordingly

- In April 2020, Anitha et al. [7] proposed a system which identifies soil dampness and according to dampness it automatically controls the watering of plants which helps farmer to utilize the water properly.
- In November 2019, Leh et al. [8] proposed a smart IoT irrigation system that uses Arduino Mega 2560 which process a data received from soil sensor that waters the plant automatically. It analyzes real time soil condition of plants via smartphone. It also uses dht 11 sensor and soil moisture sensor which checks condition of a plant for irrigation.
- In September 2019, Babayigit and Buyukpatpat [9] proposed an IoT irrigation system which includes two subsystems, i.e., control system and cloud system. In this system, environmental conditions are monitored using the temperature sensor,

soil sensor and rain sensor and then this sensor data was wirelessly gathered. Then data is analyzed and according to instructions provided by cloud and then control system and water is provided.

- In February 2018, Mishra et al. [10] proposed a coded water system with framework for terrains if you want to lessen manual labor and optimizing water usage growing productiveness of vegetation. In this setup, Arduino is used with Wi-Fi module and moisture sensor. This setup is connected with framework of cloud and then aquitization of data is done and then this is analyzed using cloud services.
- In November 2017, they [11] proposed a system which is automated which features with the help of latest technologies in electronics by using microcontroller which ON and OFF the pumping motor by detecting the soil's dampness content and then GSM telephone line is operated after counting the humidity, temperature, and soil moisture.
- In February 2017, Shrishti Rawal [12] proposed an automatic irrigation system that uses automatic irrigation to monitor and control materials. The microcontroller ATMEGA328P and Arduino UNO to perform the control. With the help of GSM-GPRS SIM900A modem, farmer can see for whether the irrigation system is ON or OFF.

7 Conclusion

The purpose of this survey paper was to gain knowledge about the concepts of secure IoT-based agriculture systems, as well as the threats that the new paradigm of technology and modern ways poses. The rising complexity of Agri-product cultivation and production methods leads to an increase in possible hazards in numerous areas. However, information technology comes with a number of security issues. This paper examines the most recent research on smart agriculture. To be considered successful, a new method or system must be able to: (i) bring down costs; (ii) saving time; (iii) build trust; and (iv) minimize risk.

Our study focused on adding the security aspect to smart irrigation systems by means of authorization and authentication. Using memory safe language like Rust helps in avoiding the memory corruption bugs, making the system robust and secure.

This paper doesn't go over what algorithms will work best for encryption, authentication and authorization on embedded devices which tend to have low computational power. There is also scope in reducing the cost by using some existing IoT cloud service providers.

References

1. Parvathi Sangeetha B (2022) IOT based smart irrigation management system for environmental sustainability in India. https://doi.org/10.1016/j.seta.2022.101973
2. Stolojescu-Crisan C, Butunoi B-P, Crisan C (2021) An IoT based smart irrigation system. IEEE Consum Electron Mag. https://doi.org/10.1109/MCE.2021.3084123
3. Rohith M, Sainivedhana R, Sabiyath Fatima N (2021) IoT enabled smart farming and irrigation system. In: 2021 5th international conference on intelligent computing and control systems. https://doi.org/10.1109/ICICCS51141.2021.9432085
4. Naeem MRH, Gawhar S, Adib MBH, Sakib SA, Ahmed A, Chisty NA (2021) An IoT based smart irrigation system. In: 2021 2nd international conference on robotics, electrical and signal processing techniques. https://doi.org/10.1109/ICREST51555.2021.9331092
5. Balasooriya TN, Mantri P, Suriyampola P (2020) IoT-based smart watering system towards improving the efficiency of agricultural irrigation. In: 2020 IEEE global conference on artificial intelligence and internet of things. https://doi.org/10.1109/GCAIoT51063.2020.9345902
6. Sholihul Hadi M, Nugraha PA, Wirawan IM, Zaeni IAE, Mizar MA, Irvan M (2020) IoT based smart garden irrigation system. In: 2020 4th International conference on vocational education and training. https://doi.org/10.1109/ICOVET50258.2020.9230197
7. Anitha A, Sampath N, Asha Jerlin M (2022) Smart irrigation system using internet of things. In: 2020 International conference on emerging trends in information technology and engineering. https://doi.org/10.1109/ic-ETITE47903.2020.271
8. Leh NAM, Ariffuddin Mohd MS, Muhammad Z, Kamarzaman NA (2019) Smart irrigation system using internet of things. In: 2019 IEEE 9th international conference on system engineering and technology. https://doi.org/10.1109/ICSEngT.2019.8906497
9. Babayigit B, Buyukpatpat B (2019) Design and implementation of IoT based irrigation system. In: 2019 4th International conference on computer science and engineering. https://doi.org/10.1109/UBMK.2019.8907066
10. Mishra D, Khan A, Tiwari R, Upadha S (2018) Automated irrigation system-IoT based approach. In: 2018 3rd International conference on IoT. https://doi.org/10.17577/IJERTV12IS030112
11. Newlin Rajkumar M, Abinaya S, Venkatesa Kumar V (2017) Intelligent irrigation system-An IoT based approach. In: 2017 International conference on innovations in green energy and healthcare technologies (IGEHT). https://doi.org/10.1109/IGEHT.2017.8094057
12. Rawal S (2017) IoT based smart irrigation system. Int J Comput Appl. https://doi.org/10.5120/ijca2017913001

An Efficient Classification Method Using GLCM and Decision Tree Classifier

Punam Gulande and R. N. Awale

Abstract The treatment of lung cancer and related diseases is a major concern of medical science. Death due to lung cancer is seen as a leading cause and therefore early detection and treatment becomes a critical necessity. A commonly used technique of imaging to diagnose lung cancer is computed tomography (CT). In disease diagnosis, envisaging techniques play an important role. These methods help detect abnormal tissue-tumor patterns in targeted cancer cells. According to the World Health Organization, among all cancers, lung cancer accounts for around 14 percent. Paper contributes the different five stages as a proposed methodology. Firstly, database like normal lung and disease lung cancers images are collected. Median filter is used for pre-processing. Edges are preserves correctly by the median filter, so it is preferring, i.e., preservation of sharp features. In the third stage, segmentation on the target image is performed to ascertain and isolate the desired cancerous entity from the background. In stage four, features such as area, contraction, energy, entropy and homogeneity are mined using Gray Level Co-occurrence Matrix (GLCM). High demarcation precision and lower processing speed can be achieved by using GLCM. The fifth stage, these mined features are assigned to 5 different classifiers to classify lung cancer from normal lung. A decision tree (DT) classifier achieved highest accuracy for detection and lung cancer classification.

Keywords CT scan image · Segmentation · Decision tree (DT) · Medical image processing

P. Gulande (✉) · R. N. Awale
VJTI, Mumbai University, Matunga, India
e-mail: psgulande_p19@el.vjti.ac.in

R. N. Awale
e-mail: rnawale@el.vjti.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_40

1 Introduction

Cancer is a type, commonly seen as a type of tumor, found in the human lungs that breeds quickly and blowouts to other human organs. The cancer is said to be a growth of abnormal cell that has potential to harm the other normal human tissue cells. According to the data of histopathology, lung cancer cell is normally bifurcated into two types, Small-scale Cell Lung Cancer (SCLC) and Non-Small-scale Cell Lung Cancer (NSCLC). The Non-Small-scale Cell Lung Cancer (NSCLC) consists of squamous cells (ACC), adenocarcinoma (ADC) and large cells. Among these types of cancer, NSCLC is the type of cancer that constitutes for approx. 80–90% of deaths cases worldwide. Based on the pathological remains, they transform into benignant and fatal cancers. In many of the cases, the lung cancer is diagnosed at the final stage, which makes curing action difficult [10].

These numbers of deaths can be reduced by the very detection of cancer cells in the early stages. Imaging techniques are used as one of the methods for early identification of such cancer cells and can therefore assist for treating it at the early stages. The methodology proposed has been classified in five different phases for lung cancer classification. The first phase is collection of the data which is required for classification purpose. This database is taken from kaggle dataset. In the second stage, the median filter is used for denoising. In the third phase, segmentation is done on the captured input images. After segmentation, in phase four, for the purpose of extraction of features, GLCM is used. These features which has been extracted are used in the fifth phase for classification purposes that will be carried by different classifiers.

This work contributes as follows:

- Design and implementation of proposed algorithm for feature selection for the image dataset.
- Classification analysis is done using five different deep learning algorithms.
- The result as well as proposed method compare with GLCM + SVM Algorithm [14], GLCM + SVM and BPNN Algorithm [4], GLCM + ANN Algorithm [1] to show the performance in the form of tabular and graphical.

The organization of the paper is done in the particular manner. Section 2 is related work, whereas Sect. 3 describes the proposed methodology part. The proposed methodology section introduces the proposed algorithm along with the feature selection and classification mechanism. The results are presented in tabular and graphical forms with discussion in Sect. 4. Conclusion is as mentioned in Sect. 5.

2 Related Work

Anifah et al. cancer detection was performed using artificial neural network (ANN) backpropagation. The database of lung cancer is having total 50 images. These images are distributed into two types—normal images and infected images. Selection of features are selected using GLCM and input to ANN [1]. Mohamed et al. a new proposal was mentioned to facilitate the segmentation method. Many images are considered in the analysis of lung cancer and thus the segmentation process for this big amount of data becomes difficult. Therefore, probability-based graphical presentation used in Bayesian Network Model (BNMB) was chosen. Excessive segmentation is an image resulting from of graphical approach [2]. Patil and Jain the new methodology used computerized tomography (CT) images. No of different steps have been taken for pre-processing and denoised the images in this methodology. Excess segmentation is removed with the help of marker controlled watershed segmentation in the second step. Binarization and masking technique were applied on the segmented image. Achieved accuracy was 85.27% by using watershed to threshold [3]. Li et al. suggested method for pre-processing and denoising the CT image by applying filters as median filter and a Wiener filter for lung cancer detection. SVM and BPNN are used as classifiers. Features were selected using the GLCM technique and these were fed to the classifiers [4].

3 Methodology

The planning of proposed methodology presented in different five stages in the below Fig. 1 having main parts, namely computed tomography images as input, filtering, segmentation and extraction of feature for classification stage.

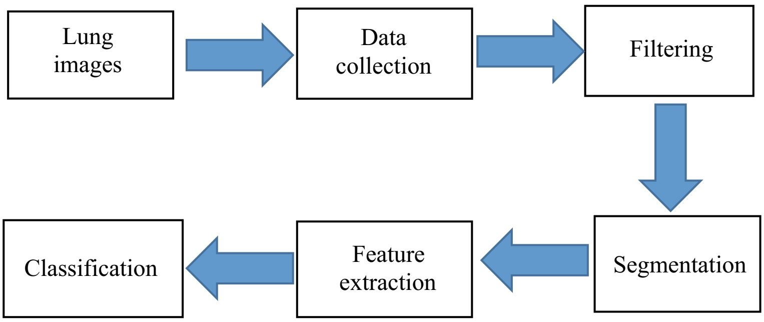


Fig. 1 Work flow of proposed method

3.1 Proposed Methodology

The proposed methodology is as per the below Fig. 1.

3.1.1 Collection of Database

Cancer and normal images required are possessed from Kaggle dataset and separated 70% for training and remaining for testing purpose. All Dicom format computed topography (CT) images are used. These images are an advance to X-rays so chosen here. All CT scans are viewed from different angles and are combination of X-ray images. Advantages of CT scans are like due to less noise and high resolution. In this proposed research, 120 and 700 images are used for analysis and detection of cancer. 120 samples are used with 3 features and out of 700 samples, 500 samples are cancerous and 200 non-cancerous with same features as that of 120 samples.

3.1.2 Pre-Processing of Image

Initially, all images are grayscale images and are represented from 0 to 255. While doing scaling, images are represented in 0 and 1 using threshold point. RGB images are difficult to process than grayscale images so these images are used to process for further processing. Filters are used to remove this noise, which is may contain in collected images which may be of any type. So filters are used to remove. In this research, median filter is used to remove noise. Under certain conditions, the median filter preserved the edges without any interference while removing noise from the image, and it is beneficial for further processing.

3.1.3 Segmentation

For identification and to disperse the cancerous object preferred from the context, the segmentation phase is used. The segmentation phase uses a method to find contours, where the outcomes of the thresholding process that was performed are looked at the widest area of cancer by these pixels. The results of pre-processing is then used for removal of the cancer-detected area from the original image in the segmentation process, where the system has the largest value of 1 pixel from the region of the pre-processed lung CT scan. An important role is played by segmentation in disease classification. Images are divided into several segments. Together, these segments overall covers an entire area of the image. Marker-driven watershed segmentation is used for the purpose of segmentation. Fast segmentation is the main objective of this technique. It provides faster processing of the results when compared with other similar techniques. Apart it helps to eliminate the over-segmentation created in the watershed transform technique [8].

3.1.4 Feature Extraction

Several techniques are available for the purpose of extraction of the features from the images. GLCM is one of the techniques used to extract texture features from the targeted image. This is most preferred technique because this technique accord with pixel related information based on the position of similar gray-level values. All different features can be extracted area, perimeter, eccentricity, contrast, energy, entropy and homogeneity. In our proposed method, three features, namely area, perimeter and eccentricity are used with fivefold and sevenfold cross validation. The texture-based feature extraction is carried out using the gray-level co-occurrence matrix (GLCM) method. All pairs of gray layers on the joint frequency separated by distance and direction of u and v nodules in the GLCM nodules. This method is used to scale and distinguish The GLCM method calculates the contrast, energy, entropy and homogeneity of the cancer object using following formulas. Here, pixel size in the given shape u and v is used and region of interest as ROI. Two vectors in X and Y directions for position of interest of area. These values are calculated by using below formulas:

Area

$$Ar = Ar(u, v)XROI[Ar] = u, YROI[Ar] = v \tag{1}$$

Perimeter

$$Pr = \Pr(u, v)Xedge[Pr] = u, Yedge[Pr] = v \tag{2}$$

Eccentricity

$$Er = \frac{\text{Major axis length}}{\text{Minor axis length}} \tag{3}$$

Contrast

$$Cr = \sum_{uv} |\boldsymbol{u} - \boldsymbol{v}| 2\boldsymbol{p}(\boldsymbol{u}.\boldsymbol{v}) \tag{4}$$

Energy

$$E = \sum_{uv} p(u.v)^2) \tag{5}$$

Homogeneity

$$Hr = \sum_{uv} \frac{p(u.v)}{1 + |u - v|} \tag{6}$$

Entropy

$$En = \sum_{uv} p(u, v) logp(u, v) \tag{7}$$

3.1.5 Classification

Five different machine learning method are applied for classification phase. Values in the form of cancer area are contrast, energy, entropy and homogeneity as parameter and these are the inputs used in this phase. While the output that is produced in the classification phase gives the result in the form of cancerous and non-cancerous. There are normally two states in the process of classification for machine learning approach, namely training and testing.

The proposed algorithm is fragment of a family of supervised learning algorithms and its goal is of building a training model which can be used for prediction of target class or values by applying learning rules for decision derived from the data which is for training. The proposed algorithm is useful for solution of regression and classification problems. The SVM (Support Vector Machine) identifies different patterns from the given data set and classifies the data for analysis [10]. First, a set of input data from which these two classes are classified is considered. These classes are assigned to the classifier for detection purposes [11]. In this work, DT and SVM classifies cancerous images and normal image and the result of 120 and 700 samples. The accuracy for the adopted classification as the provided result.

4 Simulation Results and Discussion

4.1 Image Dataset: CT Images

The image dataset for CT images of 120 and 700 samples has been downloaded from https://www.kaggle.com/datasets. Figures 2 and 3 represent common and unusual images and Fig. 4 shows the result of segmentation taken from the database for the purpose of classification. The format for all the cancerous and non-cancerous images is Dicom format.

The curve receiver operating characteristics (ROC) is depicted for the 120 samples and 700 samples as shown in Figs. 5 and 6, respectively. It takes the input as an image which is shown in fig above.

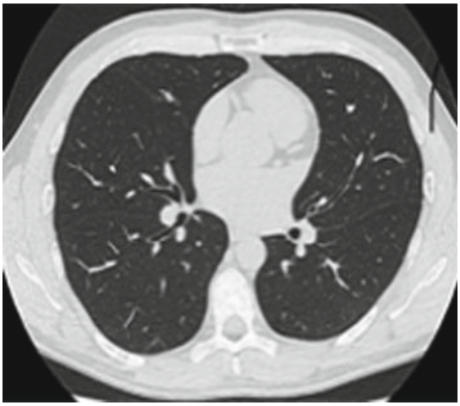


Fig. 2 Normal lung image

Fig. 3 Abnormal lung image

Fig. 4 Segmented lung image

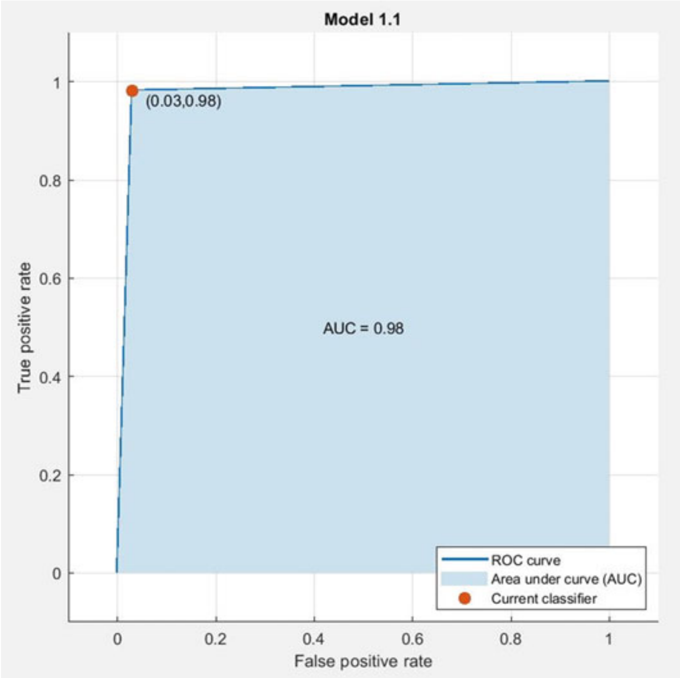


Fig. 5 ROC curve-CT image-120

Figures 7 and 8 show the scatter plot taken randomly with two columns for the 120 and 700 samples, respectively, of the CT Images. In the plot, two different colors are used for the cancerous and non-cancerous images, red color shows the non-cancerous, while blue color for cancerous images.

The accuracy of the five types of classifiers when used with the CT images for 120 and 700 samples is as given in Table 1. The same is also presented in graphical format in Fig. 9. As can be seen, for 120 sample, the accuracy is highest with TREE at 98.4. The precision and F1 score is highest with NB and LR. The recall rate is comparable among all the classifiers.

For 700 samples, the accuracy of KNN and TREE is highest. The precision is highest at 96.31 with TREE. The recall rate is highest with KNN and F1 score is comparable among KNN and TREE.

Different performance levels are included in Table 1 for image data. Of the 700 samples, it was found that the highest accuracy was achieved using a dividing tree (98.4%) with 5 features and 7 cross-sectional confirmations. The same is represented in a graphical form in Fig. 10.

A straight comparison with the existing methods, multiple task work executes to get maximum accuracy that shows good performance of the proposed method. In every step that highlights the better performance in cancer detection. However, we

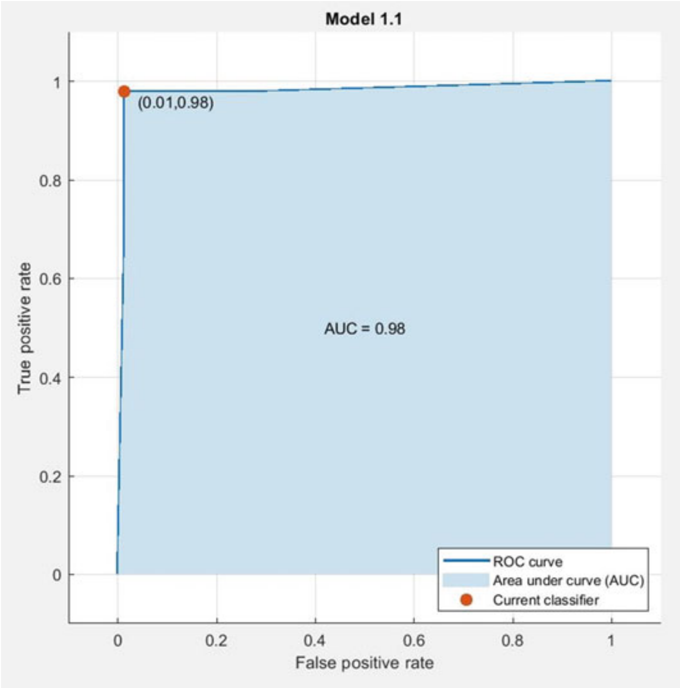


Fig. 6 ROC curve-CT image-700

have provided few of existing ones of the same kind by graphically. It shows that the better results in detection.

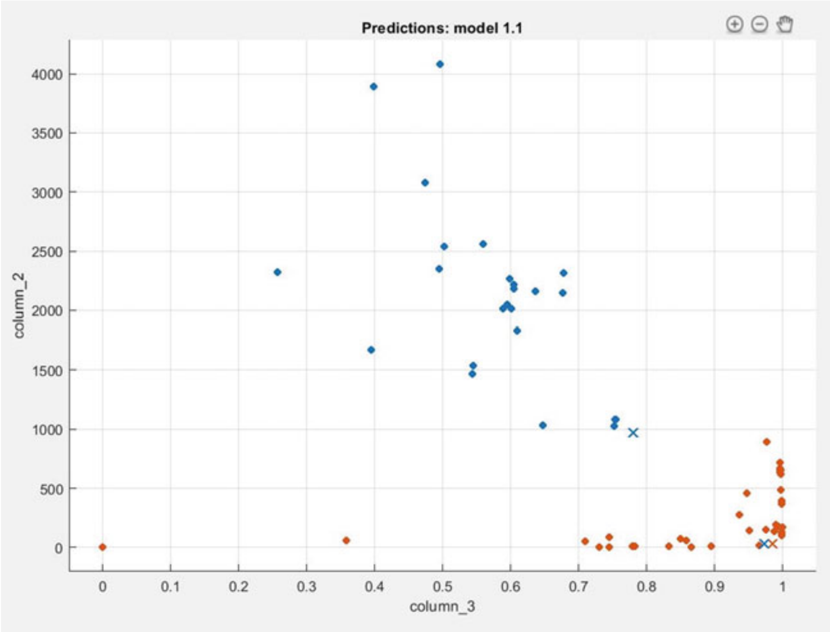


Fig. 7 Scatter plot for columns-120

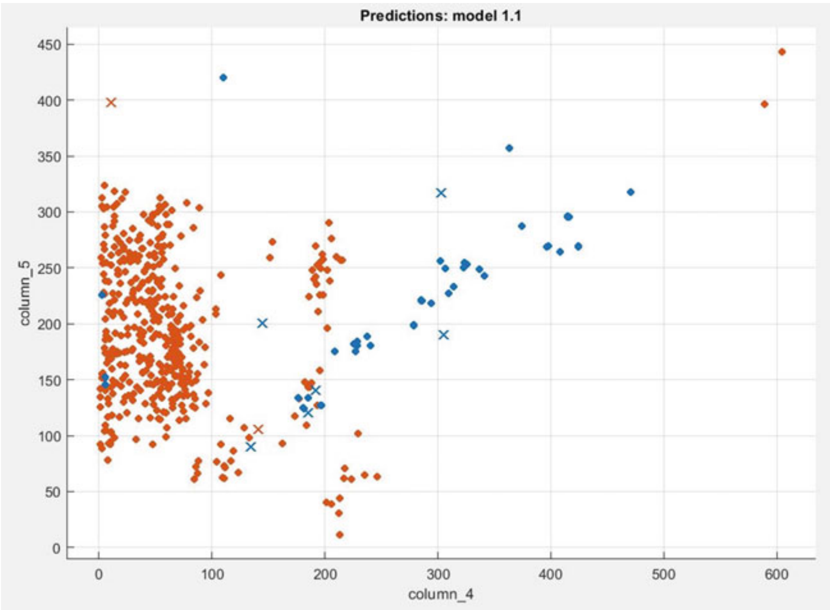


Fig. 8 Scatter plot for columns-700

Table 1 Performance analysis CT image

No. of samples	Methods/ parameter	Logistic regression	NB	KNN	SVM	Proposed method
120	Accuracy	92.3	92.8	96.9	97	98.4
	Precision	98.14	98.14	89.83	96.85	96.36
	Recall	98.14	98.14	98.14	97.88	98.14
	F1-Measure	98.14	98.14	93.80	97.36	97.24
700	Accuracy	93.26	92.31	97.96	96.08	97.96
	Precision	91.01	92.16	94.44	93.61	96.31
	Recall	85.71	80.95	98.94	93.12	96.82
	F1-Measure	88.28	86.19	96.63	93.36	96.96

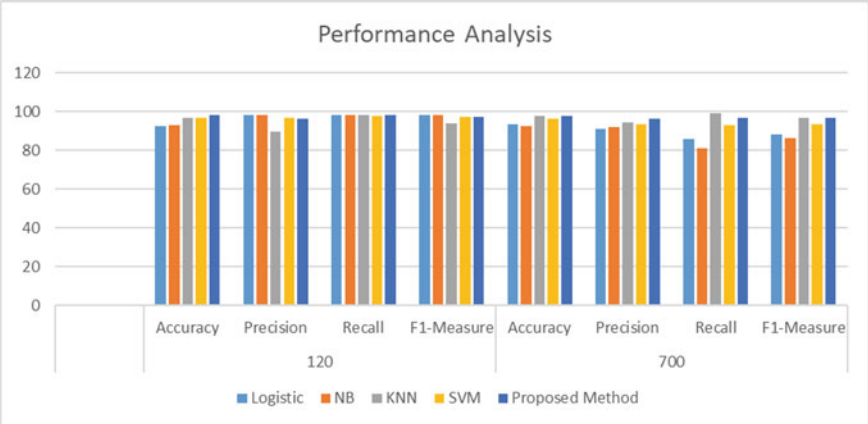


Fig. 9 Performance analysis CT image

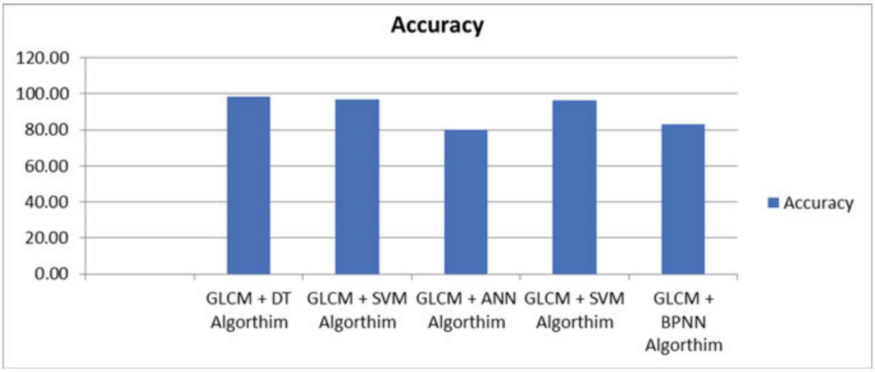


Fig. 10 Comparative study of lung cancer detection

Table 2 Comparative analysis of accuracy lung cancer detection

S. No.	Detection method	Author	Dataset	Accuracy
1	GLCM + DT	Proposed method	Lung data	98.40
2	GLCM + SVM	Usha Kumari et al. [2]	Lung data	96.70
3	GLCM + SVM	Li et al. [4]	Lung data	96.32
3	GLCM + BPNN	Anifah et al. 1	Lung data	83.07
4	GLCM + ANN			80.00

5 Conclusion

The study shows successfully classified status of lung cancer patients on the basis of CT-scan images by applying deep learning techniques. CT-scan image-based lung cancer detection system is used with 120 and 700 samples with selected features and fivefold cross validation and its one versus one approach under the proposed method. In order to enhance the accuracy of the cancer detection, the proposed method uses GLCM and decision tree algorithm. Models were validated on independent data set. The test results based on this has demonstrated that the proposed algorithm has a maximum accuracy of 98.4% with large unbalanced data and cross validation when compared with other existing methods of small balanced data, as mentioned in Table 2.

References

1. Anifah L, Haryanto, Harimurti R (2017) Cancer lungs detection on CT scan image using artificial neural network backpropagation based gray level coocurrence matrices feature. In: International conference, 2017
2. Mohamed ASED, Salem MA (2015) Probablistic-based framework for medical CT images segmentation. IEEE
3. Patil BG, Jain SN (2014) Cancer cells detection using digital image processing methods. Int J
4. Li L, Wu Y, Yang Y, Li L (2018) A new strategy to detect lung cancer on CT images. IEEE
5. Divyashree BV, Kumar GH (2021) Breast cancer mass detection in mammograms using gray difference weight and MSER detector. J Springer Nature
6. Bharathy S, Pavithra R, Akshaya B (2022) Lung cancer detection using machine learning. ICAAIC, pp 539–543
7. Ashwini SS, Kurain MZ, Nagaraja M (2021) Performance analysis of lung cancer classification using multiple feature extraction with SVM and KNN classifiers. IEEE
8. Al-Thrawneh MS (2012) Lung cancer detection using image processing techniques. Electron J Pract Technol
9. Adi K (2018) Detection lung cancer using gray level co-occurrence matrix (GLCM) and back propagation neural network classification. J Eng Sci Technol Rev
10. Jony MH, Fathema (2019) Detection of lung cancer from CT scan images using GLCM and SVM. ICASERT
11. Tuncal K, Sckeroglu B, Ozkan C (2020) Lung cancer prediction using machine learning algorithms. Turkey

12. Karthikeyan R, Kalaiyarasi M (2021) A review of lung cancer detection using image processing. Smart technologies. Communication and robotics
13. Rehman A, Kashif M (2021) Lung cancer detection and classification from chest CT scans using machine learning techniques. CAIDA
14. Ankita R, Kumari U (2019) Lung cancer image-feature extraction and classification using gLCM and SVM classifier. IJITEE
15. Dandil E, Cakiroglu M, Ekşi Z (2014) Artificial neural network-based classification system for lung nodules on computed tomography scans. In: International conference of soft computing and pattern recognition Tunis
16. Nageswaran S, Arunkumar G, Bisht AK (2022) Lung cancer classification and prediction using machine learning and image processing. Hindawi BioMed Res Int
17. Chaunzwa TL, Hosny A (2021) Deep learning classification of lung cancer histology using CT images. Scientific Reports
18. Nadkarni NS (2019) Detection of lung cancer in CT images using image processing. In: Trends in electronics and informatics
19. Venkatesh C (2022) A neural network and optimization based lung cancer detection system in CT images. Frontiers in Public Health
20. Feng J (2022) Deep learning-based chest CT image features in diagnosis of lung cancer. Hindawi Comput Mathem Methods in Med

Digital Twin in the Aviation Industry and the Role of Machine Learning: A Review

Priya Aggarwal, Bhawna Narwal, Amar Kumar Mohapatra, and Pooja Bhati

Abstract Utilising the power of machine learning and artificial intelligence in essentially every domain, the digital twin has taken a rapid leap in the realm of manufacturing, supply chain, aviation and more. Digital twin usage has rapidly increased in past years because of its integrity, dynamics, interactive nature and central behaviour. The digital twin has integrated itself into many fields such as smart factories, smart cities and smart machines. It guarantees security in the aviation industry which is the major concern of this paper. The objective of this paper is to familiarise readers with the notion of digital twin technology and at the same time incorporate machine learning with it. Finally, the future predictive trends of the digital twin have been discussed.

Keywords Artificial intelligence · Aviation industry · Digital twin · Integrity · Machine learning

1 Introduction

Today, the aviation industry is running faster than ever as a bullet train in moving people from one destination to another. Cyber-technology and machine learning have significantly promoted the growth of the aviation industry as it caters for the needs of passengers, viz. flight bookings, e-check-in and more. These technologies have

P. Aggarwal (✉) · B. Narwal · A. K. Mohapatra
IT Department, Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi, India
e-mail: priya013mtit21@igdtuw.ac.in

B. Narwal
e-mail: bhawnanarwal@igdtuw.ac.in

A. K. Mohapatra
e-mail: akmohapatra@igdtuw.ac.in

P. Bhati
MAE Department, Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi, India
e-mail: poojabhati@igdtuw.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_41

uncovered the aviation industry to vicious and highly priced threats [1] withal the risk of handling and sharing confidential data prevails. Thereupon, to address cyber-security issues International Civil Aviation Organisation (ICAO) [2], advised each contracting state, employing cyber technology, to develop measures to protect data processing and communication systems used for aviation purposes. We have entered a new era of combining digital twins with the aviation industry [3]. In digital twin, we use information mirroring model. This means we can collect more information at the physical end and make a mirror of the same at the virtual end for simulation and processing of the data. Figure 1 is an example of an information mirroring model.

While research on the aviation industry from the perspective of cybersecurity using digital twins is still underway, the study using machine learning algorithms seems to be the next in the cybersecurity domain and engineering. After introducing digital twin technologies: How the aviation industry can be benefited from ML-based cybersecurity? In ML, one can explicitly design an algorithm using a pre-decided set of features from the data set collected physically, learning from data and predicting both the trained and untrained data.

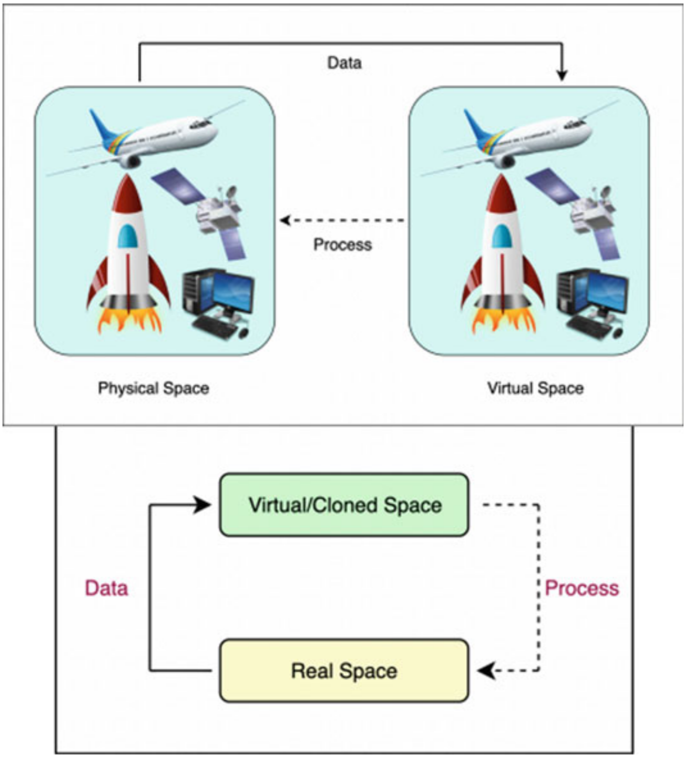


Fig. 1 Digital twin representation

In 2018, Bécue et al. [4] designed an aerospace electronic card for avionics production data lake and data lake techniques in their study for digital twin by analysing the impact of attacks in manufacturing and monitoring aviation electronics in digital twin/virtual environment to save time, and cost and withstand the ability to determine the robustness towards cyber threats. In 2021, Kuehner et al. [5] performed a meta-review on how digital twin was founded on common grounds. They compared several concepts of digital twin models, and shadows to their level of integration. They researched on many various concerns and categorised them to the number of occurrences across reviews. In 2019, Mukherjee and DebRoy [6] conceptualise the idea of building a bridge between the real world and the virtual world of printing. They created a DT which reduces the number of analyses of trial and error, defect mitigation, and reduces the time between design and production. They used the technology of the mechanist model of 3D printing. In 2022, the authors of [7] discussed the use of digital twins in the field of military targeting the aviation industry of the military sector. They stated that digital twin stands out in the designation of effectual cost solutions and increasing efficiency and the quality of the product. According to their conceptualised study, the present level of DT is anticipated to rise with the extensive usage of DT technology in the future. They also concluded that more hybrid studies can lead to being conducted on how DT technology will be taking over with other coming-up technologies like AI, ML, Data Science, IoT, blockchain and 5G, and how these technologies will have an impact in future.

2 Digital Twin in Different Domains

2.1 Artificial Intelligence and Machine Learning for Cybersecurity

AI is a cutting-edge technology widely applied for improving the strength of building aviation industry systems stronger at each stage of their lifecycle and enhancing the safety of aviation systems and their unique ability to modify, learn, enhance and anticipate difficult situations. In contradiction, ML is an application of AI which can explicitly design an algorithm using a pre-decided set of features from the data set collected physically and learning from data and predicting both the trained and untrained data.

2.2 Aviation Cyber Security

Aviation is considered to be the safest, fastest and most well-interconnected transportation around the world. Aviation cybersecurity protects the crucial systems of airlines from dangerous cyber-attacks and threats which could apparently cause damage to both the airline and people. Impact of IT in aviation—Today Wi-Fi is easily available on every aircraft you travel and electronic flight bags and entertainment systems are provided to users for news and sports. With technologies like ML, IoT and cloud computing, airlines can robotise the manual process [8]. IoT has benefited airlines with advantages, namely high flight air time with less time on the ground for maintenance which does cost-saving for airlines. IoT and digital twin lets engineers examine the mechanical parts for performance without disassembling them. IoT has improved luggage service tracking lost baggage even if it's on the aircraft. In 2019, Navid and Prathamesh [9] proposed improved solutions in aviation cybersecurity. They stated that every airline should conduct formal risk assessments on regular basis. They focused on understanding and assessing the potential attack and danger, providing the incident response report, taking part in research and development, coming up with cyber standards and defining operational designs and principles. The vulnerability and threats face insider threats, network access control, DDoS attack and many more. They focused on five prevention techniques for the above attacks named as (a) building a culture of security (b) a proactive approach (c) supply chain risks (d) establishing international standards (e) a critical role of the board of directors (Fig. 2).

Fig. 2 Aviation cybersecurity components with connectivity [15]

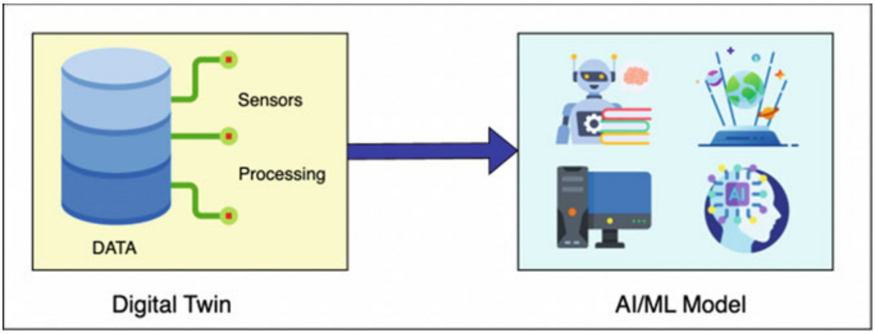


Fig. 3 Digital twin and AI/ML model representation

2.3 Machine Learning and Digital Twin

Digital halves are virtual clones of effects within the real world. They are realised by numerical modelling of plots, countries and attributes of the real world, which also makes it a dynamic virtual dupe. When a digital binary receives information from numerous places, it is constantly learning further about the thing or exertion its copying. Digital twin takes new information and combines it with real-time IoT data, also using advanced modelling and machine literacy ways, and the whole ecosystem is represented in a time-varying manner. Ever since the word "learning" was coined, machine learning has been associated with it. Using machine learning is the key to accelerating automation and making machines smart enough to learn on their own without human help. Before using a machine in a manufacturing process, it must be trained. In our case, old models are useful for training. Using supervised learning techniques, you can predict how people will act in the future based on what they are doing now. It is necessary to better understand the names of the parts of the proposed model. By changing a property when a developer changes part of a model, they can evaluate how well the model is performing. Research shows that digital twins can be used to make adjustments, and multivariate linear regression can be used to find the most successful and appropriate architecture [10, 11] (Fig. 3).

2.4 Simulations and Digital Twin

The digital double concept is often compared to simulation. To the untrained eye, they may look the same, but to fully understand the potential of digital twins, we need to examine their actual behaviour, structure and function. The simulation is a subset of the digital twin in the real world. In other words, a simulation typically analyses only one process, but a digital twin can run as many linked simulations as needed to study different processes. Moreover, simulation is not well suited for real-time computing. If your system contains a lot of data, performance will not be optimal. This creates a demand for digital twins. When the simulation cannot handle large amounts of data or data that updates in real time, it cannot take advantage of this type of data and miss an important goal [12].

2.5 Digital Twin and Data Science

We are surrounded by the aid of using facts each day. Our strategy uses real-time data, which is ultimately considered big data and used for digital doubles. When there is a lot of constantly changing data, it is difficult to understand. However, with data science, you can gain knowledge from data and use it to your advantage. The field of data science is increasingly being used for a variety of applications, including analysing corporate data and predicting future trends. In addition, it can be used to demonstrate the benefits of real-time data analysis using digital copies. Although digital twins can use data in real time, they cannot derive knowledge from the data or use the data to perform statistical analysis, which can be an advantage. Data science techniques such as machine learning, deep learning, statistical and probabilistic analysis and big data processors such as SPARK and Hadoop are typically combined to learn from the data and predict what is useful for the model at the right time. Data science is behind everything they do here [13].

2.6 Digital Twin and Cyber Security

Cybersecurity is an essential part of the aviation industry and digital twins of the aviation industry. Even if the test passes, serious defects may occur after the model is released or used in production. Even if tests are done correctly during development, it still happens. When this happens, setting up a digital twin can save your day. In addition, the digital twin architecture is available for all models and is not limited to aerospace products. This feature allows you to perform a security check. When enough tests are performed, it becomes a real example of cyber defence. We recommend that you use the digital version and the actual model after testing is complete. Implementing security protocols directly in the model makes it difficult to detect errors that cause unexpected problems. Digital double allows developers to test the security of their programmes in a more realistic environment.

3 Applications of Digital Twin

Today DT is being used everywhere in our real life. A DT is used in many different ways in real life and displays the most salient and basic applications are shown. Figure 4 depicts the graph of how digital twin has been used explicitly in the

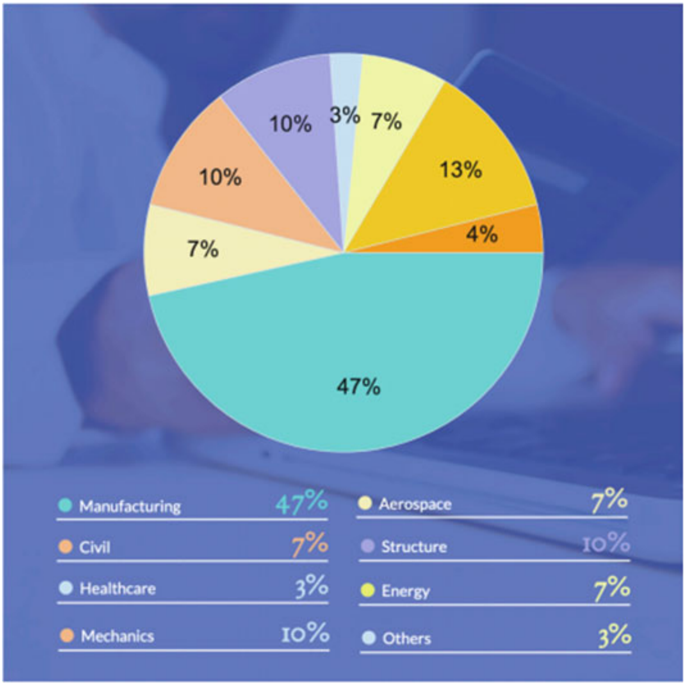


Fig. 4 Applications of DT in different sectors

different industries with the manufacturing industry at a percentage of 47%, the aviation industry using 7%, civil at 7%, structure at 10%, health care at 3%, energy at 7%, mechanics at 10% and other using DT technology 3%. The application of DT aerospace has advanced with time. In 2010, NASA released a roadmap for space technology which set a goal for achieving DT in the aerospace industry especially NASA by 2027 [11]. The report stated four scenarios of digital twins.

- Used for "test flights" of the aircraft before take-off. Analysis of the influence of various functional parameters, study and testing of appropriate strategies for the treatment of various anomalies.
- When the sensor structure shows a decrease in performance, the cause of the failure is determined and post-failure response actions are analysed.
- It shows the exact flight of the plane. Real-time monitoring of loads, temperatures and structural damage that reflects real-world flight conditions.

4 Prospects

Digital twin is still developing and taking its place with pre-existing technology, it can be used to achieve various security features. Aviation is one of the fastest-growing industries that is exploiting the power of digital twin technology. Many industries implementing twin technology have gained significant results. As we have discussed, the DT is a burgeoning technology and has been predicted to be one of the most promising technology for the upcoming years [12]. Digital twin technology can take smart innovations to the maximum extent and make industrial production models more intelligent, efficient and sustainable. Digital twin is a key to the development of the industry of artificial intelligence. Fore-by, digital twin can extensively enhance medical infrastructure [13], artificial organs can be created using digital twins using which doctors can study and analyse the functioning of organs virtually while getting the same results. Drug research is also getting famous in virtual human bodies where the molecular level of drugs and clinical trials can be studied in real time. Moreover, digital twin provides an effective platform to deploy smart cities [14] and improve the quality, architecture, planning and construction of traditional city framework. Hence, digital twin is a robust, cost-effective, secure and adaptable alternative to conventional technology and physical assets [15] (Tables 1 and 2).

5 Conclusion

An empty field is the starting point for a double word and is the main field for running applications. Digital transformation is the only way for the future growth of the world economy. Digital twin is a promising and budding technology which finds its application in numerous fields including health sectors, education, aviation industries, etc. In the field of the aerospace sector, the digital twin has pre-existing and notable advantages, from equipment maintenance to real-time data processing. The digital twin can significantly reduce latency as well as time and space costs. A variety of such services are also used by relatively established companies and the private sector. This technology can be easily effectuated and scaled up using interactive twin interfaces and accessible applications.

Table 1 DT applications comparison

References	Domain	Objective	Physical twin	Data analysis
[16]	Smart city and urban waste	DT for distribution of water	Water distribution system (WDS)	Big data
[17]	Electrical energy	DT for network micro-grid resiliency against cyber security attacks	Micro-grid and PCC	IoT, cloud computing
[18]	Automotive industry	DT for privacy enhancement mechanism	Smart car ecosystem	ML
[19, 20]	Automotive industry	DT for autonomous vehicle and driving	Car	IoT
[21]	Manufacturing domain	DT for optimising energy consumption	SMT PCB assembly line machine	IoT
[22]	Medicine	DT framework for health in smart cities	Human	ML (CNN)
[23]	Aerospace industry	DT for fixating optimisation of aircraft skins	Aircraft	Big data, AI
[24]	VR/AR industry	DT for 5G IoT from VR/AR	Smart home, smart watches, smart gloves, smart clothes, smart home	AI
[25]	Electrical energy	DT framework for power grid online analysis	Online analysis architecture	ML, neural network
[26]	Manufacturing domain	DT for bibliometric analysis	Articles	Ml

Table 2 Characteristic study of manual data acquisition and DT data acquisition [27]

Characteristics	Manual data acquisition	Digital twin
Optimisation measures using automated derivation	Low	High
Efforts for installation	Low	Medium
Continuous acquisition of data	Low	High
Efforts of repetition	Medium	Low
Events observation in terms of high simultaneity factor	Low	High
Motion data capturing capacity	Low	High
Optimisation measures using repeated derivation	Low	High

References

1. Haass J, Sampigethaya R, Capezzuto V (2016) Aviation and Cybersecurity: Opportunities for Applied Research. TR News, (304). Retrieved from https://commons.erau.edu/publication/299
2. Abeyratne R (2011) Cyber terrorism and aviation—national and international responses. J Transp Secur 4(4):337–349. https://doi.org/10.1007/s12198-011-0074-3
3. Fuller A, Fan Z, Day C, Barlow C (2020) Digital twin: enabling technologies, challenges and open research. IEEE Access 8:108952–108971
4. Bécue A et al (2018) CyberFactory#1—Securing the industry 4.0 with cyber-ranges and digital twins. In: 2018 14th IEEE international workshop on factory communication systems (WFCS), pp 1–4. https://doi.org/10.1109/WFCS.2018.8402377
5. Kuehner K, Scheer R, Straßburger S (2021) Digital twin: finding common ground–a meta-review. Procedia CIRP. 104:1227–1232. https://doi.org/10.1016/j.procir.2021.11.206
6. Mukherjee T, DebRoy T (2019) A digital twin for rapid qualification of 3D printed metallic components. Appl Mater Today 14:59–65. https://doi.org/10.1016/j.apmt.2018.11.003
7. Nam KT, Samer A, Larissa M, Sarah W, Scott C, Scott B, Hooman R (2022) Evolving applications of artificial intelligence and machine learning in infectious diseases testing. Clinic Chemistr 68(1)
8. Dawson M, Mohamed E, Marwan O (2016) Security solutions for hyperconnectivity and the Internet of things. IGI Global
9. Kagalwalla N, Churi PP (2019) Cybersecurity in aviation: an intrinsic review. In: 2019 5th international conference on computing, communication, control and automation (ICCUBEA), pp. 1–6. https://doi.org/10.1109/ICCUBEA47591.2019.9128483
10. Potgantwar A, Aggarwal S, Pant P, Rajawat AA, Chauhan C, Waghmare VN (2022) Secure aspect of digital twin for industry 4.0 application improvement using machine learning (August 11, 2022). Retrieved from https://ssrn.com/abstract=4187977 or https://doi.org/10.2139/ssrn.4187977
11. Thelen A, Zhang X, Fink O, Lu Y, Ghosh S, Youn BD, Todd M, Mahadevan S, Hu C, Hu Z (2022) A Comprehensive review of digital twin—part 1: modeling and twinning enabling technologies.
12. Wu et al. (2021) Research and design of a digital twin-based enterprise architecture digital control platform for provincial electrical power company. In: 2021 6th international conference on control, robotics and cybernetics (CRC), pp 186–191. https://doi.org/10.1109/CRC52766.2021.9620120
13. James L (2021) Digital twins will revolutionise healthcare: Digital twin technology has the potential to transform healthcare in a variety of ways—improving the diagnosis and treatment of patients, streamlining preventative care and facilitating new approaches for hospital planning. Eng Technol 16(2):50–53. https://doi.org/10.1049/et.2021.0210
14. Mohammadi N, Taylor JE (2017) Smart city digital twins. IEEE Symposium Series on Computational Intelligence (SSCI) 2017:1–5. https://doi.org/10.1109/SSCI.2017.8285439
15. Gontar P, Hendrik H, Michelle R, Julia B, Dehais F, Klaus B (2018) Are pilots prepared for a cyber-attack? a human factors approach to the experimental evaluation of pilots' behavior. J Air Transp Manage
16. Conejos FP, Martínez AF, Hervás CM, Alonso Campos JC (2020) Building and exploiting a digital twin for the management of drinking water distribution networks. Urban Water J 1–10. https://doi.org/10.1080/1573062x.2020.1771
17. Saad A, Faddel S, Youssef T, Mohammed OA (2020) On the implementation of IoT-based digital twin for networked microgrids resiliency against cyber attacks. In: IEEE transactions on smart grid, vol 11, pp 5138–5150. https://doi.org/10.1109/TSG.2020.3000958
18. Damjanovic-Behrendt V (2018) A digital twin-based privacy enhancement mechanism for the automotive industry. Int Conferen Intell Syst (IS) 2018:272–279. https://doi.org/10.1109/IS.2018.8710526

19. Almeaibed S, Al-Rubaye A, Tsourdos A, Avdelidis NP (2021) Digital twin analysis to promote safety and security in autonomous vehicles. In: IEEE communications standards magazine, vol 5, pp 40–46. https://doi.org/10.1109/MCOMSTD.011.2100004
20. Winkler D, O'Connor RV, Messnarz R (2012) Systems, software and services process improvement. Commun Comput Infrom. https://doi.org/10.1007/978-3-642-31199-4
21. Karanjkar N, Joglekar A, Mohanty S, Prabhu V, Raghunath D, Sundaresan R (2018) Digital twin for energy optimization in an SMT-PCB assembly line. IEEE Int Conferen Int Things Intell Syst (IOTAIS) 2018:85–89. https://doi.org/10.1109/IOTAIS.2018.8600830
22. Laamarti F, Badawi HF, Ding Y, Arafsha F, Hafidh B, Saddik AE (2020) An ISO/IEEE 11073 standardized digital twin framework for health and well-being in smart cities. IEEE Access 8:105950–105961. https://doi.org/10.1109/ACCESS.2020.2999871
23. Hu F (2022) Digital twin-driven reconfigurable fixturing optimization for trimming operation of aircraft skins. Aerospace 9(3):154. https://doi.org/10.3390/aerospace9030154
24. Zhang Z, Feng W, Zhongda S, Xinge G, Tianyiyi H, Chengkuo L (2022) Artificial intelligence-enabled sensing technologies in the 5g/internet of things era: from virtual reality/augmented reality to the digital twin. Adv Intell Syst 22:2100228
25. Zhou M, Yan J, Feng D (2019) Digital twin framework and its application to power grid online analysis. CSEE J Power Energy Syst 5(3):391–398
26. Sheuly SS, Mobyen UA, Shahina B (2022) Machine-learning-based digital twin in manufacturing: a bibliometric analysis and evolutionary overview. Appl Sci 12(13):6512. https://doi.org/10.3390/app12136512
27. Uhlemann TH-J, Christoph S, Christian L, Stefan F, Rolf S (2017) The digital twin: demonstrating the potential of real time data acquisition in production systems. Procedia Manufact 9:113–120

Ship Detection in Remote Sensing Imagery for Arbitrarily Oriented Object Detection

Bibi Erum Ayesha, T. Satyanarayana Murthy, Palamakula Ramesh Babu, and Ramu Kuchipudi

Abstract It is quite difficult to locate ships in satellite images. Because ships are dispersed across the sea and in large ports, identifying them requires the processing of images from various remote sensing systems. The ability to recognize potential pirate threats correctly is vital to the mission's success. As a result, improving ship identification accuracy in marine surveillance systems makes a significant contribution. We present a novel deep learning strategy based on the You Only Look Once v7 (YOLOv7) deep learning architecture as part of this research. Furthermore, rotation can be detected using the YOLO detection technique without the need for angle regression. To summarize, we offer a method for simultaneously recognizing numerous remote sensing photographs exploiting the strengths of our fully convolutional lightweight network. It outperformed other algorithms by a substantial margin on both the publicly accessible HRSC2016 dataset and our own private large-area remote sensing (LARS) image collection, and it did so at an unparalleled rate.

Keywords DCNDarknet25 · Yolov7 · LARS

1 Introduction

Keeping a watch on traffic, approaches based on deep learning, like YOLO, and older methods. It is worth noting that ship identification is difficult because of the complicated environments, wide range of ship sizes, and different ship orientations. The data

B. E. Ayesha (✉) · T. Satyanarayana Murthy · P. R. Babu · R. Kuchipudi
Chaitanya Bharathi Institute of Technology(A), Hyderabad, India
e-mail: erumayesha11@gmail.com

T. Satyanarayana Murthy
e-mail: tsmurthy_it@cbit.ac.in

P. R. Babu
e-mail: prameshbabu_it@cbit.ac.in

R. Kuchipudi
e-mail: ramukuchipudi_it@cbit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_42

quality and availability may have an impact on how well ship detection algorithm function. As a result, ship identification techniques have a long way to go before they're perfect. It is worth noting that ship identification is difficult because of the complicated environments, wide range of ship sizes, and different ship orientations. The data quality and availability may have an impact on how well ship detection algorithm's function. As a result, ship identification techniques have a long way to go before they're perfect. When huge amounts of data are presented to a system with insufficient processing capability, a fast and efficient approach is first and foremost required. Second, it is hard to isolate ships against busy backdrops. Various ships at different times have incredible variations in shapes and looks, which makes ship detection even more challenging.

Convolutional neural networks, a relatively new and promising branch of the machine learning field. They have also been combined with a multilayered network architecture built with traditional neural network methods.

Several conventional single- and two-stage algorithms have obtained great results in horizontal object recognition because, when viewing real sceneries, objects are frequently perceived flat, and a horizontal bounding box may properly characterize the item position. When airplanes and satellites survey the landscape from above, horizontal bounding will include both the target ship and a portion of the nearby ship if the two vessels are too close together at an acute angle. The standard object detection technique seems unable to handle this real-world issue. The target ship and a portion of the nearby ship will be obscured by the box, which will delay processing. So far as we can tell, the standard object detection method is not up to the task of dealing with this real-world issue. The rotational object detection method has emerged as a solution to this issue.

We propose approaching the challenge of automated ship recognition from a different perspective. You Only Look Once (YOLO) is a popular architecture that underpins our method's capacity to identify the most distinctive traits.

2 Literature Survey

In order to successfully detect ships in both open ocean and port environments, Yang et al. [1] developed a set of rules for pyramid networks with rotational density features (R-DFPN). To address problems caused by the ship's restricted width, they suggested a network design called a Dense Feature Pyramid Network (DFPN).

Liu et al. [2] presented the HRSC2016 dataset, which contains optical remote sensing data with a high enough resolution to be used for vessel identification. They examined the ship hierarchy, the amount of bounding data, the unbiasedness, the extensive helping tools, and so on as part of the HRSC2016 study. They developed a new metric called SF to assess the effectiveness of sea-land separation.

Xu et al. [3] devised a precise but efficient method for detecting things with many orientations. To appropriately depict a multi-oriented item, instead of quickly relapsing the four vertices, they only skim the vertex of the even jumping box on each corresponding side.

Jiang et al. [4] proposed a Rotational Region CNN to solve the issue (R2CNN). To measure imprecise text districts and get the precise slanted box, R 2CNN recalls the Text Locale Proposition Organization (Text-RPN) and a do different jobs refinement network, respectively. Secondly, they addressed the issue of scene text identification by introducing numerous ROIPoolings. Lastly, they post-processed the detection candidates using an inclined Non-Maximum Suppression (NMS).

YOLO9000, an object detection system capable of recognizing over 9000 different object categories, was introduced by Redmon and Farhadi [5]. They suggested a method for training on item detection and classification simultaneously.

Redmon et al. [6] presented YOLO, an all-in-one approach for detecting objects. In addition to being straightforward to implement, their model also benefits greatly from being trained directly on whole images.

Typically, Rezatofighi et al. [7] Convergence over Association (IoU) is the most widely utilized product evaluation metric. They address one of IoU's shortcomings by providing a consolidated form, which can be viewed as both an additional misfortune and a sign of its effectiveness.

Ming et al. [8] provides a more successful label assignment strategy based on a dynamic anchor learning (DAL) mechanism that evaluates anchors' localization potential in depth using the newly provided matching degree. Using the DAL and a small number of horizontal preset anchors, they were able to improve detection performance for objects of any orientation.

Yang et al. [9] created a single-stage enhanced end-to-end finder for reorienting objects with enormous perspective proportions, thick dispersions, and unpredictable trajectories. They developed an element refinement module to address the issue of component misalignment in current refined single-stage finders.

Jianqi Ma a pooling layer for Turn Locale of Interest was developed for extending an uncertainly positioned proposal to an element map for a text neighborhood classifier.

Nan Xue proposed modeling the geometry change with the RoI Transformer module, which avoids the issue of misalignment between area features and objects. This strategy significantly improves oriented object detection on the difficult DOTA and HRSC with only a minor increase in computing cost.

Zikun Liu presented an essentially closed structure transport turned jumping box space for transport placement as well as a method for developing a few fundamentally appropriate opponents in this space. They started by determining if it was possible to precisely cover all boats with the name of turned leaping boxes.

By merging a graph neural network (GNN) with the You Only Look Once version 7 (YOLOv7) architecture, Krishna Patel presented a game-changing new method to deep learning. The main goal of this strategy was to clarify what the ship was doing in the harbor.

With a specific, high-priority target in mind, Zhongzhen Sun devised a scheme to ease the transfer of SAR photos. To begin, they used the suggested finder's located box jumping to re-calibrate the SAR Boat Location Dataset (SSDD).

Ji et al. [10] was proposed to compress models for learning in light detectors. Deep learning detection techniques may be difficult to execute on edge devices due to their high computational cost and memory usage.

Yiding Li suggested a Lite Faster-R-CNN, a lightweight network that recognizes ships in SAR pictures faster. Its core network for feature extraction, as well as task networks for locating and identifying, has been greatly improved.

Xingyu Chen presented a more powerful detector for ships that can be oriented in any direction. We need to fine-tune the finder's display from the standpoint of element isolation. They used a cutting-edge backbone to connect feature pyramids.

To improve detection accuracy, Xin Hou and Qizhi Xu devised a method that adapts detection based on ship size. The data revealed that the proposed strategy outperformed the conventionally used transport identification methods in terms of location accuracy [11–14].

Wang et al. [15] propose a self-designed deep convolutional neural network to reduce false warnings for all observed vessels.

Yunlong Gao proposed the improved YOLOv4 (ImYOLOv4), an attention-based ship detector capable of identifying vessels in SAR photographs rapidly and correctly. They employed the YOLOv4 algorithm first, and then CSPDarknet53 to remove multiscale inclusion maps to acquire such exact boat location.

Ch. Anusha developed a system for detecting ships in satellite data that uses deep learning with YOLOv3. This improves detection precision while also mitigating the effects of the environment.

A unique algorithm, SMMA-YOLOv5, has the potential to improve model placement precision without significantly increasing model size. To begin, we replace the original C3's convolutional block with one that employs window multi-head self-consideration (W-MSA) to analyze global input and zero in on the target area.

2.1 Problems and Solutions

A ship detection problems and solutions table would normally include a list of frequent problems encountered in ship detection, as well as potential solutions or ways to reduce these problems. Additional information such as the sources of the problem, the types of ships or surroundings where the problem is most likely to occur, and any relevant literature or study on the topic may be included in the table.

Table 1 presents an overview of some of the concerns that may emerge when spotting ships in satellite photography, as well as potential solutions to these issues. Researchers and practitioners can use the table as a reference resource.

Table 1 Problems and solution

Reference	Problem	Solution	Models	Metrics
Yang et al. [1]	Traditional ship identification systems usually suffer from the complexity of application settings, the difficulty of intensive object detection, and the detection region's redundancy	Rotation Dense Feature Pyramid Networks (R-DFPN), which can locate ships in a variety of situations such as the sea and a harbor. The Dense Feature Pyramid Network (DFPN) is presented as an attempt to address issues posed by the ship's restricted breadth	Dense Feature Pyramid Networks (DFPN) and Rotation Dense Feature Pyramid Networks (R-DFPN)	Recall–82% Precision–91.0% F-measure–89.6%
Liu et al. [2]	The difficulty in recognizing ships against complicated backdrops is the most significant hurdle to ship recognition; this issue needs to be addressed and researched further. There are also few open-source applications and no publicly accessible ship remote sensing dataset	In addition to the traditional bounding-box labeling approach, the HRSC2016 high-resolution ship dataset includes three-level classifications for ships, ship categories, and ship types. "The segmentation mask for each image in the "Test set" is also included, as is the ship head position for all ships with "V"-shaped heads	Fast R-CNN	Separation fitness (SF), avg time,
Liu et al. [16]	When dealing with enormous amounts of data on a platform with limited computational resources, a quick and effective solution is critical. Removing ships from complicated backdrops is challenging	A two-cascaded linear model was used to score latent candidates, followed by binary linear programming to select a limited set of highly prospective candidates	BING	Recall, accuracy

(continued)

Table 1 (continued)

Reference	Problem	Solution	Models	Metrics
Xu et al. [3]	Objects in arial shots, scene captions, and pedestrians in fisheye images were recognized using a multi-oriented object detector	In terms of missing data, the proposed technique outperformed all other strategies	Fast R-CNN, Multi-Oriented Object Detection	Precision, Recall, F-measure, FPS
Sun et al. [17]	Aim for the ship target in the high-resolution image with a wide aspect ratio, random direction, and dense arrangement	Recalibrate the SAR Ship Detection Dataset (SSDD) with the proposed detector's bounding box. The mosaic method is used to augment the data, and the loss function is changed to improve model performance	The YOLOv5 detection network is equipped with a Circular Smooth Label (CSL)	Precision, Recall, Average Precision (AP)
Hou et al. [18]	Ship detection from optical remote sensing pictures remains a difficult task due to the wide range of ship sizes	A size-adapted approach was developed to improve detection accuracy for ships of varied sizes. Ship candidates were immediately pulled from feature maps generated by three shallow networks of varied scales to adapt ship sizes	CNN	Precision, Recall,
Wang et al. [15]	Buildings and azimuth ambiguity both exhibit similar scattering mechanisms of ships, resulting in false alarms in ship detection	Deep convolutional neural networks that can learn discriminative features that are self-designed are used	Self-designed deep CNN	Classification accuracy

(continued)

Table 1 (continued)

Reference	Problem	Solution	Models	Metrics
Gao et al. [19]	Due of intricate backdrops and multiscale ships, deep networks fail to extract representative target properties, limiting ship detection capabilities to some extent	The attention mechanism has been used to improve YOLOv4. Because of its fast detection speed, we use the off-the-shelf YOLOv4 as our main framework to achieve the best balance of detection accuracy and detection speed	Improved YOLOv4, CSPDarknet53, Retina Net, Center Net	Precision, recall, F1
Patel et al. [20]	The precision of ship detection is an important issue as it is in charge of detecting any pirate threats	We present a YOLOv7 + GNN-based technique for automatic ship detection from the High-Resolution Satellite Image Dataset (HRSID),	GNN, YOLOv7	Accuracy

2.2 Pros and Cons

A ship detection pros and cons table would typically list the benefits and drawbacks of several methodologies or methods for spotting ships in satellite imagery. This table can assist researchers and practitioners in weighing the pros and drawbacks of various methodologies and selecting the best approach for their specific application. Table 2 summarizes the advantages and disadvantages of various strategies for detecting ships. Researchers and practitioners can make informed decisions about the strategy to adopt for their unique application by examining the costs and benefits of alternative strategies. Furthermore, this table can be used as a starting point for further ship identification research and development, suggesting places where improvements or novel approaches may be required.

3 Dataset and Modality

The HRSC2016 dataset is a high-resolution satellite imagery dataset that contains 30 scenes covering different land cover types such as urban areas, forests, agricultural land, and water bodies. In terms of modality, the HRSC2016 dataset is primarily visual, featuring RGB pictures with a spatial resolution of 0.5 m per pixel, among the highest available. The dataset includes over 13,000 annotated objects, each of which is labeled with a bounding box and information on the object's class, size, orientation,

Table 2 Pros and Cons

References	Pros	Cons
Yang et al. [1]	The R-DFPN boasts cutting-edge performance in ship detection in difficult settings, particularly when detecting densely organized ships	Despite the best efforts, there are still some issues. R-DFPN has a substantially lower precision than Faster-R-CNN and FPN due to more false alarms
Zikun Liu	HRSC2016's distinctive features, such as ship hierarchy, boundary information, tools, and so on, were examined. In addition, for sea-land segmentation, a new performance metric SF was added	Ship recognition, covering, among other things, the difficulty of recognizing ships in a cluster, the problem of ship recognition with a small number of samples, and the problem of rotational RoI pooling
Sun et al. [17]	It can achieve comprehensive ship detection performance in difficult port environments and broad sea areas, which has a high application value	It is a small library and may not be as robust as other libraries out there due to its limited functionality and lack of documentation
Hou et al. [18]	In terms of detecting accuracy, it demonstrates expressive performance	To train and run, a large amount of computational power and memory are required. The computing cost is high, and the interpretability is limited
Patel et al. [20]	It classifies ships more accurately and gives ideal outcomes	Require a lot of computational power and memory to train and run and limited generalization
Gao et al. [19]	ImYOLOv4 has a promising performance in recognizing ships in SAR photos while remaining quick	Computationally expensive, which may be a drawback if dealing with a huge dataset or needing to execute object detection in real-time

and aspect ratio. Many studies on satellite symbology and item placement have made use of the HRSC2016 dataset. And is freely available for research purposes. Its high spatial resolution and diverse set of land cover types make it a valuable resource for developing and testing algorithms for object detection in remote sensing imagery.

4 Conclusion

The practical usefulness of the proposed technique lies in its potential to improve the efficiency, accuracy, and reliability of ship detection in remote sensing imagery, which can have significant implications for various applications in the maritime and other domains. Although, the approach may not be suitable for all types of remote sensing systems, such as those with very high or very low resolutions, or those with different sensor types. Future directions for this research is, improving the performance of the approach on more diverse datasets and in real-world scenarios, such as in challenging weather conditions or with low light levels. Applying the approach to other types of objects in remote sensing imagery, such as vehicles, buildings, or natural features and investigating the use of different deep learning architectures or other machine learning techniques to improve the accuracy and efficiency of the approach.

References

1. Yang X, Sun H, Fu K, Yang J, Sun X, Yan M, Guo Z (2018) Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks. Rem Sens 10(1):132
2. Liu Z, Yuan L, Weng L, Yang Y (2017) A high-resolution optical satellite image dataset for ship recognition and some new baselines. In: International conference on pattern recognition applications and methods, vol 2. SciTePress, pp 324–331
3. Xu Y, Fu M, Wang Q, Wang Y, Chen K, Xia GS, Bai X (2020) Gliding vertex on the horizontal bounding box for multi-oriented object detection. IEEE Trans Pattern Anal Mach Intell 43(4):1452–1459
4. Jiang Y, Zhu X, Wang X, Yang S, Li W, Wang H, Luo Z (2018) R 2 cnn: Rotational region cnn for arbitrarily oriented scene text detection. In: 2018 24th international conference on pattern recognition (ICPR), IEEE, pp 3610–3615
5. Redmon J, Farhadi A (2017) YOLO9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7263–7271
6. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788
7. Rezatofighi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S (2019) Generalized intersection over union: a metric and a loss for bounding box regression. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 658–666
8. Ming Q, Zhou Z, Miao L, Zhang H, Li L (2021) Dynamic anchor learning for arbitrary-oriented object detection. In: Proceedings of the AAAI conference on artificial intelligence, vol 35, pp 2355–2363
9. Yang X, Yan J, Feng Z, He T (2021). R3det: Refined single-stage detector with feature refinement for rotating object. In: Proceedings of the AAAI conference on artificial intelligence, vol 35, pp 3163–3171
10. Ma X, Ji K, Xiong B, Zhang L, Feng S, Kuang G (2021) Light-YOLOv4: an edge-device oriented target detection method for remote sensing images. IEEE J Select Topics Appl Earth Observat Remote Sens 14:10808–10820
11. Ma J, Shao W, Ye H, Wang L, Wang H, Zheng Y, Xue X (2018) Arbitrary-oriented scene text detection via rotation proposals. IEEE Trans Multimedia 20(11):3111–3122
12. Ding J, Xue N, Long Y, Xia, GS, Lu Q (2019) Learning RoI transformer for oriented object detection in aerial images. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 2849–2858
13. Li Y, Zhang S, Wang WQ (2020) A lightweight faster R-CNN for ship detection in SAR images. IEEE Geosci Remote Sens Lett
14. Chen X, Tang C (2022) Arbitrary-oriented ship detection based on deep learning. In: 2022 international conference on frontiers of artificial intelligence and machine learning (FAIML), IEEE, pp. 200–203
15. Wang Y, Wang C, Zhang H (2018) Ship discrimination with deep convolutional neural networks in SAR Images. In: IGARSS 2018–2018 ieee international geoscience and remote sensing symposium, IEEE, pp 8444–8447

16. Liu Z, Wang H, Weng L, Yang Y (2016) Ship rotated bounding box space for ship extraction from high-resolution optical satellite images with complex backgrounds. IEEE Geosci Remote Sens Lett 13(8):1074–1078
17. Sun Z, Lei Y, Leng X, Xiong B, Ji K (2022) An improved oriented ship detection method in high-resolution SAR image based on YOLOv5. In: 2022 photonics & electromagnetics research symposium (PIERS), IEEE, pp 647–653
18. Hou X, Xu Q, Ji Y (2018) Ship detection from optical remote sensing image based on size-adapted CNN. In: 2018 fifth international workshop on earth observation and remote sensing applications (EORSA), IEEE, pp 1–5
19. Gao Y, Wu Z, Ren M, Wu C (2022) Improved YOLOv4 based on attention mechanism for ship detection in SAR images. IEEE Access 10:23785–23797
20. Patel K, Bhatt C, Mazzeo PL (2022) Improved ship detection algorithm from satellite images using YOLOv7 and graph neural network. Algorithms 15(12):473

Wireless Room Surveillance System

Pinna Murali Krishna, Achanta Tharun Kumar, Gandi Vinay Suhaas, Battini Pranay, and Shakti Raj Chopra

Abstract The wireless room surveillance system described in this project leverages various electronic components, such as NodeMCU, buzzer, LEDs, metal detector, MQ-2 gas sensor, MQ-9 gas sensor, and PIR sensor, to enhance security and monitor environmental conditions within a room. The system is designed to detect unauthorized access, presence of metal objects, and the presence of hazardous gases, providing real-time alerts to users for prompt action. The NodeMCU acts as the main controller unit, integrating the various sensors and actuators. The PIR sensor is employed to detect human presence and motion within the monitored area, enabling the system to trigger appropriate actions. When an intruder is detected, the buzzer is activated to generate an audible alarm, and the LEDs are illuminated to provide a visual indication of the security breach. To detect the presence of metal objects, a metal detector is utilized. It emits an electromagnetic field and measures changes in the field caused by nearby metallic objects, alerting users to potential security threats. Furthermore, the system incorporates gas sensors, specifically the MQ-2 and MQ-9 sensors, to detect the presence of harmful gases like methane and carbon monoxide. These sensors continuously monitor the air quality within the room, enabling the system to identify potentially dangerous gas levels and trigger appropriate safety measures. The collected data from the sensors is wirelessly transmitted to a remote monitoring station using the NodeMCU's Wi-Fi capabilities. This allows users to monitor the room's status in real-time via a web or mobile application, providing a convenient and accessible interface for surveillance and management.

Keywords NodeMCU · Buzzer · LED's · Metal detector · MQ-2 gas sensor · MQ-9 gas sensor · PIR sensor

P. M. Krishna (✉) · A. T. Kumar · G. V. Suhaas · B. Pranay · S. R. Chopra
Lovely Professional University, Phagwara, India
e-mail: pinnamuralikrishna@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_43

1 Introduction

Nowadays, we can see that there is lot of energy wastage and a lot of crime rate is increasing so that we need some things other than human interference so that we can improve the security system with the machinery other than human being which will be more cost efficient and the security will be more in case of machinery and we use here the microcontroller and the IOT devices we can use for the continuous surveillance of the room and the security of the persons entering the room can also can be checked using the microcontroller and we can raise alarms if we require, e.g., of fire in room we can raise the alarm for the help in room and we can turn off the power supply and control the things according to our purpose by the people who are in in-charge so due to this the things get more crisp and clear and all time things go right and we doesn't need continuous human surveillance and we can also store the data for some days in case of any bad things happened.

Going into the depths of digital communication using Wi-Fi, we use NodeMCU for many IOT-based projects because it has an inbuilt Wi-Fi module and is much faster in terms of usage [1]. The module is based on the ESP8266, a low-cost Wi-Fi microprocessor with a full TCP/IP stack and microcontroller capabilities. The maker, Espressif systems, first announced it. The ESP8266 NodeMCU is a sophisticated gadget that combines some capabilities of a standard Arduino board with the capability of connecting to the Internet. It has always been wise to incorporate automation into the pertinent project. The absence of built-in Wi-Fi functionality is a minor downside of these modules, though. Therefore, in order to make these gadgets work with the Internet channel, we need to incorporate an external Wi-Fi protocol into them. The ESP8266 Wi-Fi Soc serves as the foundation for the well-known NodeMCU. Based on the ESP-12E, this is version 3 of the device (an ESP8266 based Wi-Fi module). You may prototype your Internet of Things (IoT) device using the open-source NodeMCU firmware and development kit, which can be programmed using the Arduino IDE [1].

The extended machine-to-machine exchange at the centre of the Internet of Things is built on systems of information, social event sensors, and cloud computing [2]. They claim that it will make every aspect of our existence, from streetlights to seaports portable, virtual, and quick to link. However, this is what we mean when we say that people don't imagine it sufficiently. Machine-to-machine correspondence, in which devices or machines are capable of intelligent communication, has been the subject of a great deal of research. The job of IOT may be accomplished in two methods that can be stated as follows: it accumulates all the fundamental resources, such as names, addresses, and other related properties, utilizing automatic categorization, sensors, and satellite positions. IP addresses and domain names are used to identify and locate networking equipment on the Internet as its fundamental assets, while services like information retrieval and resolutions become essential components of the Internet's infrastructure in order to support all of its applications. The maintenance of all the resources of the devices that are connected through the Internet is a difficult undertaking since there is a high demand for the Internet of Things. One discipline used to maintain and retrieve information pertaining to the device is information retrieval [2].

The ESP8266 Wi-Fi module is a low-cost platform for controlling and communicating with Wi-Fi- or Internet-based devices [3]. It functions independently as well as in conjunction with other devices. It is connected to the microcontroller, as previously mentioned, and is utilized frequently in Internet of Things applications. There are a few readily available models, including the ESP-01, ESP-02, ESP-03, and ESP-14, among others, on the market, which can be found in genuine electronic stores (offline) or online easily. The module's functions are the same, but the general-purpose input–output mechanism (GPIO) is different. A module called ESP-01 makes it possible to connect the controller device to the Internet. This module, which is a system on chip (SoC), does not necessarily require a controller to drive its input and output (I/O). There are two distinct colours available for this kind of printed circuit board: black and blue. The setting is affected by variations in memory size. The other one is 1 Mb, while the blue one is only 512 Kb. Due to its limited GPIO, this module has the potential to function as a minicomputer. GPIOs, which are also referred to as GPIO-0, GPIO-1, GPIO-2, and GPIO-3, are also referred to as VCC, ground, reset, and chip enable. VCC, or direct current input voltage, is 3.3 V. It is frequently connected to the GPIO-1 and GPIO-3 for serial data transmission and reception. Because it has built-in AT firmware, the ESP8266 Wi-Fi module uses AT command operation for communication [3].

Wi-Fi technology was employed as the network backbone to link the hardware interface modules and server [4]. In addition to increasing system mobility and scalability, Wi-Fi is used to improve system security (by utilizing secure Wi-Fi connections). Repeaters or managed wireless LAN will totally eliminate that problem, even if a user wants to install new hardware interface modules outside of a central access point's coverage area. Managing, overseeing, and keeping an eye on distrusted system components are the server's primary duties. As a result, the server can receive triggered events from sensors and receive commands from actuators to carry out the duties assigned to the hardware interface modules. In set-up mode, the user can change the macros to carry out complex action sequences, add and delete hardware interface modules, and construct simple macros with simple triggers. Direct wire connections are used to link sensors and actuators to hardware interface modules. The capacity to regulate energy management systems, such as lighting, thermostats, heating, ventilation, and air conditioning (HVAC) systems, and security systems, is provided by hardware interface modules, such as door locks, cameras, motion detectors, and fire alarms [4].

1.1 System Model

Here in this project, we're going to use a NodeMCU, and the sensors that we're going to use are MQ-2 gas sensors for domestic gas and HFC-32 gas sensors for R32 refrigerant gas leakage, which we can detect with these sensors. We're also going

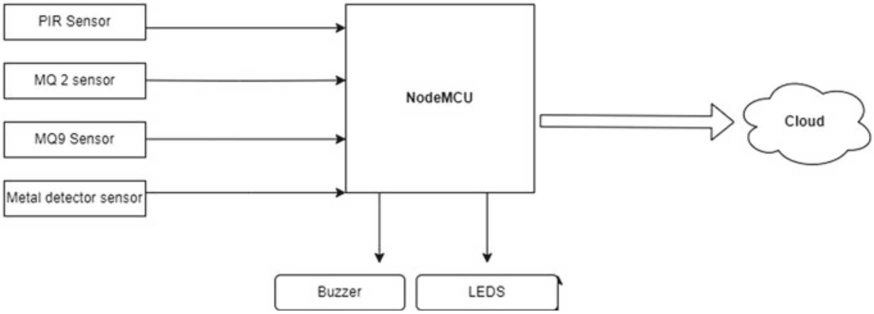


Fig. 1 Block diagram

to use metal detection sensors at the doors and PIR sensors to detect people in the room, so we've designed the project so that we can continuously monitor the room and receive security alerts. In this project, whenever a person enters a room, a metal detector detects them and checks them to see if they are carrying any potentially dangerous items. If they are, the detector raises an alarm, stops the person, and sends the report to security. When we move in, if no one is in the room, the detector notifies security to turn off the electricity to the room after the security check is finished from the security, and after the acknowledgement. In the event that a sensor or human security guard makes a mistake, then the person inside the room can press the button inside the room; once it is pressed, it raises the alarm to the security, and if the button is pressed twice, then the main security will be alarmed and shows the room number on the display from where the alarm is raised, and if there is no response from the security after three to five press, the microcontroller will automatically generated with and interrupt and opens the door for the human present in the room, and by this project, we can save electricity and better security for the users.

1.2 Block Diagram

See (Fig. 1).

1.3 Circuit Diagram

See (Fig. 2).

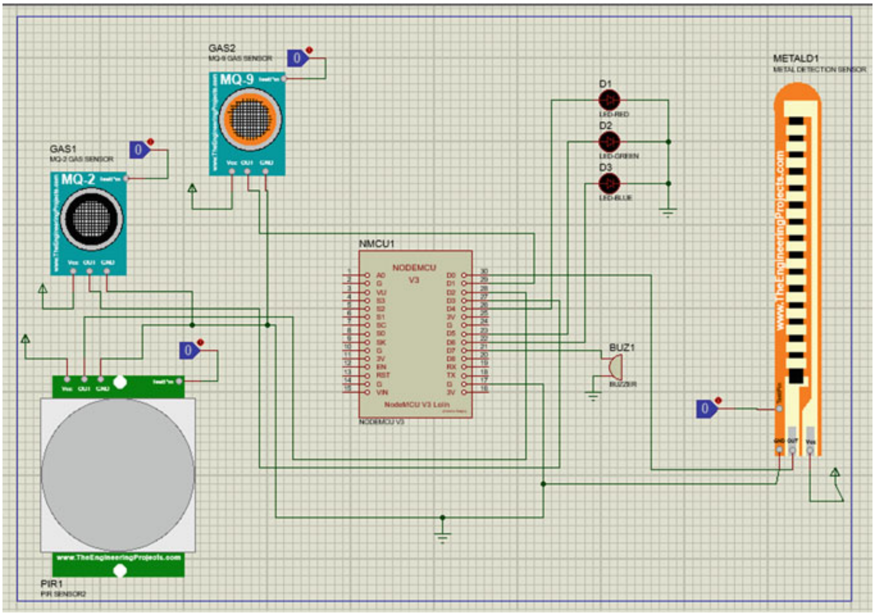


Fig. 2 Circuit diagram

1.4 Flow Chart

See (Fig. 3).

2 Ease of Use

We can reduce the electricity and provide the good security and safety for the people [5]. These days, it is more crucial than ever to safeguard buildings against fire. Monitoring commercial and residential buildings continuously is a smart technique to lower the risk of property and life loss due to fire. In the recent years, automatic fire alarm systems have been widely installed in certain locations. The information from a large number of tiny fire detectors should be sent to the control centre of a building or a block. Yet, the wiring cost for conventional wired fire alarm systems is relatively high. Wiring costs are significantly reduced by networking without existing infrastructure. Wireless sensor networks (WSNs) have been widely used in the past several years to monitor the environment, the health of structures, and industrial processes. For these applications, it offers affordable solutions. It is made up of compact, low-power, and inexpensive gadgets with a few built-in radio communication, sensing, and computation functions. Therefore, WSN is particularly suitable for detector-to-detector communication in fire alarm system [5, 6]. Since COVID-19 has caused the

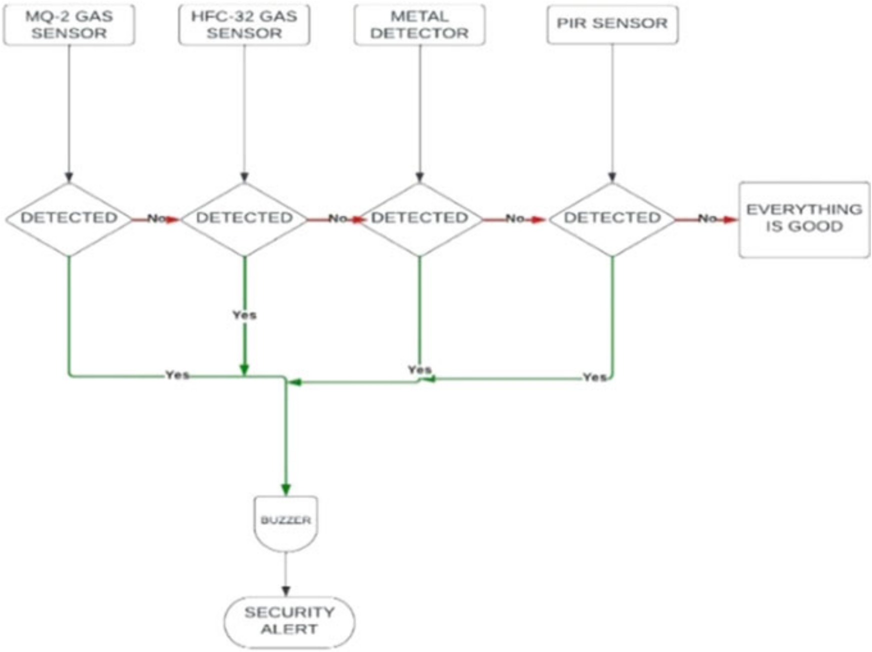


Fig. 3 Flow chart

present pandemic crisis, everyone must take precautions including wearing protective gear, keeping a safe distance, disinfecting their hands, and refraining from touching anything that is not absolutely necessary. However, there is a potential of disease transmission when we touch objects like tables, doors, cars, and other things. People frequently interact at offices and colleges, which increases the risk of disease transmission through touching contaminated surfaces like doors or other objects. The project is focused on automation, and we will build an automatic door mechanism (for a door to a house or college). In this scenario, the door opens when the door's sensor detects a cut, at which point a mechanism opens the door and closes automatically. As a result, there is no direct human contact with the door, which will assist stop or slow the spread of pandemic disease COVID-19 [6, 7]. Electric relays are widely used in electronics, which control our environment. They come in many shapes, perform many different roles, and each has their own peculiar behaviours. Applications and Principles of Electric Relays is now the only exhaustive reference that covers the entire spectrum of electric relays. This large-scale project is notable not only for its scope but also for its practical technique, which places more focus on the practical and operational aspects than the theoretical and mathematical ones. The first section of the book covers the fundamental components of relays and their essential functioning elements, such as contact and magnetic systems. Then, a distinct chapter is dedicated to each type of relay. The author describes the characteristics of several relays belonging to a type that operate on different principles as well as the operating and building principles for each type. The best introduction to the range of electric relays is provided in the quick-reference manual Electric Relays: Principles and Applications, which is written for experienced engineers. It is extraordinarily comprehensive and indubitably useful [7].

2.1 Buzzers

Our everyday activities depend heavily on time, particularly when it comes to sectional meetings or conference situations when careful time management is necessary [8]. In addition to discussing numerous electronics parts utilized as I/O devices, such as an LCD, A 7-segment display, LED and buzzer, the main topic of this article focuses on the creation and implementation of an advanced digital timer with an audio-visual device. In order to update the audience, inform them via audio-visual methods, and keep track of the times designated for events, this device is essential in our daily operations. Using the proposed methodology, you can effectively manage your time and avoid wasted time during seminars and similar events. Presenters benefit most from being time sensitive since it forces them to organically modify so that the allowed time is sufficient to cover their presentation. People with disabilities, such as the deaf and blind, can benefit from the digital timer and warning system presented here by taking control of their actions and being fully involved in the situation. To simplify diagnosis and troubleshooting of the mechanism, the system is designed in the form of various modules interconnected using integrated circuits [8].

2.2 Sensors

Wireless sensor networks (WSNs) have a lot of potential as a technology that will enable a variety of applications, from straightforward data collection to acquiring in-depth knowledge about a specific area of interest [9]. Because of the WSN technology's rapid progress, more complex WSN applications may now be made. The majority of WSN applications can be classified as belonging to one of the following groups: smart homes, industry and business, environment, or the military. Since the surveillance system is one of the typical WSN applications, numerous research have already been done and systems developed with this objective in mind.Two parts make up the overall system:

- Wireless sensor nodes
- A surveillance control system (Master node).

A network connection between the sensor and the master node is required. The sensor node can be readily deployed anywhere in the house because it is wireless. The nodes' modular design allows for the addition or removal of any number of sensor nodes [9].

2.3 Wi-Fi-Enabled Sensor

Both a sensor node and a NodeMCU are components of each module (ESP8266) [10]. The open-source, inexpensive NodeMCU Wi-Fi module is used in Internet of Things applications. The sensor is connected to an IoT module, which interprets sensor output and informs the master node via Wi-Fi on the sensor's status.

The sensor nodes in this system come in several varieties. Motion sensor nodes, metal sensor nodes, gas sensor nodes, and fire detection sensor nodes all employ passive infrared technology (PIR). PIR nodes can be installed in sensitive locations such as entry-restricted zones and they will be able to identify intruders when they enter. In the event of a fire, a fire node will detect it and turn on a fire ball, buzzer, and fire safety equipment.

2.4 Surveillance Controller System

The master node has a buzzer for alarms, a webcam for video surveillance, a GSM for remote notification. This node oversees a number of operations, including video streaming, GSM call and SMS alert management, email alert management, and the monitoring of sensor and camera feeds. A Wi-Fi server configuration has been made for the NodeMCU. In this case, the device connects to Wi-Fi via a USB Wi-Fi dongle.

Master node operations are driven by sensor updates from sensor nodes. As soon as the camera detects the status of the sensor, it powers on and begins recording the event. Viewers can view live video [10].

2.5 System Design and Working

The system is built as an embedded system with two components: Wi-Fi-enabled sensor nodes and a surveillance controller system [11].

The following describes each component's hardware architecture:

(1) Design of the PIR Sensor Node Hardware

For interacting with the sensor, it has a NodeMCU (ESP8226) module. A cheap, clever, open-source Wi-Fi module for Internet of Things uses is called Node MCU. It has a micro-USB connection for power and flash storage, ten GPIO pins, and a PCB antenna. It is very simple to operate and has a small size. A three-pin gadget that runs on 5 V, the PIR sensor. It tracks changes in the amount of infrared radiation that people emit to detect movements. An invader is present when the output pin has a logically meaningful value. When a sensor output is received, it updates a webpage with sensor status. The webpage may be seen in a browser using NodeMCU's IP address.

Due to the fact that the sensor and master node are both connected to the same Wi-Fi server, the master node may be able to get the sensor data from the website.

2.6 Fire Detection Sensor

It contains a fire detection temperature sensor (LM35). The LM35's analogue output is coupled with the NodeMCU's analogue pin. When a fire is found, a buzzer rings and fire protection equipment is activated. The function of fire safety devices is controlled by relays. You will need a 12 V power supply as the selected security device requires 12 V to operate. The NodeMCU only supports a maximum voltage of 5 V, so it receives a 12 V supply regulated to 5 V. A 12 V relay is used to control the operation. The fail-safe relay is controlled by a transistor connected to a digital output on the microcontroller. In the event of a fire, the transistor relay driver circuit and fuse will operate [11] (Figs. 4, 5, 6, 7, 8, and 9; Table 1).

Fig. 4 MQ-2 gas sensor

Fig. 5 MQ-9 gas sensor

Fig. 6 Metal detector

Fig. 7 PIR sensor

Fig. 8 Buzzer

Fig. 9 NodeMCU

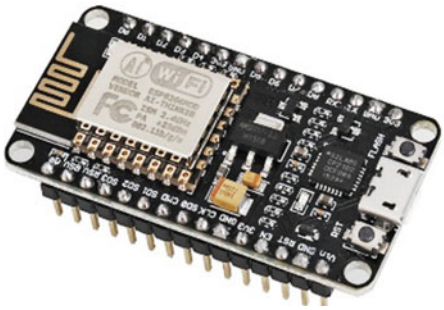


Table 1 V_{in} and V_{out} values of sensors

Sensor or actuator	V_{in}	V_{out}
MQ-2	5 v	0–5 v
MQ-9	5 v	0–5 v
PIR	+5–+12 v	0–3.3 v
Metal detector	+5 v	–
Buzzer	+5–+12 v	100–2000 Hz

2.7 MQ-2 Gas Sensor

Utilize the (MQ-2) gas sensor module to locate gas leaks (in houses and factories) [12]. It may be used to recognize substances like propane, alcohol, smoke, H_2, CO, or LPG. Due to the high sensitivity and short response time, measurements can be performed as quickly as possible. The sensitivity of the sensor can be adjusted with a potentiometer. Gas leak detection when using sensors is the first step in recognizing potentially hazardous gas leaks. These detectors usually use sound alarms to alert users when they discover potentially harmful compounds.

2.8 Metal Detector Sensors

An instrument called a metal detector scans the nearby surroundings for metal. Metal detectors can be used to locate metal objects buried underground or metal inclusions hidden inside other things. They generally consist of a mobile unit with a sensor probe that may be moved across a surface, such the ground, or an object. The proximity of the sensor to a piece of metal can be detected by a changing tone in headphones or a moving needle on an indicator. A distance indicator is frequently provided by the instrument; for example, a louder tone or higher needle movement indicate that the metal is closer. A different common kind is stationary "walk through" metal detectors, which are used at the entrances of prisons, courts, airports, and mental hospitals to look for hidden metal weapons. To detect metal, a metal detector's

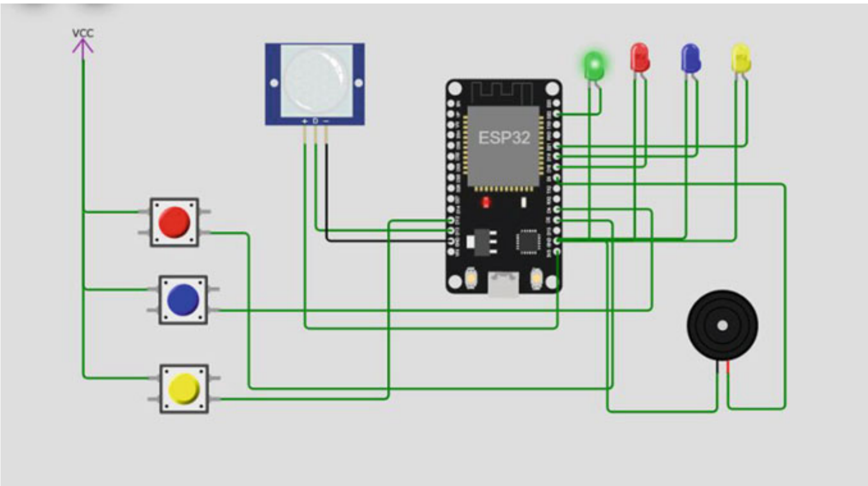


Fig. 10 Output on Wokwi.com

search coil creates an electromagnetic field in the ground. Any metal object (target) in the electromagnetic field acquires energy and creates its own electromagnetic field. The detector's search coil collects the resent fields and alerts the user with a target response. Metal detectors can distinguish between different types of targets, ignoring unwanted ones [12].

3 Result and Conclusion

In this project, we have used NodeMCU for the wireless connection and we have used multiple sensors for checking of the human presence and metal detection for security and we can send the data continuously to the user so that he can monitor the things from any location and he can also raise the alarms accordingly to the incident happened or else the security present near the incident can also be directly check the alarm on what basis the alarm has been raised and they can move for the further processing in this way we can provide the security for the home or any other workplace automatically as well as human controllable. So hereby, we can conclude that this project which we have done can provide the security for the home as well as any other places like malls and other places so we can we increase the no. of sensors and no. of actuators according to the requirement (Fig. 10).

In the above image, the button is mentioned in the place of MQ-2 gas sensor MQ-9 gas sensor and metal detection sensor.

References

1. Saputra L, Kuncoro P, Yuan L (2017) Implementation of air conditioning control system using REST protocol based on NodeMCU ESP8266. In: 2017 international conference on smart cities, automation & intelligent computing systems (ICON-SONICS), IEEE, pp 126–130
2. Walia NK, Parul K, Deepti M (2016) An IOT by information retrieval approach: smart lights controlled using WiFi. In: 2016 6th international conference-cloud system and big data engineering (confluence), IEEE, pp 708–712
3. Wibowo FW (2020) Wireless communication design of internet of things based on FPGA and WiFi module. In: Journal of physics: conference series, vol 1577, IOP Publishing, p 012035
4. ElShafee A, Karim AH (2012) Design and implementation of a WIFI based home automation system. Int J Comput Inform Eng 6(8):1074–1080
5. Zhang L, Gaofeng W (2009) Design and implementation of automatic fire alarm system based on wireless sensor networks. In: Proceedings. The 2009 international symposium on information processing (ISIP 2009), Academy Publisher
6. Rai B et al (2021) Automatic door opening and closing system using arduino. Int J Recent Adv Multidisciplin Topics 2.7:195–196
7. Gurevich V (2018) Electric relays: principles and applications. CRC Press
8. Ajao LA et al (2016) Development and implementation of microcontroller-based improved digital timer and alarm system
9. Sanjay SS, El-Ocla H (2022) Home security system using wireless sensors network. Wireless Personal Commun 1–17
10. Sruthy S, George SN (2017) WiFi enabled home security surveillance system using Raspberry Pi and IoT module. In: 2017 IEEE international conference on signal processing, informatics, communication and energy systems (SPICES), IEEE, pp 1–6
11. Abdulla AI, Abdulraheem AS, Salih AA, Sadeeq MA, Ahmed AJ, Ferzor BM, Sardar OS, Mohammed SI (2020) Internet of things and smart home security. Technol Rep Kansai Univ 62(5):2465–2476
12. Zhang T, Chowdhery A, Bahl P, Jamieson K, Banerjee S (2015) The design and implementation of a wireless video surveillance system. In: Proceedings of the 21st annual international conference on mobile computing and networking, pp 426–438
13. Yorozu Y, Hirano M, Oka K, Tagawa Y (1987) Electron spectroscopy studies on magneto-optical media and plastic substrate interface. IEEE Transl. J. Magn. Japan 2:740–741 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
14. Young M (1989) The technical writer's handbook. University Science, Mill Valley, CA

Skin Cancer Classification Using Machine Learning and Convolutional Neural Networks

Nidhi Patel, Aarnav Marathe, Vihaan Jasani, and Mansing Rathod

Abstract The timely and accurate diagnosis and treatment of skin cancer are crucial as it is a common and potentially deadly condition. One effective machine learning technique for accurately classifying skin lesions and improving skin cancer diagnosis is known as convolutional neural networks (CNNs). The article offers a detailed examination of the machine learning and CNN techniques employed for skin cancer classification. The discussion includes several studies that have utilized these methods to classify various types of skin lesions, such as basal cell carcinoma, melanoma, and squamous cell carcinoma. Nevertheless, the lack of high-quality datasets presents a significant challenge in skin cancer classification. To address this problem, one of the large-scale datasets developed is the ISIC dataset, which features over 20,000 annotated images of skin lesions. Skin cancer classification has been tackled using different machine learning algorithms, such as decision trees, support vector machines, random forests, as well as CNNs. CNNs are particularly useful for accurately classifying skin lesions because they can learn complex features from images. Several studies have utilized pre-trained CNNs, such as VGGNet and Inception, to classify skin lesions, achieving high accuracy rates of 85–95%. Ensemble methods, such as bagging and boosting, have also been used to improve skin cancer classification accuracy. These methods combine the predictions of multiple models to achieve better performance than a single model. Transfer learning has been effective in reducing overfitting and improving model generalization performance. This review also highlights studies that have utilized multi-modal data, such as dermoscopy and clinical images, to improve skin cancer classification accuracy by providing complementary information about skin lesions.

N. Patel (✉) · A. Marathe · V. Jasani · M. Rathod
Department of Information Technology, K.J. Somaiya Institute of Technology, Mumbai, India
e-mail: patel.nb@somaiya.edu

A. Marathe
e-mail: aarnav.m@somaiya.edu

V. Jasani
e-mail: vihaan.j@somaiya.edu

M. Rathod
e-mail: rathodm@somaiya.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_44

Keywords Skin cancer · Machine learning · Classification · Convolutional neural network

1 Introduction

Irregular growth of cells causes skin cancer. One of the primary reasons for the development of skin cancer is significant exposure to ultraviolet (UV) radiation. The UV rays destroy the cells in your skin and cause mutations in them. When these mutations lead to an increase in the rate of growth of the skin cells, it is known as skin cancer. It is fortunate that early classification of skin abnormalities into malignant or benign can help people address this issue in its early stages and this would increase the chances of being cured. Image preprocessing is a major part of skin cancer detection technologies when using machine learning and deep learning technologies. This process aims to remove noisy factors that invariably accompany an image. An almost necessary step done in this process is the removal of skin hair that tends to overlap with the image of the mark or skin abnormality. We also utilize post processing in order to handle the clicked image in such a way so as to give maximum visibility or the suspicious area. The picture is parted into different sections based on color, shade, texture, and size which help in eliminating the least significant portions of the image.

2 Literature Review

The skin is composed of three layers: the outer layer, known as the epidermis, the inner layer, called the dermis, and a deep layer of fat known as the hypodermis. When healthy cells grow uncontrollably and form a mass, it is called a tumor. Tumors can be benign, meaning they grow but do not spread to other parts of the body, or malignant, also called cancerous tumors, which can spread to other parts of the body. Skin cancer has four basic types. Basal cells are formed in the lower part of the epidermis, and about 80% of skin cancers arise in this area, referred to as basal cell carcinomas. The majority of the epidermis comprises flat cells known as squamous cells, and 20% of skin cancers start in this area, known as squamous cell carcinomas, often found near the mouth or old scars. Only 2–5% of these cancers spread to other parts of the body. Melanocytes are cells responsible for pigmentation, located in the area where the epidermis meets the dermis. Melanoma, the deadliest form of skin cancer, arises in these cells.

Unnecessary noises in the images are removed during the initial stages. Body hair is also considered troublesome to analysis and hence is factored out. Median filtering techniques are effective in filtering out hair and noise. RGB images increase the cost of computation; hence, these are reduced to grayscale images. SVMs and GLCM techniques were used in this, and SVMs evaluated better than GLCM techniques.

It makes use of the ubiquitous ABCD rule along with a total dermatoscopic value (TDV) for classification. Otsu's method along with gradient vector flow (GVF) and color-based image segmentation are analyzed for their ability to segment images. Median filtering technique and Dull Razor software is beneficial for image acquisition. Convolutional neural networks (CNNs) have been utilized in the experiments, where the training input size was 28 × 128. The SVM model is described in Eqs. 1 and 2.

$$\min_{o}\left(\sum_{i=1}^{m}[y^{i}\cos t_{1}\left(\theta^{T}x^{i}\right)+1-y^{i}\cos t_{o}\left(\theta^{T}x^{i}\right)\right)+\frac{1}{2}\sum_{i=1}^{n}\theta_{i}^{2} \tag{1}$$

$$S=y^{i}\cos t_{1}\left(\theta^{T}x^{i}\right)+1-y^{i}\cos t_{o}\left(\theta^{T}x^{i}\right) \tag{2}$$

In this study, the skin lesion dataset of 44,108 images with 33,126 and 10,982 was used for training and testing, respectively. The dataset comprises two primary features: whether the mole is cancerous, indicated as malignant, or non-cancerous, indicated as benign. The images were captured by advanced digital cameras, and in order to mitigate noise effects, CNN's grading and training were performed using a segmentation mask. In addition, the segmentation mask was utilized to even out the region surrounding the lesion, thereby minimizing the influence of normal skin texture on classification. The Canny edge detector was used to locate object boundaries in images, and non-maximum suppression was performed to narrow out the edges by identifying the pixel with the highest value on an edge. Additionally, a Gaussian filter was applied to the skin based on the segmentation mask's details to optimize the input image for CNN. The skin color image datasets required basic data preprocessing, where pixels were stored as integers in array format. To facilitate easy evaluation, the images were converted to black and white. Preprocessing steps that have been recommended are RGB to grayscale, median filtering, and contrast enhancement. The research involves a range of image processing techniques to remove noise and hair, convert RGB to grayscale, and smooth using a Gaussian filter. The investigation employs methods like SVM and GLCM, and the results of the simulation indicate that SVM is the most suitable approach. The article focuses on thresholding methods, which include edge, mean, and max–min techniques for region growing and merge split blocking. For image interpretation, gray values are utilized, and edge detection is employed to compare boundary estimation and region-based segmentation. Several segmentation methods, including Otsu's method, gradient vector flow (GVF), and color-based image segmentation using K-mean clustering, are utilized in the skin cancer detection process. The authors propose a multi-stage approach that involves image acquisition, as well as various techniques such as median filtering, Dull Razor software, contrast enhancement, and image resizing. For image segmentation, k-means clustering and ROI clustering are used, and feature extraction techniques such as ABCD rule-based, GLCM, and HOG are implemented. The geodesic active contour (GAC) is used to perform the segmentation process. Paper [1] describes an artificial intelligence (AI) system for detecting skin pathologies. The system uses a convolutional neural network (CNN) to analyze images of skin lesions and classify them as benign or malignant. CNN was trained on a dataset of over 10,000 images and achieved an accuracy of 95%. It resulted in a study according to which, it is observed that the segmentation task is not affected by the issue of imbalanced data classes, and hence, the results obtained are quite promising and comparable to the latest techniques. However, when it comes to the classification task, we still face limitations due to the data classes imbalance problem, even after implementing various methods to reduce it. Nevertheless, the findings from our experiments hold significant value [2]. The evaluation of quantification of asymmetries on the spectral bands of malignant melanoma using six sigma threshold as a preprocessor is being conducted. This paper [3] introduces a novel approach to measure the asymmetry of pigmentation in skin lesions. The areas affected by melanoma caused by excessive melanin secretion are segmented using six sigma thresholding. The next step in this study aims to automatically detect the level of fuzziness present in the borders of the melanoma [4]. The article presents a novel method for skin lesion classification and detection of novel classes, referred to as "learning a meta-ensemble." This method consists of a two-step algorithm. The first step uses a stacking network for base learners, which has shown to outperform simple averaging. The second step incorporates the CS-KSU Module Collection, which has been successful in detecting novel classes. Additionally, replacing the ResNet-18 components in the CS-KSU Module Collection with other advanced pre-trained models has the potential to further enhance the algorithm's performance [5]. This paper discusses the potential of utilizing other datasets for medical imaging classification and the importance of selecting appropriate auxiliary datasets. The findings indicate that using an ImageNet pre-trained model provides better initialization for MTL than random or fine-tuned models. However, it is crucial to carefully select the auxiliary dataset. The study focuses on the similarity between white-light colonoscopy and fNBI colonoscopy, as both have similar imaging mechanisms and feature fine-grained variability [6]. The use of AI classifiers in analyzing skin cancer images is discussed in this study. A thorough search of three databases yielded 526 publications, with a majority found in ISI Web of Science. The paper emphasizes the importance of the image analysis and processing methods employed, as they affect the performance of the classifier. [7] The aim of this research was to classify different technologies that utilize AI to identify and categorize skin cancer. A total of 906 papers were collected from three databases, and after analysis, 53 papers were deemed relevant. Out of these papers, 14 utilized shallow AI techniques, while 39 used deep AI techniques. The studies employed up to 11 different metrics to evaluate their models, with accuracy being the primary metric in 39 of the studies. Generally, studies with smaller datasets reported higher accuracy. However, comparing the methods directly was challenging due to variations in evaluation metrics and image types used. The performance scores of the models investigated in this paper were influenced by various factors, including the size of the dataset, the number of diagnostic classes, and the techniques employed [8]. A detection system based on convolutional neural networks. In the preprocessing stage, labeled data is used to ensure accuracy. During the training phase, the system achieves a steady rate of around 97% accuracy in matching intents to question–answer pairs. However, as more data is collected and entered into the database, the accuracy may decrease slightly. Overall, the model performs the task with considerable accuracy [9]. The proposed strategy in the mentioned article aims to develop an efficient automated system for skin cancer classification using CNNs. The results of their experiments show that their proposed system outperformed other classifiers in terms of accuracy and other evaluation metrics. This suggests that their approach could be a viable option for developing a reliable and accurate automated system for skin cancer classification [10]. The utilization of a genetic algorithm-based color image segmentation technique for detecting skin cancer in skin lesion images was presented in the paper. The algorithm was successful in segmenting the images, but future research will involve an improved region merging approach to isolate the lesions from the rest of the image background [11]. This study employs deep learning neural networks to classify skin cancer. The Inception-v4 convolutional neural network architecture is used to train and test the model on a large dataset of skin lesion images. The study finds that the Inception-v4 architecture outperforms other CNN architectures and achieves an area under the ROC curve of 0.94. The researchers also compare the performance of their model with dermatologists and observe that the deep learning neural network performs similarly to experienced dermatologists [12, 13]. These papers demonstrate the ability of deep neural networks to achieve dermatologist-level accuracy in the classification of skin cancer using images and a hybrid deep neural network approach for skin lesion classification that combines both convolutional and recurrent neural networks.

3 Existing Methodology

Comparison of more detailed models can be shown in Table 2, which highlights in greater details the tools used along with preprocessing techniques in some places (Table 1; Figs. 1 and 2).

4 Evaluation Metrics

To determine the accuracy of a skin cancer classification method using convolutional neural network (CNN), evaluation metrics play a crucial role. There are several commonly used metrics to evaluate a CNN model for skin cancer classification. The following are some of those metrics. Accuracy: This metric measures the percentage of correctly classified samples. It is a commonly used metric to evaluate a classification model. However, it may not be appropriate for imbalanced datasets where one class dominates. Precision: It measures the fraction of true positive predictions out of all positive predictions. It is useful when reducing false positives and ensuring that the positive predictions are correct. Recall: Recall is a performance metric that

Table 1 Overview of different machine learning techniques applied to cancer classification in existing methods

S. No.	Architecture	Accuracy
1	Inception-v4	89%
2	Inception-ResNet v2	92%
3	ResNet-152	94%
4	DenseNet-201	95%
5	Ensemble (Inception-v4, Inception-ResNet v2, ResNet-152, DenseNet-201)	96.5%
6	Deep learning neural networks	91% sensitivity for melanoma detection
7	Hybrid deep neural networks (DNN) CNN, DBN, SVM	Overall accuracy: 89.5%
8	Support vector machine	93.42%,
9	Decision tree	83.16%
10	K-nearest neighbor	88.06%
11	Random forest	91.94%
12	*Proposed method	92%

Table 2 Overview of classification methods with preprocessing techniques

S. No.	Algorithm	Accuracy (%)	Preprocessing technique	Dataset size
1	Support vector machine (SVM)	96.67	Feature extraction using histogram of oriented gradients (HOG)	200 dermoscopy images
2	K-nearest neighbor (KNN)	96.33	Feature extraction using color histogram, HOG and local binary patterns (LBP)	180 images
3	Ensemble classifier	98.3	Integration of SVM, KNN, and random forest algorithms	1200 images
4	Deep belief network (DBN)	91.77	Feature extraction using color histogram, LBP, and Gabor filter	1000 images
5	*Proposed method (CNN)	93.25	Transfer learning with Inception V3 model	2000 images

quantifies the proportion of actual positives that are correctly identified as positives. It is particularly useful for minimizing false negatives and ensuring that all actual positives are detected.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3}$$

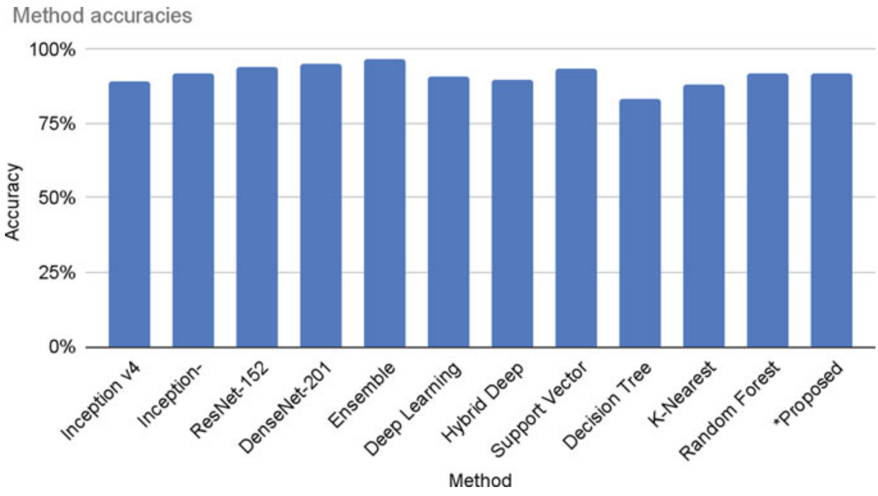


Fig. 1 Methods and accuracy

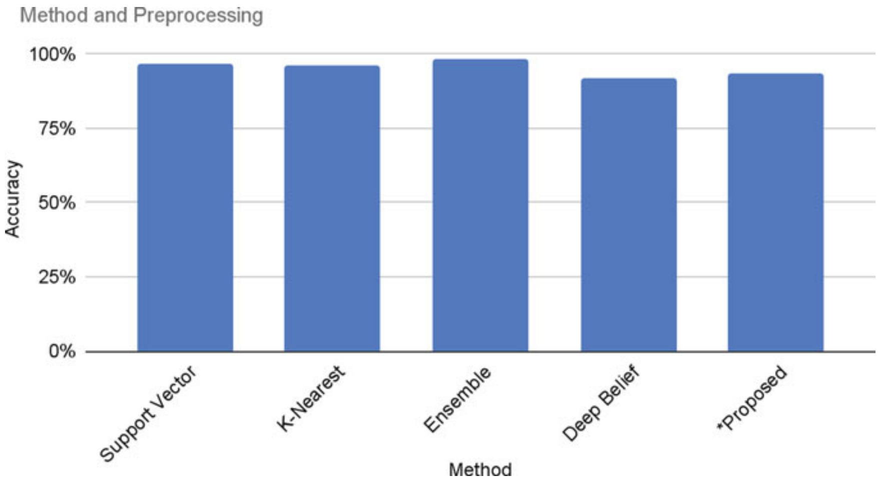


Fig. 2 Method and preprocessing

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

On the other hand, F_1-score is a metric that takes into account the harmonic mean of both precision and recall, thereby enabling a balance between these two performance measures. A model's capability to differentiate between positive and negative classes can be quantified by using the area under the receiver operating characteristic curve (AUC-ROC) metric.

5 Proposed System

A software system designed for skin cancer classification using machine learning and convolutional neural networks involves several key components and processes. These include data collection and preprocessing, model selection and training, model evaluation and testing, deployment, and continuous improvement. Data collection involves obtaining a large dataset of skin images and annotating them with labels indicating whether they are cancerous or non-cancerous. Figure 3 shows the structure of our application. Users can upload the image from the mobile application. It will then come to our backend APIs. The API used is made using flask which is a Python-based Web application framework. It will be loaded in a model which is pre-trained in order to detect skin cancer among other skin anomalies. Our model will predict the result which will then be sent to the users' device through the architecture shown above. The user can then make use of this prediction in order to decide whether or not he/she should seek further medical counsel. Figure 4 shows the structure of our mobile application. It entails the user interface components of our proposed mobile application. We have three major sections which will each have their respective subsections. To create the project we began by creating a machine learning model that would be able to take in input images. These images will be colored and of high resolution. Our model should be able to process these images and accurately classify the images based on the type of skin cancer it is able to predict. We then mapped the model which was loaded into a flask with routes. We called these routes from the mobile application. This product is merely a preliminary examination which people can access from the comfort of their homes. The application provides a fair level of accuracy but does not claim to be the final authority on the matter. This authority is exclusive to medical experts sanctioned as such by the government.

6 Results and Discussion

The dataset utilized in this study is HAM10000 (human against machine with 10,000 training images), which is widely employed for skin cancer classification via machine learning and convolutional neural networks. The dataset comprises 10,015 dermatoscopic images depicting both benign and malignant skin lesions, which were gathered from the Department of Dermatology at the Medical University of Vienna in Austria. The dataset contains seven classes of skin lesions, including basal cell carcinoma, melanoma, and seborrheic keratosis. Each image is labeled with a unique identifier and a diagnosis, as determined by an expert dermatologist. In addition, each image

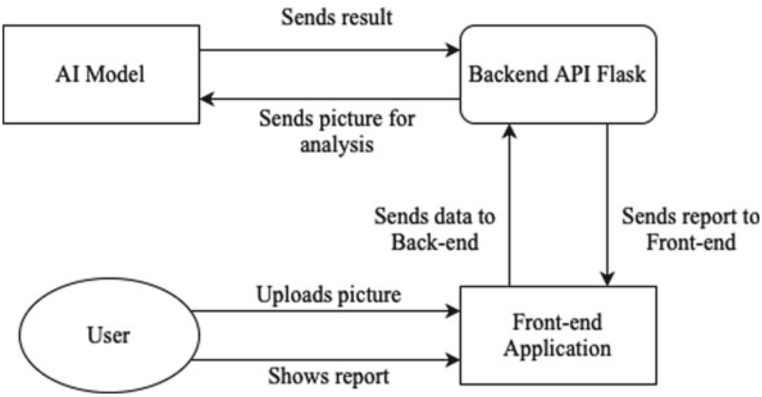
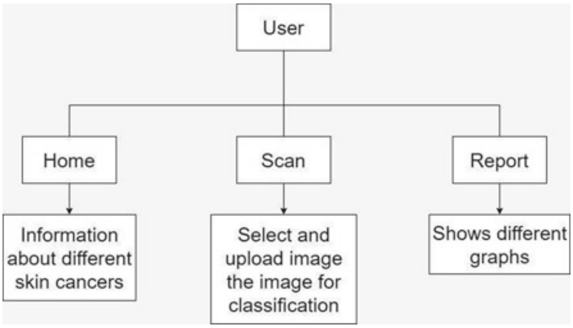


Fig. 3 Proposed method architecture

Fig. 4 Flowchart of system interface



is accompanied by clinical metadata, including patient age, sex, and anatomic location of the lesion. To start, the model takes the image column of the DataFrame and converts it into a numpy array. The images' pixel values are normalized by scaling them between 0 and 1. The images are then assigned corresponding labels, and a one-hot encoded format is applied using the to_categorical() function. The dataset is then divided into training and testing sets using the train_test_split() function. A Sequential() class from Keras is used to define the CNN model. It includes three convolutional layers of different filter sizes, each followed by a MaxPooling2D() layer to reduce the dimensions of the output feature maps.

Figure 5 shows one trend in the training accuracy values is a steep increase in accuracy from epochs 1–10, followed by a gradual increase from epochs 10–30. From epochs 30–40, the training accuracy values remain fairly stable between 86 and 89%, followed by a slight increase from epochs 40–50. Another trend is that the training accuracy values fluctuate between epochs, indicating that the model's performance is not entirely consistent. The training loss values in Fig. 6 indicate how well the model is performing during the training process. Ideally, we would like the loss to decrease as training progresses, which would indicate that the model is learning and improving. Looking at the provided array, we can see that the training loss values start relatively high at 92% and gradually decrease over time, reaching a minimum of 31%. This indicates that the model is indeed learning and improving over the course of the training process. We can also observe that there are fluctuations in the loss values, which is expected and can be attributed to various factors such as batch size, learning rate, and model architecture. However, the general trend is that the loss values are decreasing over time, which is a positive sign. Overall, the decreasing trend in the training loss values indicates that the model is learning and improving as it trains, which is a strong point of the model. The validation loss array for the skin cancer image classification model displays a downward trend.

This implies that the model's performance in the validation set improved as the number of epochs increased. The trend shows a significant dip in the loss values during the initial epochs, followed by a steady decrease with fluctuations. The loss values gradually converge after around 25 epochs, indicating that the model performance may not improve significantly with more epochs. The validation loss values

Fig. 5 Change in accuracy over epochs

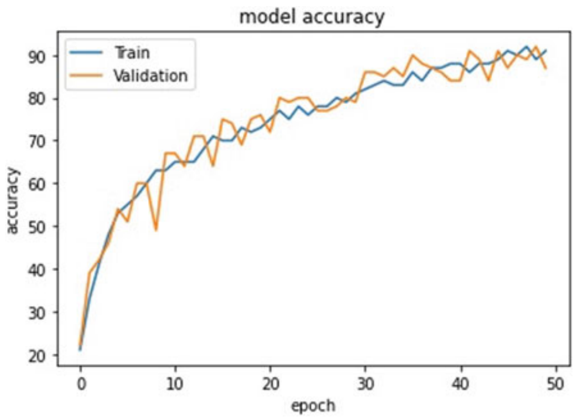
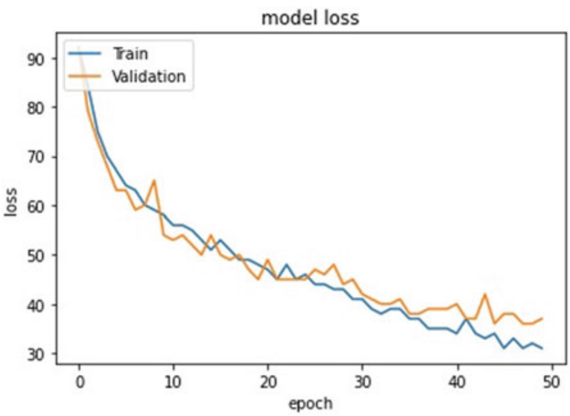


Fig. 6 Change in loss over epochs



appear to be slightly higher than the training loss values, but the difference is not significant. This suggests that the model is not overfitting or underfitting. The model is generalizing well to unseen data, as the validation loss is decreasing along with the training loss.

7 Conclusion

We found that the data classes imbalance problem still exists in the classification task, although the segmentation task yielded promising results. Investigated a new method for quantifying the asymmetry of skin lesion pigmentation, and read a proposal of a two-step algorithm that showed promising results for skin lesion classification and novel class detection. Read that an ImageNet pre-trained model serves as a better initialization for MTL in medical imaging classification. Realized the significance of the adopted image analysis and processing methods in skin cancer detection. These studies show the potential of AI-based technologies in detecting and classifying skin cancer, but also highlight the challenges that need to be addressed, such as data classes imbalance problem, image analysis and processing methods, and varied evaluation metrics. Based on the analysis of the training and validation accuracy and loss arrays, we can conclude that the CNN model made for classifying skin cancer images has performed well. The model seems to have learned well from the training data as evidenced by the steady increase in training accuracy from 21 to 92%. However, there is a possibility of overfitting as the training accuracy is much higher than the validation accuracy. Although the validation accuracy increased gradually from 22 to 92%, there were some fluctuations throughout the training process, which indicates that the model might not generalize well to unseen data. Similarly, the training loss decreased steadily from 92 to 31, while the validation loss fluctuated a bit but eventually reached a low value of 37. The drawbacks are that the dataset used in this research contains 10,015 images, which is a relatively small dataset for deep learning models. Using a larger dataset may have improved the generalization performance of the model. Other evaluation metrics that could have been used to assess the performance of the model, such as precision, recall, and F_1-score. Real-life application and clinical trials are not discussed.

References

1. Gavrilov D, Lazarenko L, Zakirov E (2019) AI recognition in skin pathologies detection. In: 2019 International conference on artificial intelligence: applications and innovations (IC-AIAI), 2019, pp 54–542. https://doi.org/10.1109/IC-AIAI48757.2019.00017
2. Lin C-H, Lin C-S, Chou P-Y, Hsu C-C (2021) An Efficient data augmentation network for out-of-distribution image detection. IEEE Access 9:35313–35323. https://doi.org/10.1109/ACCESS.2021.3062187

3. Mishra J, Ghosh S, Raj A (2013) Uncertain query processing using vague sets: a better approach. In; Third international conference on computational intelligence and information technology (CIIT 2013), 2013, pp 52–60. https://doi.org/10.1049/cp.2013.2572
4. Bagchi S, Banerjee A, Bathula DR (2020) Learning a meta-ensemble technique for skin lesion classification and novel class detection. In: 2020 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW), 2020, pp 3221–3228.https://doi.org/10.1109/CVPRW50498.2020.00381
5. Shang H, Sun Z, Fu X, Zhang Z, Yang W (2019) What and how other datasets can be leveraged for medical imaging classification. In: 2019 IEEE 16th International symposium on biomedical imaging (ISBI 2019), 2019, pp 814–818. https://doi.org/10.1109/ISBI.2019.8759148
6. Reichrath J, Leiter U, Garbe C (2014) Epidemiology of melanoma and nonmelanoma skin cancer—the role of sunlight. In: Reichrath J (ed) Sunlight, vitamin D and skin cancer. Springer, New York, NY, 89–103
7. Takiddin A, Schneider J, Yang Y, Abd-Alrazaq A, Househ M (2021) J Med Internet Res. https://www.jmir.org
8. Abuared N, Panthakkan A, Al-Saad M, Amin SA, Mansoor W (2020) Skin cancer classification model based on VGG 19 and transfer learning. In: 2020 3rd International conference on signal processing and information security (ICSPIS), pp 1–4
9. Amelio A, Pizzuti C (2013) A genetic algorithm for colour image segmentation. In 16th European conference on applications of evolutionary computation, evoapplications 2013, pp 314–323
10. Arbelaez P, Maire M, Fowlkes C, Malik J (2011) Contour detection and hierarchical image segmentation. IEEE Trans Pattern Anal Mach Intel l, 33(5):898–916
11. Tschandl M et al (2018) Skin cancer classification using deep learning neural networks. J Am Acad Dermatol 78(6):1170–1172
12. Esteva A, Kuprel B, Novoa R et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:115–118. https://doi.org/10.1038/nature21056
13. Yu Y et al (2018) Skin lesion classification using hybrid deep neural networks. J Imaging 4(11):134

AI-Based Usability Prediction for 3BHK Apartment Unit

Kshitij Kumar Sinha, Manoj Mathur, and Arun Sharma

Abstract Anything that adds value to life should be treasured. A well-designed space improves quality of life, thereby becoming the value component in itself. The design value associated with a space can be spelled through *Firmitas* (structural), *Utilitas* (functional), and *Venustas* (aesthetics), the three principles of architecture as explained and promoted by Vitruvius. The *Utilitas* or the functional component of a design demands careful value assessment, as most of its parameters remain hidden from the normal interpretation. The gap area of the 'missing utility' component in the residential real estate segment is established by investigating the current practices of different stakeholders. Taking 3BHK (three bedroom hall kitchen) typologies offered in residential apartment segment in Delhi NCR as the prime area of study, this research proposes an AI-based tool driven on fuzzy inference system (FIS) to predict the usability (also referred as utility value or as *Z* value) of a 3BHK apartment layout based on the use value of its bedrooms, hall, and kitchen. The proposed *Z* value has the potential to substantiate the pricing model in the real estate industry.

Keywords Utility value · Design layout · Apartment typology · Fuzzy inference system

K. K. Sinha (✉) · M. Mathur
School of Planning and Architecture, New Delhi, India
e-mail: Kshitijkumarsinha@igdtuw.ac.in

M. Mathur
e-mail: m.mathur@spa.ac.in

K. K. Sinha · A. Sharma
Indira Gandhi Delhi Technical University for Women, New Delhi, Delhi, India
e-mail: arunsharma@igdtuw.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_45

1 Introduction

Design creates culture. Culture shapes values. Values determine the future. - Robert L Peters

Design is not just what it looks like and feels like. Design is how it works. – Steve Jobs

Design has the power to change the world. The values associated with a good design are transformative in nature. It enhances efficiency, increases productivity, and upholds the quality of life. Just like cosmic energies, which are there but could not be seen; design is omnipresent as it impacts our life in many ways. Architecture, which is often equated as the science of living therefore, should also find clear reflection of its design value in the selling, buying, or renting of the built structures. Vitruvius has laid down the principle of understanding architecture of a building under three components—*Firmitas* (related to the structural component or the firmness of the built form), *Utilitas* (related to the functional aspects of the spaces created in a built form), and *Venustas* (referring to the aesthetics or the beauty component of the built form). It can further be said that the design value associated with an architectural design, derives its essential features from these three attributes.

This paper is an attempt to show how design value could be assigned to an architectural design (in this case a residential apartment layout) from its utility perspective. The paper, investigating the current practices by interviewing the experts from different domains of Indian Real Estate Industry, establishes the missing reflection of design value in the pricing of residential apartments. In addition, it contributes by proposing an innovative way of determining the utility score (denoted as '*Z*' Value) of a 3BHK (three bedroom hall kitchen) typology apartment layout based on over 79 different qualitative and quantitative parameters of bedrooms (35 parameters of each bedroom, 25 parameters of kitchen & 19 parameters of hall—referred also as dining/drawing). This paper is an extension to the earlier paper [1]; wherein fuzzy logic-based tool of predicting the *Z* value of bedrooms was developed based on 29 parameters. Applying the same logic to 5 other spaces of more than 170 layouts of various three BHK typologies, the *Z* value of the entire apartment is calculated. The study presents a structured way of showing how well individual spaces (in this particular case—the three bedrooms, dining/drawing room and kitchen) are designed with respect to the overall utility of an apartment, thereby becoming a comparative utility assessment tool for various spaces associated with a 3BHK designs offered by different developers. The findings are stated as crucial as it demonstrates and establishes why design value should be clearly reflected in the pricing of the commodities.

2 Enquiring the 'Practice': Expert Interviews and Literature Survey

In order to investigate the practice and use of utility value assessment, a detailed interview was conducted with professionals of relevant knowledge and expertise. Stakeholder perspective of developers, real estate advisors (REA), real estate financial institutions (REFI's), architects, project management consultants (PMC), and users were collected to understand the present dynamics, identify the gap areas and to assess the usage of the term utility by them in their decision making.

The investigation, which was initially targeted at architects who have past experience of working with developers of international repute, after a few interviews became organic and demanded expansion of the panel to include professionals from different profiles also (as listed in the table above). The interview revolved around the central spine of key investigation points such as:

- Design-based strategies adopted by developers for attaining competitive advantage
- Ways of assessing the use value of an apartment
- Positives and negatives of the current design practices
- Factors considered while making design decisions
- Importance given to indoor environment quality (IEQ) parameters like thermal comfort, visual comfort, acoustic comfort, and indoor air quality in the project briefs or design briefs
- Promotion of 'design'-driven culture by leading developers of the country
- Green building rating trends
- Relevance of research studies (like post occupancy studies) and their usage by various stakeholders
- Factors to consider for calculating the 'use value' of an apartment.

Through interviews it was found that in case of real estate financial institutions (REFI's) the question of utility value does not appear anywhere in the key decision of deciding the product mix. In case of real estate advisors (REA), the stress on utility value gets overshadowed by the market assessments findings and number crunching exercise to show the profitability of the project. In developer's case, it was observed that utility of the design is limited by the knowledge and risk taking capacity (both driven by past experience and intuition) of the people handling such projects. Is it therefore justified, to carry out a detailed investigation of utility exploration which could present new set of avenues and improvements in spaces that are being offered for living. In the case of architecture practice/consultancy, it was evident that the design value associated with the apartment layouts have not been communicated as the major selling points. This is mainly because of the missing numerative assessment tools for calculating design values.

Broom [2] relates the principle of utility to generating benefits, advantages, pleasure, pleasantness, or happiness. Kapteyn [3] emphasized the significance of any product possessing inherent traits that explain its broad usefulness. Kaklauskas et al.

[4] went on to say that the usefulness might be used to estimate the market worth of real estate. Phipps [5] emphasized the need for ever more complex utility function models that account for changes to attribute utilities and importance weights in addition to functional form modifications. A claim made in support of this idea by Lancaster [6] was that a good has no inherent utility. Instead, it is made up of traits that have served as utility. The market price a consumer pays for a certain good is affected by the value of these traits.

The decision to buy or rent a home is not easy, especially in a country like India. This involves lifelong savings, it demands careful risk assessments, it asks for daily sacrifices, and it deals with so much solicited and unsolicited advice and so on. The complexity in fact increases in metropolitan cities of the country like Delhi NCR. The list may not end up soon, and even after ticking all the boxes, one may still find themselves falling short of making a perfect match to meet the requirements. This is evident through the alterations done post occupancy of the dwelling units. It so happens that even after every ounce of money, time, and energy the end user ends up having a very unpleasant experience in the entire exercise. The dilemma could range from the initial investment stage justifying the amount spent while moving in to the operational expenditure which one may have to bear while occupying the property, to the aspects of profits that the investment may give back. This covers aspects of 'value for money', 'utility value', and 'return on investment' (ROI) concepts. The values associated with the property, till recently, have been dependent on its location or the way it has been marketed. The utility value, which puts a value to a space for its design to be of utility, has remained hidden from the end customers/occupants. What best is communicated to the customer is the specifications of materials used in the apartment and the typology of spaces that they are investing in. As a result, the dilemma's listed above start to seep in.

Sinha et al. [1] demonstrate a fuzzy logic-based system for calculating the utility value of a bedroom of 3BHK layout. It further argues that the Hedonic pricing model can be substantiated with the utility value of the layouts, thus bringing a check on the speculative market pricing practices and justifying the 'value for money' and 'return on investment' components associated with a property.

3 Methodology Adopted

This paper can be understood as an outcome of two parallel work packages developed simultaneously for validation of the findings. The first package was the investigation works and comprised of conducting detailed interviews with real estate experts working in the various profiles of developers, architects, project management consultancy, real estate advisory (REA), real estate financial institutions (REFI), user having design understanding and architectural firms. The interviews helped in establishing the current practices in the industry from different stakeholders' perspective when it comes to utility value of apartments. Also, the interviews were framed for identifying qualitative and quantitative parameters for the apartment spaces like dining drawing room (also referred as hall at some places) and kitchen. The expert opinion was also taken to assess the utility of 3BHK apartment layouts. The layouts were assigned a score in percentage (later converted to a scale of 0–1) which later on became the output variable of Z value of apartments.

The second package of works can be classified as extension works and focused on calculating the utility value (Z value) of the dining drawing and kitchen spaces. This is classified as extension works because a system to calculate the Z value of bedroom was already developed in the previous works of Sinha et al. [1] based on the three inputs, namely quantitative parameters (X value), qualitative value (Y value), and area of apartment. Once the quantitative and qualitative parameters for dining drawing spaces and kitchen spaces were deduced, the Z value of these spaces was also computed in the similar manner as it was deduced for bedroom.

3.1 Input/Output Variables

The Z value of all three bedrooms (also denoted as bedroom 1, bedroom 2, and bedroom 3), dining drawing (hall), and kitchen became the five input variables for developing the FIS. These values were further classified into three ranges, namely low, medium, and high. The Z value of the overall apartment (obtained through the expert survey) acted as the output variables. These values were also categorized as very low, low, medium, high, and very high (refer Fig. 2 for ranges).

The expert opinion was also taken to assess the utility of 3BHK apartment layouts. The layouts were assigned a score in percentage (later converted to a scale of 0–1) which later on became the output variable of Z value of apartments.

3.2 Set of Rules

Mapping of low, medium, and high Z values for bedroom 1, bedroom 2, bedroom 3, dining drawing (hall), and kitchen was done with the very low, low, medium, high, and very high Z value of overall apartment obtained through the expert survey was done to prepare a data set of over 170 layouts of 3BHK apartments. Out of these, some cases were more or less similar. As a result, we have 123 unique cases, including the variation in overall Z value of the apartment (i.e., the output variable). However, if we ignore the variation of the output variable and only examine the input variable, we were able to find 97 unique cases. The noise in the data was further removed by dropping the ambiguous cases when observed with reference to the expert survey findings. So, in total 86 unique set of rules were made to be utilized further for deployment of fuzzy inference System.

The result which was obtained when validated with expert's assessment (carried out at a later stage for 25 plans other than the previously shortlisted) presented a strong co-relation value of 0.82. The findings also include various trends to the Z

value of bedrooms, dining drawings, and kitchen when observed in reference to the Z value of the apartment, becoming a very efficient tool for comparison of 3BHK layouts based on their utility score.

Few suggested rules include:

- If bedroom 1 and bedroom 2 has low Z score, while bedroom 3, dining drawing and kitchen have medium Z score, then the overall Z value of the apartment will be low.
- If bedroom 1 has low Z score, bedroom 2 and bedroom 3 have medium Z score, and the Z score of dining drawing and kitchen are high; then, the Z value of the apartment will be medium.
- If bedroom 1 and bedroom 2 has medium Z score, while bedroom 3, dining drawing and kitchen has high Z value; then the Z value of the apartment will be high.

4 The Proposed Fuzzy Inference System

Çekmiş [7] states that in order to deal with the ambiguity, vagueness, and imperfection prevalent in architectural and urban environment, fuzzyness is a beneficial tool. Using a fuzzy logic framework, Nadin [8] explained how their analysis applies the aesthetic criteria as fuzzy control rules. Ciftciogh [9] asserts that fuzzy logic has been used successfully in architecture projects where decision making necessitates not only the study of quantitative data but also the evaluation of non-numerical data and abstract notions. Adding further, that for architectural design, fuzzy logic is especially fascinating.

As per Durmisevic [9], fuzzy logic, although not commonly used, has been acknowledged as a workable mathematical model for architectural design to process non-quantifiable data. Valiyev et al. [10] uses fuzzy logic to evaluate inside daylight confirming deployability of fuzzy information systems for architectural design evaluations. Sinha et al. [1] have shown how fuzzy logic could be applied for calculating the utility score of a bedroom with quantitative value, qualitative value and area forming the input variables on a scale of low, medium, and high and deployed through a rule set of 27 rules to predict the output as utility score or the Z value of the bedroom of an apartment layout in five ranges of very low, low, medium, high, and very high (Figs. 1, 2, 3, 4 and 5).

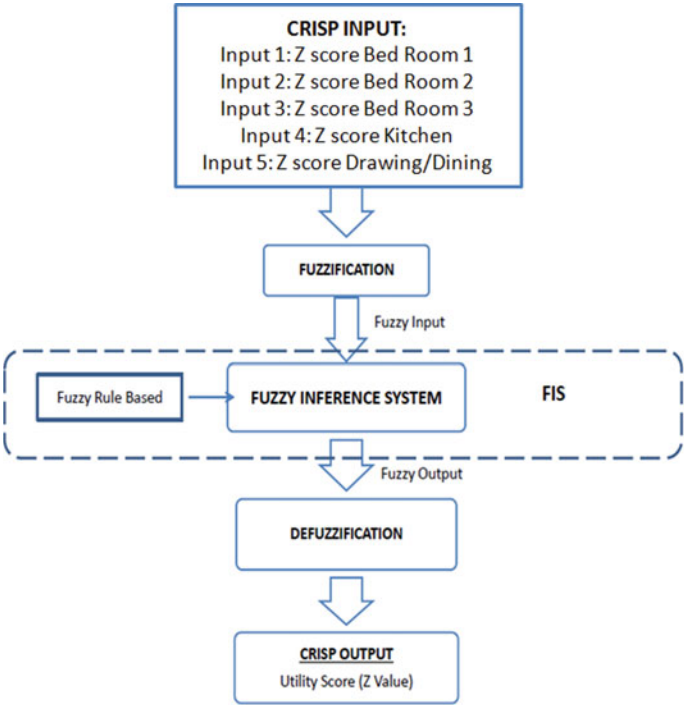


Fig. 1 Proposed FIS architecture. *Source* Author

5 Results and Discussions

5.1 Relationship Between Z Value of Apartment and Z Value of Bedroom (1, 2, and 3)

In the above graph (refer Fig. 6), it can be observed that the z value for bedroom 1 is almost always higher than the Z value of the whole plan. Majority of the Z values for bedroom 1 lie between the range of 0.8–1.0, with over half of them falling between 0.9–1.0, indicating good design practices. However, Z value of bedroom 1 may not be a good indicator of the expected Z value of the complete plan. Plans with Z value as low as 0.5 also have bedroom 1 with a Z value > 0.8. Therefore, it can be inferred that in almost all cases, bedroom 1 is well-designed, irrespective of the rest of the housing.

4.73% of sample projects have Z values of the bedroom lower than the Z value of the project ($z < Z$)

56.8% of bedroom 1 z values falling between 0.9–1.0

4.73% of bedroom 1 z values lower than 0.7.

System	Input	Output
Name='Z Value Apartment' Type='mamdani' Version=2.0 NumInputs=5 NumOutputs=1 NumRules=86 AndMethod='min' OrMethod='max' ImpMethod='min' AggMethod='max' DefuzzMethod='centroid '	**Input 1:** Name='Z_BR1' Range=[0 1] NumMFs=3 MF1='low':'trimf',[0.6 0.65 0.8] MF2='medium':'trimf',[0.76 0.8 0.9] MF3='high':'trimf',[0.88 0.95 0.99] **Input 2:** Name='Z_BR2' Range=[0 1] NumMFs=3 MF1='low':'trimf',[0.5 0.62 0.75] MF2='medium':'trimf',[0.72 0.78 0.85] MF3='high':'trimf',[0.82 0.9 0.99] **Input 3:** Name='Z_BR3' Range=[0 1] NumMFs=3 MF1='low':'trimf',[0.5 0.6 0.7] MF2='medium':'trimf',[0.68 0.75 0.825] MF3='high':'trimf',[0.8 0.9 0.99] **Input 4:** Name='Z_KT' Range=[0 1] NumMFs=3 MF1='low':'trimf',[0.465 0.625 0.8] MF2='medium':'trimf',[0.78 0.85 0.9] MF3='high':'trimf',[0.875 0.925 0.99] **Input 5:** Name='Z_DDH' Range=[0 1] NumMFs=3 MF1='low':'trimf',[0.35 0.55 0.75] MF2='medium':'trimf',[0.725 0.78 0.85] MF3='high':'trimf',[0.825 0.9 0.99]	Name='Z_APT' Range=[0 1] NumMFs=5 MF1='very_low':'trimf', [0.45 0.52 0.6] MF2='low':'trimf',[0.58 0.65 0.7] MF3='medium':'trimf',[0 .68 0.75 0.8] MF4='high':'trimf',[0.7 8 0.85 0.9] MF5='very_high':'trimf' ,[0.88 0.9 0.92]

Fig. 2 Parameters of fuzzification process. *Source* Author

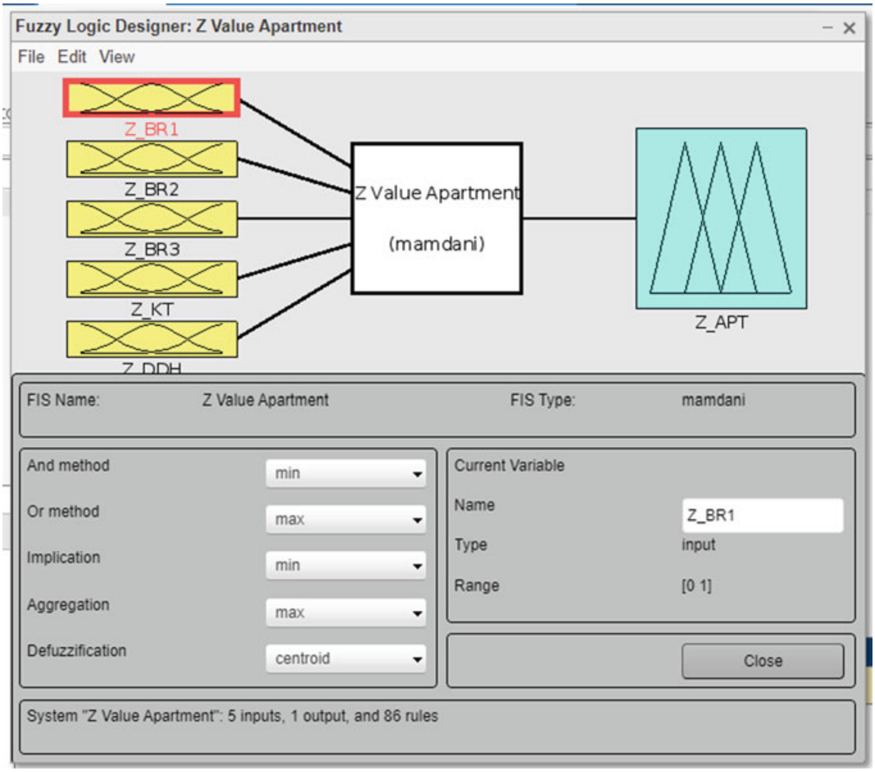


Fig. 3 Fuzzification of input into output (*Z* value). *Source* Author

The *Z* values for bedroom 2 show a considerable amount of variation with highest as well as lowest peaks indicating a wide range (refer Fig. 7). Similar to bedroom 1, very few plans have a lower *Z* value for bedroom 2 than the *Z* value for the plan. Curious overlap of the *Z* value of bedroom 2 and *Z* value of the plan can be witnessed in the above graph. It can be inferred that bedroom 2 is a good indicator of the *Z* value of the plan because of this overlap.

19.53% of sample projects have *Z* values of the bedroom lower than the *Z* value of the project ($z < Z$)

18.34% of bedroom 2 *Z* values falling between 0.9–1.0

11.83% of bedroom 2 *Z* values lower than 0.7.

Bedroom 3 exhibits the maximum variation in the *Z* values recorded (refer Fig. 8), with a significant amount of troughs in the above graph representing poorly-designed bedroom 3. Thus, it can be inferred that bedroom 3 is one of the most compromised spaces in a housing design. Conversely, bedrooms exhibiting peak values (> 0.9) are almost all from plans with higher *Z* values (> 0.8). Therefore, bedroom 3 also acts a good indicator to differentiate between averagely designed and well-designed houses.

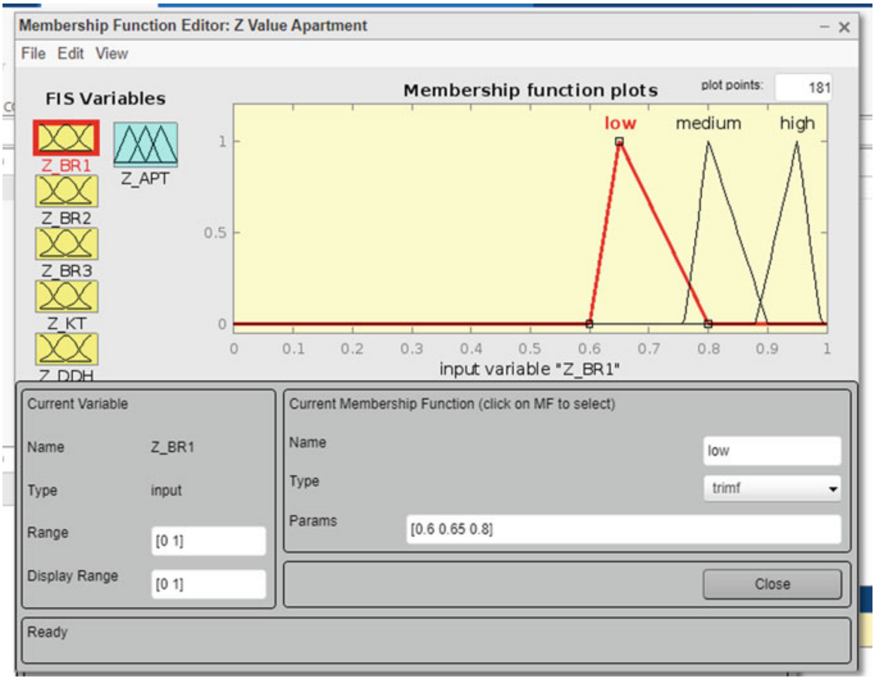


Fig. 4 MF for complexity attribute used in FIS. *Source* Author

33.73% of sample projects have Z values of the bedroom lower than the Z value of the project ($z < Z$)

14.79% of bedroom 3 Z values falling between 0.9–1.0

25.44% of bedroom 3 Z values lower than 0.7.

5.2 Relationship Between Z Value of Apartment and Z Value of Drawing Dining (Hall)

The Z value for drawing/dining is almost always higher than the Z value of the plan, indicating that the functionality of the drawing room is kept in mind in the majority of designs (refer Fig. 9). Additionally, Z value peaks and troughs correspond to respective Z value peaks and troughs in the above graph. This highlights a direct correlation between the Z value of the living/drawing room and the Z value of the plan.

12.43% of sample projects have the Z value of the drawing/dining room lower than the Z value of the project ($z < Z$).

55.62% cases have z values falling between 0.9–1.0

4.73% cases have z values lower than 0.6.

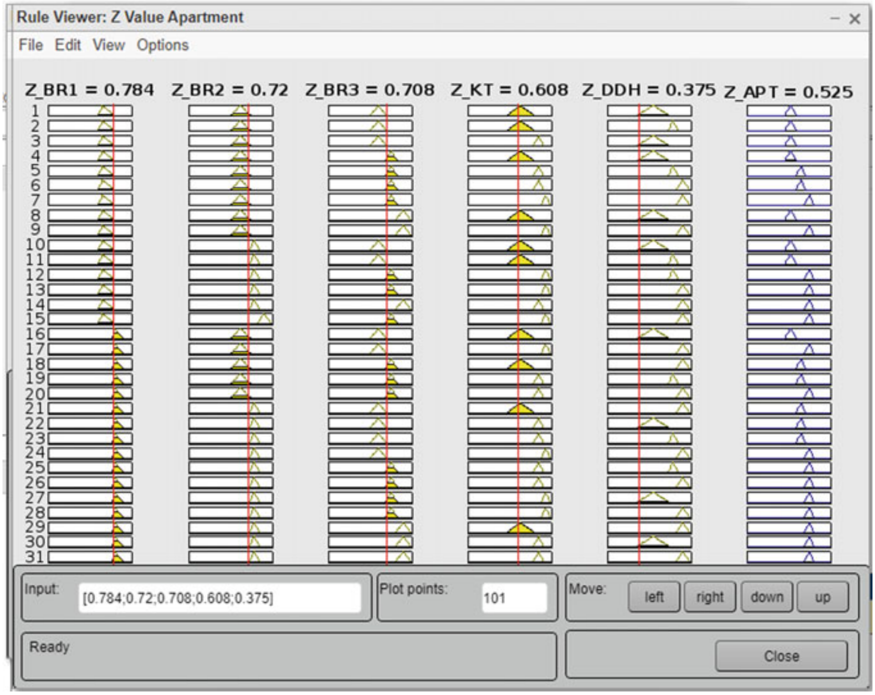


Fig. 5 Snapshot of rule viewer showing some of the rules. *Source* Author

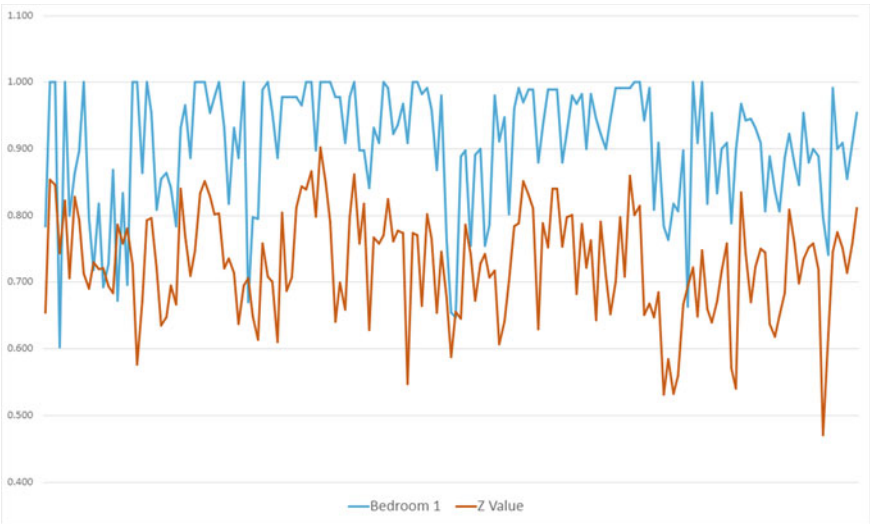


Fig. 6 Trends observed with z value of bedroom 1 to overall Z value of apartments. *Source* Author

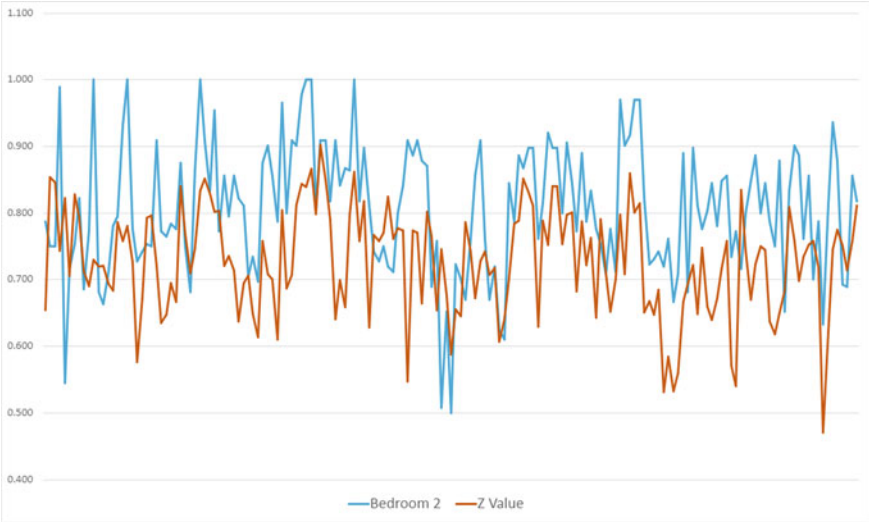


Fig. 7 Trends observed with z value of bedroom 2 to overall *Z* value of apartments. *Source* Author

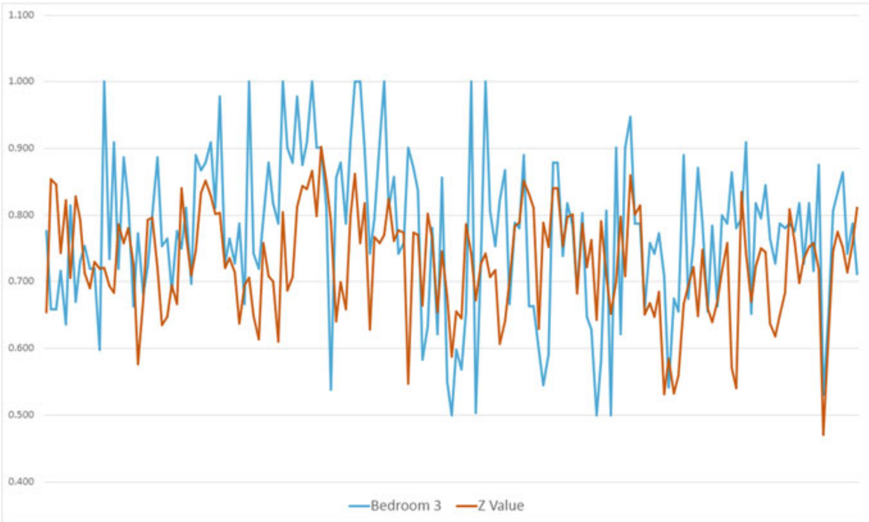


Fig. 8 Trends observed with *z* value of bedroom 3 to overall *Z* value of apartments. *Source* Author

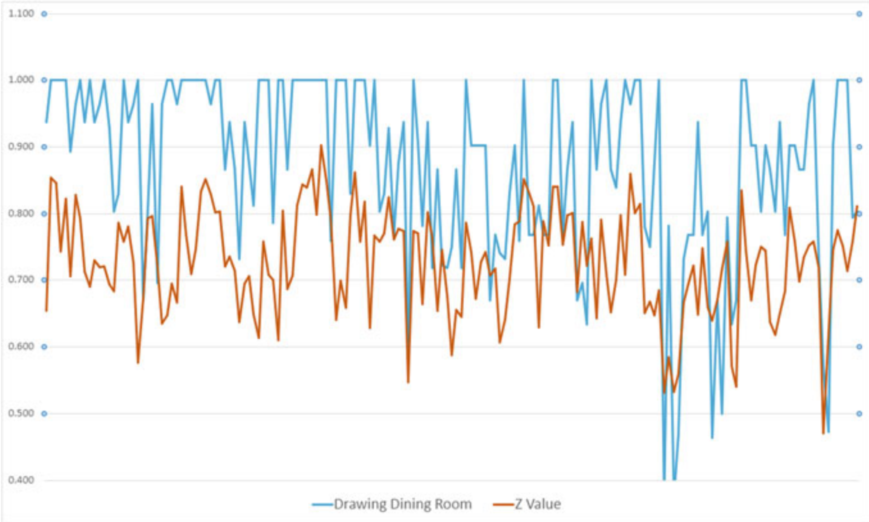


Fig. 9 Trends observed with z value of drawing/dining (hall) to overall Z value of apartments. *Source* Author

5.3 Relationship Between Z Value of Apartment and Z Value of Kitchen

See Fig. 10.

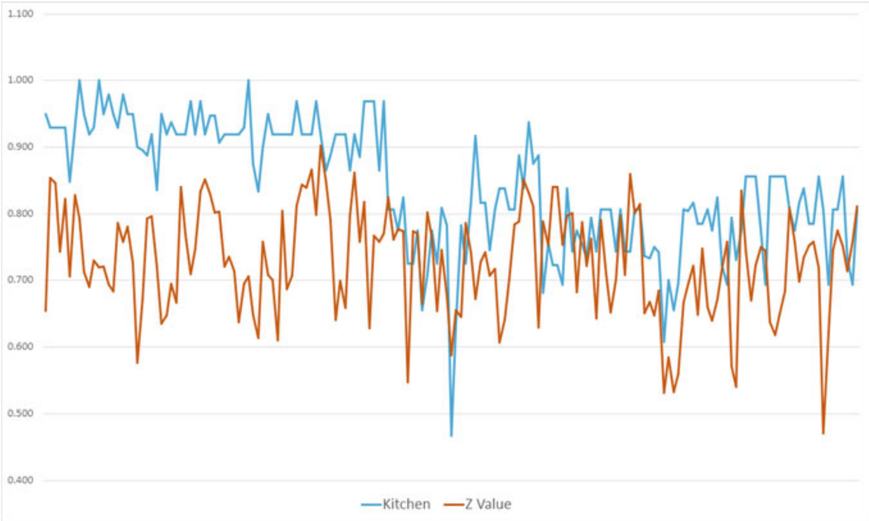


Fig. 10 Trends observed with z value of kitchen to overall Z value of apartments. *Source* Author

The graph represents kitchen having a Z value higher than the Z value of the plan for a majority of the plans reviewed. It can be inferred that kitchen design is a crucial component of housing and is almost always given priority. Kitchens with higher Z value belong to plans with a higher Z value, indicating a direct relation.

12.43% of sample projects have Z value of the kitchen lower than the Z value of the project ($z < Z$).

35.5% cases have Z values falling between 0.9–1.0

6.51% cases have Z values lower than 0.7.

6 Conclusion and Further Scope

The apartment layouts although belonging to the same typology may have similar specifications in terms of the finishes and inventory. However, the above study clearly shows the clear demarcation in the utility value of the apartment layout belonging to the similar typology. This thus establishes that the value of the commodity cannot remain isolated from its use value. The finding of this research encourages integrating the utility score in the pricing model of the apartments.

The further scope to this research includes (but not limit to) deploying machine learning (ML) component for predicting the usability of apartments. The present study was restricted to fuzzy inference system due to limited dataset. In addition to this, the proposed model may have a reflection of various user groups while predicting the utility score of the apartment.

Acknowledgements We are thankful to Ar. Deepak Bhatia, Architect and Founder, Good Life Homes, New Delhi, for sharing valuable inputs to consider while calculating the utility score. Also, we acknowledge the support given by Muskan, Faria, Avantika, and Nandini for helping us conduct the assessment of layouts. The authors are thankful to all the experts who participated in the successful conduct of the area.

References

1. Sinha KK, Mathur M, Sharma A (2022) Assessment of bedroom utility: fuzzy logic based approach. In: 4th International conference proceeding on artificial intelligence & speech technology (AIST). IEEE, Delhi, pp 530–535
2. Broome J (2000) Utility. Econ Philos 7(01):1–12
3. Kapteyn A (1985) Utility and economics. Economist (Leiden) 133(01):1–20
4. Kaklauskas A, Zavadskas EK, Banaitis A, Satkauskas G (2007) Defining the utility and market value of a real estate: a multiple criteria approach. Int J Strateg Prop Manag 11(2):107–120
5. Phipps AG (1983) Utility function switching during residential search. Geogr Ann. Ser B, Hum Geogr 65(1):23–38
6. Lancaster KJ (1966) Change and innovation in the technology of consumption. Am Econ Rev 56:14–23

7. Çekmiş A (2016) Fuzzy logic in architectural site planning design. Procedia Comput Sci 102:176–182
8. Naz A, Nadin M (2018) Application of fuzzy logic in design of an aesthetics-based interactive architectural space. Int J Appl Res Inf Technol Comput 9(2):113
9. Ciftcioglu Ö, Durmisevic S (2001) Fuzzy logic in architectural design. In: EUSFLAT conference, pp 265–268
10. Valiyev A, Imamguluyev R, Ilkin G (2021) Application of fuzzy logic model for daylight evaluation in computer aided interior design areas. Adv Intell Syst Comput, Springer International Publishing, pp 715–722

Impact of Issued Loans on the Income of Self-help Group Women in Rajasthan

Pooja, Priyanka Harjule, Basant Agarwal, and Sandesh Nayak

Abstract This paper focuses on a study conducted to examine the impact of loans issued to self-help group (SHG) members on their income in Rajasthan. It is essential to understand the role of loans in income generation, particularly for women who are often excluded from formal financial services. The data is collected across the entire state and has been analyzed using descriptive and inferential statistics. It presents the findings of the study, including an analysis of the relationship between issued loans and income and the factors that influence this relationship (years of joining). It was found that 73.14% of SHG members' income increased after joining SHG. The study found that 10.46% of the SHG families earn more than or equal to 100,000 and 3.27% of the SHG families earn more than or equal to 150,000 annually in order to enhance NRLM policy. The study also highlights the importance of encouraging SHG members to use loans to finance income-generating activities to increase their earning potential.

Keywords Self-help group (SHG) · Income-generating activities · Financial inclusion

Pooja (✉) · P. Harjule
Department of Mathematics, Malaviya National Institute of Technology, JLN Marg, Jaipur, Rajasthan 302017, India
e-mail: poojayadav234422@gmail.com

P. Harjule
e-mail: priyanka.maths@mnit.ac.in

B. Agarwal
Department of Computer Science and Engineering, Central University of Rajasthan, Kishangarh, Ajmer, Rajasthan 305817, India

S. Nayak
Department of Mines and Geology, Government of Rajasthan, Jaipur, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_46

1 Introduction

Self-help groups (SHGs) have emerged as an effective tool for poverty reduction and women's empowerment in developing countries. These groups give a platform for the poor to come together and access financial services such as microfinance, savings and credit, which were previously unavailable to them. They are typically formed by women from economically poor backgrounds who come together to pool their savings and provide loans to one another at affordable interest rates. One of the critical objectives of SHG is to improve the economic status of its members, particularly in terms of income generation and financial inclusion [1]. In this context, loans issued by SHG play a crucial role in enabling members to undertake income-generating activities. The impact of issued loans on the income of SHG members has become a considerable subject of research. This article aims to contribute to this research by examining the impact of issued loans on the income of SHG members of Rajasthan and whether the capacity of rural women in Rajasthan for business growth increases after joining an SHG? Rajasthan has a high number of SHGs. These groups are primarily concentrated in rural areas, where access to formal banking services is limited. The Rajasthan government has also launched several initiatives to promote the formation and growth of SHG, alleviate poverty and empower women. One of the key components of the SHG model is the provision of loans to members. These loans are typically small in size and are used for income-generating activities such as farming, livestock rearing and small-scale businesses. The repayment of these loans is usually made weekly or monthly, with an affordable interest of 7% per annum [2]. The impact of these loans on the income of SHG members in Rajasthan is an important research question, given the significance of SHG in the state's development agenda.

Not many studies exist in the literature on the income-loan gap analysis of SHG members in Rajasthan. But there were various case studies for different districts in Rajasthan linked to women's upliftment in SHG [3, 4]. In addition, some research was conducted throughout the state to study the influence of government programmes such as SAKHI on the economic progress of SHG women [5]. Some studies also highlighted other variables contributing to the income loans gap, such as a lack of education, insufficient infrastructure and limited credit access [6]. There were also other studies that looked at the impact of various socioeconomic constraints on the income of an SHG member [7]. The main objectives of this study are as follows:

1. To study the "Income-Loan Gap" analysis among the SHG members across Rajasthan.
2. Impact analysis of the provided loans on the economic growth of SHG members.
3. Performance analysis of the loans and purpose of loans on the income of SHG members.
4. Significance of NRLM concerning the total loan amount of an SHG member.

This paper is divided into five sections. Section 1 gives this study's introduction, background, motivation and objectives. Section 2 provides the methodology, data collection process and sampling. Section 3 gives a detailed analysis of the main findings, followed by discussions. A conclusion is given in Sect. 5. Finally, Sect. 6 provides the acknowledgement.

2 Methodology

2.1 Data Collection and Description

A survey tool was developed in order to collect the data. A mobile application named "Kobocollect" was designed for the survey tool. This application was installed on the mobile phones of 30 community-based monitors (CBMs) women who conducted the door-to-door survey across 33 districts in Rajasthan using their mobile phones. The snowball sampling technique is used to finalize the sample. The sample was selected from different regions, ensuring uniform representation from both urban and rural areas. The sample size was determined based on the assumption that it must collect information about those SHGs that are three or more years old and have a membership range of ten or more members. Additionally, it is also assumed throughout the report that the term "annual income" refers to the annual income of the SHG member (not of her family). Table 1 presents the finalized sample of this study.

3 Results

3.1 Years of Joining Versus Number of Issued Loans

It is essential to investigate a relationship between the number of provided loans and the year of joining SHG. The data regarding the number of loans issued in relation to the years of membership are shown in Fig. 1. Figure 1's dot denotes the existence of

Table 1 Desired sample size for the survey

Sample size	Calculation
Two blocks (Oldest) from each District of Rajasthan	$33 \times 2 = 66$ Blocks
Three villages (Random) from each selected block	$66 \times 3 = 198$ villages
Five SHG (Oldest) from each selected village	$198 \times 5 = 990$ SHG
Five SHG members (Random) from each selected SHG	$990 \times 5 = 4950$ SHG members

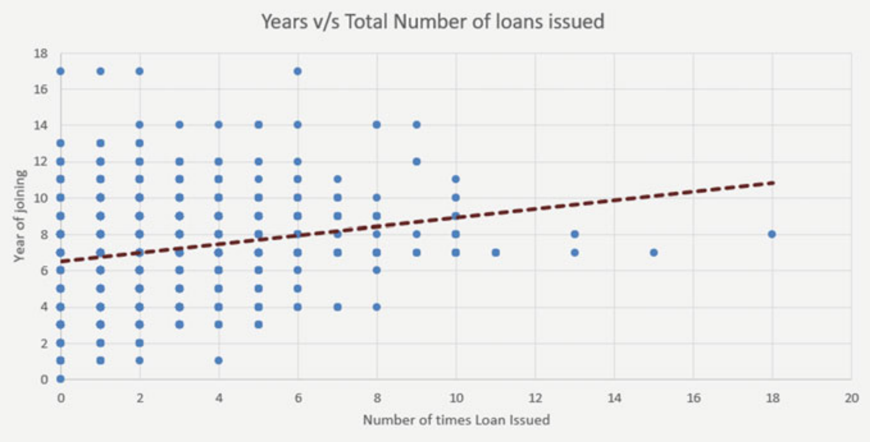


Fig. 1 Years of joining versus number of issued loans

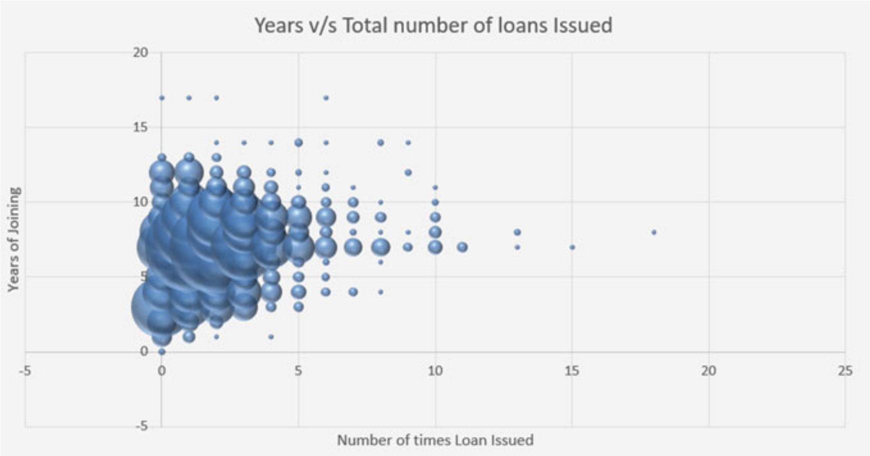


Fig. 2 Years of joining versus number of issued loans

at least one SHG member who corresponds to the particular coordinate of the graph of Fig. 1. The trend of the graph is shown by the dark red line, which is linear and slightly proportional to the years of joining of a SHG member. Moreover, 83.56% of SHG members have issued at least one loan.

The data of Figs. 1 and 2 are identical. However, the dot in Fig. 2 has different sizes, indicating that as the size of the dot increases, more SHG members are available to correlate to that dot. Therefore, it can be inferred from Fig. 2 that the majority of SHG members have a membership duration between 5 and 10 years and have taken loans 2–5 times.

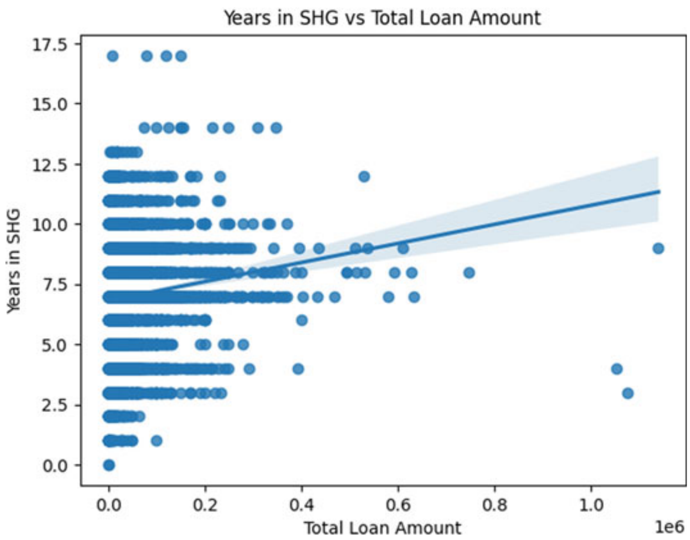


Fig. 3 Years of joining versus total amount of issued loan

3.2 Years of Joining SHG Versus Total Amount of Issued Loans

The amount of loans issued by an SHG member can be influenced by several factors, including the years of membership. A significant area of investigation is to find the relationship between the number of years spent in an SHG and the total amount of loans granted. Figure 3 displays a linear relationship between both variables. Additionally, it is also apparent from the graph that both variables are directly proportional to one another.

According to National Rural Livelihoods Mission (NRLM) [8], if a candidate receives a total loan amount of 1 lakh to 1.5 lakhs over a ten-year period, then he or she is presumed to be out of poverty. So, corresponding to it, for SHG members who took total loans of more than or equal to 1 lakh, Fig. 4 depicts the relationship between years of membership and the total amount of provided loans. The graph's horizontal line, which depicts the trend of the data, reveals that there is no correlation between the two variables. But from further analysis, it can be concluded that 19.91% of members have taken loans of more than or equal to 100,000 with an average of 8 years of joining.

Caste has been a deeply inherited social hierarchy in India, with certain castes traditionally occupying higher social and economic positions than others. Studies have shown that caste discrimination can affect access to credit in India. Therefore, caste could impact the total amount of issued loans and years of joining. Figure 5 provides the linear regression model between the years of joining and the total amount of issued loans, highlighting the impact of caste on both variables. Figure 5 shows

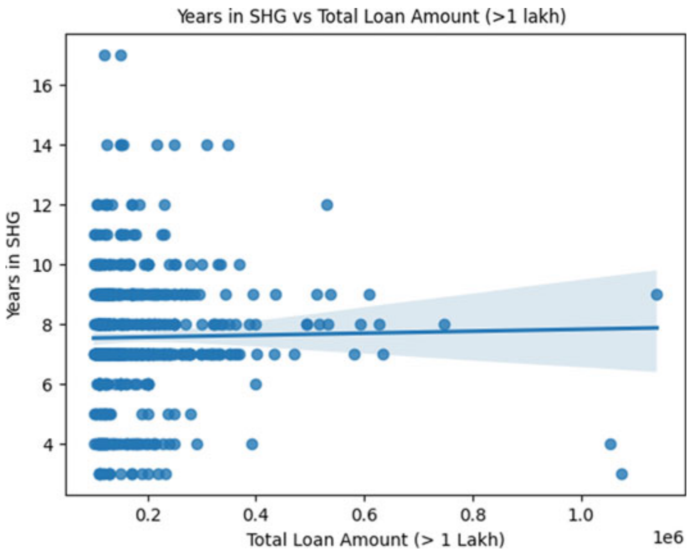


Fig. 4 Years of joining versus total amount of issued loan (> 1 lakh)

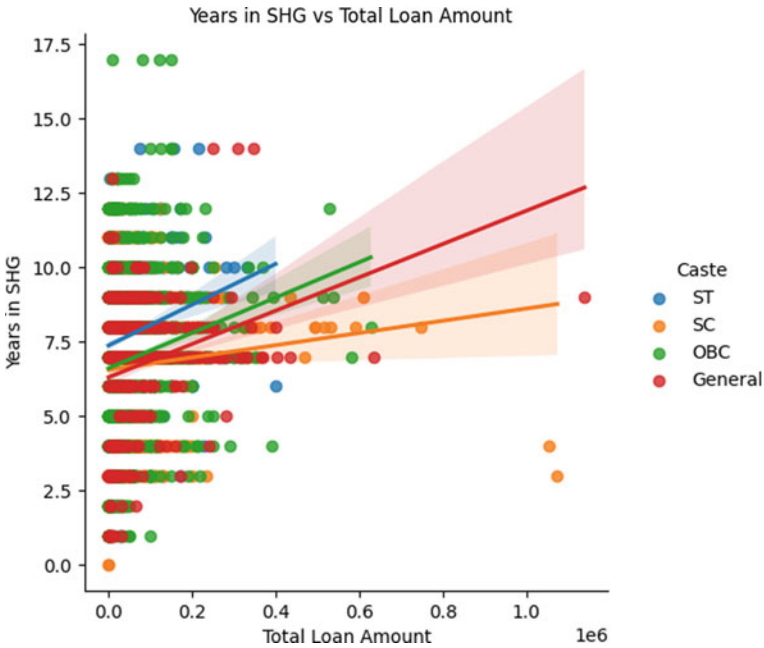


Fig. 5 Impact of caste on years of joining versus total amount of loan issued

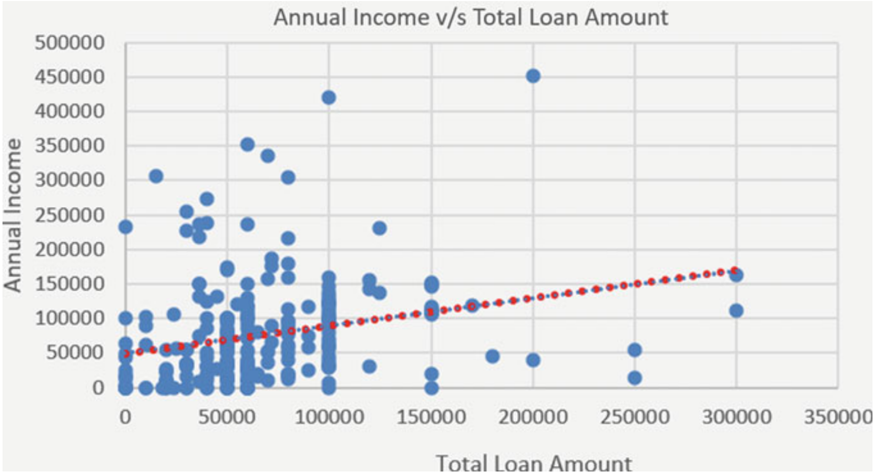


Fig. 6 Annual income versus loan amount

that the relationship is linear for members of each caste. Still, the total amount of loans is directly associated with the years of joining for ST candidates as compared to OBC and general candidates.

3.3 Annual Income Versus Loan Amount

Figure 6 illustrates the relationship between an SHG member's yearly income and the total loan amount granted. According to the data, 96.73% of SHG members earn less than 150,000 annually, and 80.09% of them have also issued loans totalling less than 100,000.

The graph's trend demonstrates that an SHG member's annual income is approximately proportional to the sum of loans; that is, an SHG member's annual pay rises in parallel with the number of loans they get. From further analysis, 81.48% of the loans are invested in financial inclusion activities, including farm, livestock and non-farm economic activities, and the annual income of 73.14% of the SHG members has increased.

3.4 Comparison in Annual Income of SHG Members Depending upon the Years of Joining

It is important to determine if an SHG member's years of membership impact their annual revenue from the point of view of NRLM. Based on how long they have been members, the SHG membership is divided into three groups: less than five years, between five and nine years and more than nine years. Two sample t-test with a significance level of 0.05 is used to determine how the following three categories impact the SHG's annual income. So, the null hypothesis and alternate hypothesis for the first group are as follows.

H_0 : *There is no significant difference between the annual income of those SHG members who have less than five years of joining and those who have membership between five to nine years.*

H_1 : *There is a significant difference between the annual income of those SHG members who have less than five years of joining and those who have membership between five to nine years.*

There is a significant difference (p-value = $4.58 \times 10^{-2} < 0.05$) in the annual income of the SHG members who have less than five years of membership and those who have membership between five to nine years. Additionally, the average annual income for the SHG members who have less than five years of membership is 47,000, and the average annual income for the SHG members who have membership between five to nine years is 60,000. A rising trend in the annual income of the SHG members is discovered here.

Now for the second group, the null hypothesis and alternate hypothesis are as follows:

H_0 : *There is no significant difference between the annual income of those SHG members who have membership between five to nine years and those who are more than nine years.*

H_1 : *There is a significant difference between the annual income of those SHG members who have membership between five to nine years and those who are more than nine years.*

There is a significant difference ($p\text{-}value = 9.67 \times 10^{-4} < 0.05$) in the annual income of the SHG members who have membership between five to nine years and who are older than nine years. Moreover, the average yearly income for the SHG members who have been members for five to nine years is 60,000, and the average yearly income for the SHG members who have been members for more than nine years is 45,000.

3.5 Comparison in Total Amount of Issued Loan of SHG Members Depending upon the Years of Joining

The influence of the total amount of issued loans for three categories defined in Sect. 3.4 is evaluated using the Mann–Whitney U-test with a significance threshold of 0.05. So, for the first group, the null hypothesis and alternative hypothesis are stated as:

H_0: *There is no significant difference between the total amount of issued loans of those SHG members who have less than five years of joining and those who have* *membership between five to nine years.* H_1: *There is a significant difference between the total amount of issued loans of those SHG members who have been members for less than five years and those with membership between five to nine years.*

There is a significant difference between (p-value = $2.0978 \times 10^{-54} < 0.05$) the total amount of loan for SHG members who have been members for between five to nine years and those who have been members for less than five years. Additionally, the average issued loan amount for SHG members with less than five years of membership is 66,000, while the average for SHG members with five to nine years of membership is 77,000. An increasing pattern is observed here in the total amount of issued loans. The null hypothesis and alternate hypothesis for the second group are as follows:

H_0 : *There is no significant difference between the total amount of issued loans of those SHG members who have membership between five to nine years and those who are more than nine years.*

H_1 : *There is a significant difference between the total amount of issued loans of those SHG members who have membership between five to nine years and those who are more than nine years.*

There is a significant difference (p-value = $2.429 \times 10^{-4} < 0.05$) in the total amount of issued loans of the SHG members who have membership between five to nine years and who are older than nine years. Additionally, the average issued loan amount for SHG members who have been a part of the group for five to nine years is 77,000, and for those who have had a membership of more than nine years is 65,000.

4 Discussions

The study's findings indicated that a member's yearly income and the total amount of loans provided to that member were significantly influenced by how long they had been a part of the SHG. In accordance with the NRLM concept, 19.91% of the SHG members are taken to be free from poverty since they have given out loans totalling at least 100,000. Moreover, the unique finding of the survey is the data itself. This kind of data was not previously available for the state of Rajasthan. The findings of this study discussed that providing loans to SHG members can significantly improve their means of earning and sustaining themselves. By offering financial support, the SHG may encourage their members to engage in income-generating activities that were previously beyond of their reach. The results imply that SHG can be a valuable instrument for reducing poverty and raising women's economic standing in rural regions.

5 Conclusion

Overall, the study emphasizes the positive impact of loans issued to SHG members on their income levels in Rajasthan. This increase in annual income is mainly observed due to the investment of loans in various income-generating activities, including handicrafts and small businesses. Inputs, including raw materials and tools, were generally purchased with loans in order to increase the production and profitability of their businesses. The study's findings encourage the SHG members to make investments which are aiming to promote the financial inclusion and empowerment of women through SHG in Rajasthan.

The future work will include a detailed analysis of the sources of loans and income of SHG members. Further, we will relate these findings to support the policy-making at Rural Development Department.

Acknowledgements The present study received funding from the Government of Rajasthan under research project number 1000114252, entitled "Study impact on members of Self Help Groups (SHG) in Rajasthan through survey and MIS data with applications of data analytics techniques".

References

1. Galab S, Chandrasekhara Rao N (2003) Women's self-help groups, poverty alleviation and empowerment. Econ Polit Wkly 38:1274–1283
2. Women SHGs development fund—operational guidelines, Department of Financial Services, Ministry of Finance, Government of India
3. Agrawal HG (2018) Empowerment of women through self help group—a case study of Baswara District (Rajasthan). Int J Res Innov Soc Sci 2:2454-6186
4. Bhatnagar S, Rathore SS (2015) Social and economical empowerment of rural women through self help group formation in Jaipur District of Rajasthan. J Krishi Vigyan
5. Choudhary AS (2015) Economic empowerment of rural women entrepreneurs in Rajasthan through self-help group: a case of SAKHI. Adv Econ Bus Manag 2:2394-1545
6. Kachhawa D, Dhangi KL (2012) Empowerment of tribal women through their mobilisation in planning and execution of self-help group activities of Udaipur district of Rajasthan State. J Progressive Agric 2 0976-4615
7. Ritu R, Ravinder M, Shubham M (2021) Analysis of constraints faced by the members of women dairy self-help group in Rajasthan. Indian J Extension Educ
8. National Rural Livelihood Missions Government of India. https://trlm.tripura.gov.in/sites/default/files/FrameworkofimplementationofNRLM.pdf

Internet of Things: A Possible Solutions for Monitoring Quality of Animal Feed During the Supply Chain

Alpa R. Barad and Ankit R. Bhavsar

Abstract The Indian Animal feed market reached a value of INR 873.3 Billion in 2021. Looking forward IMARC Group expects the market to reach INR 1493.8 billion by 2027, exhibiting at a CAGR of 9.6% during 2022–2027 (IMARC impactful insights in Indian Animal Feed Market: Industry Trends, share, size, growth, opportunity and forecast 2022–2027. https://www.imarcgroup.com/indian-animal-feed-market). Currently India is one of the largest animal feed producers in the world. Animal feed includes raw, processed, and semi-processed products that are fed to livestock. Some of the most common feed includes green grasses, cereal grains, and silage crops. Animal feed is a most important part of the agriculture sector (Food and Agriculture organization of the United Nations in Environmental performance of animal feeds supply chain 2017). As the global population increases day by day, the demand for animal food is gradually increasing. At the same time, the feed industry faces several challenges initiated by changing environment conditions, increasing demand and high commodity prices (Ravindra in IoT Applications in Agriculture 2020. https://www.iotforall.com/iot-applications-in-agriculture). The Internet of Things (IoT) and image processing have a big contribution in smart farming. IoT is a system which is built for monitoring the crop field with the help of sensors and sending this data to farmers who connected with it. While farmers can access this data from anywhere. Image processing is a method to access images of crops and transforming this image into digital image and extract valuable output from the image. In this paper, problems faced in supply chain management and quality of green grass for maintaining the animal feed sector. After examining the problem we proposed a possible solution for the particular issues in animal feed. The solution uses the latest technology like image processing and IoT for animal feed. It shows how these technologies are integrated into the problem and find the solution for the particular issues as well as maintain the health of animals.

A. R. Barad (✉) · A. R. Bhavsar
GLS University, Ahmedabad, India
e-mail: alpabarad93@gmail.com

A. R. Bhavsar
e-mail: ankit@glsica.org

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_47

Keywords Agriculture · Animal feed · Image processing · IoT · Supply chain management

1 Introduction

In India, agriculture is growing progressively, especially the livestock sector is a very valuable subset of agriculture in the Indian economy. Over the past few decades, individual lifestyle and necessity have gone through immense changes. So, smart farming is very important for handling the challenges of agricultural production in terms of fertility, ecological impact, food security, and sustainability [4]. As populations are increasing day by day, the demand for animal products is consistently increasing; therefore, farmers should take care of animal feed products. Furthermore, we get food from plants and animals so it is our responsibility to take care of these resources because our future is dependent on these resources and in consequences the biggest challenges of an era. Hence to cope up with these demands, farmers and researchers should work on farming management with sustaining the time availability and high nutrition quality across the globe.

Today animal feed products have numerous varieties to eat like silages, hays, sprouted grains and legumes, oils and mixed rations, etc. To better survive animals, the feed is the most important part of animal feed globally. The foremost necessary factor of animal feed production is feed, but nowadays some farmers or growers are yet to believe and accept their alternatives by products [5]. In addition due to their major use of other materials the nutrients values are changed simultaneously. So presently farmers and growers should have to take care and focus on quality of food and supply chain management which are mandatory factors of smart farming.

To get nutrients from the animal, it should have feed which includes water, carbohydrates, protein, minerals, and vitamins [5]. To get enough nutrients and better quality products of animals, there are few things which are non-negligible "ANIMAL FEED". Furthermore to get the best quality of animal products, it is necessary to feed them best. For this, stakeholders are regularly supplying different types of green grasses like elephant grass, green maize, sorghum, lasoon grass, green bajri, and Napier grass to the consumer.

In the process of supplying animal feed to the particular consumer, there is a particular chain followed from farmer to stakeholder to consumer. If the quality of green grass is not maintained, then the green grass is rejected by consumers and comes back to the farmer. In case of green grass rejection, there is huge loss of stakeholder as well as farmers as well as the environment. So to reduce the loss of time as well as economy, there farmers and stakeholder should maintain the necessary quality as well as follow the farming management system.

2 Supply Chain Management in Smart Farming

Supply chain means flow and movement of goods from the producers to final customers [6]. Supply chain is a chronological order that focuses on meeting final customer requirements. It not only involves suppliers and producers but also transporters, warehouses, retailers, and consumers. Supply chain management first appeared in 1980s as an inventory management approach. Supply chain management is the integrated planning, implementation, coordination, and control of all agri business processes that are necessary to manufacture and deliver, as skillfully as possible, products that satisfy customer fondness and demand [6].

Supply chain management plays a vital role in the animal feed industry as it has numerous operations that are used to simplify a chronological supply of the product to consumers [7]. SCM has been implemented to a point where middle authorities have been taken into consideration. This has been done through ensuring that the mitigation of risks is in place. The organization has implemented systems with steps that involve middle management, and senior management is in place regarding approval before claiming. There is plenty of inside authority in place as the supply chain plays a critical role [8]. There are a number of manufactures that have their own storage facility on their place to cope up with demand of consumers because animal feed has a low shelf life so producers cannot store for a long period of time. when any one link in the supply chain management of animal feed is not working optimally. For e.g., farmers or stakeholders are not working properly; then, the whole supply chain is difficult to survive [8]. Due to various issues in the agriculture sector in India, supply chains of various products have difficulties to transfer.

2.1 Working of Supply Chain Management in Animal Feed

Supply chain management (SCM) refers to the suppliers who try hard to design and implement supply chains that are well organized and cost effective as possible. Supply chain involves everything from manufacturing to final product, and then, the information system of product to manage is very difficult. Every product entered in the market is the result of the effort of a number of organizations that consist of supply chain. Although the supply chain has existed for the past many years, many of them recognized their value-added business [6].

Animal feed products such as green grasses are collected at various locations by various sources. First person, the farmer who is manufacturer and takes care of green grass so he sells the products to stakeholders. Secondly the stakeholder, who is in contact with the farmers via other sources and does some formalities like agreement and contracts so the stakeholder sells the products to customers. Thirdly, the consumer who can access the final product. As a result, supply chain management is the management of upstream and downstream, with value added at each stage (Fig. 1).

Fig. 1 Working of supply chain management in animal feed

3 Issues on Existing Supply Chain Management and Quality of Agriculture Product

To discover the real structure of existing problems in animal feed and quality of animal feed first to identify a number of issues in the existing structure. To supply animal feed to the supplier at that time, it is necessary to maintain the quality of green grass, but eventually every situation has its problem and here it is.

Firstly, every season has its own pros and cons so when supplier supplies green grass in summer then he has to be more careful about their quality because in summer the temperature is high so if he cut the green grass so early then the quality is not maintained and fug is placed in green grass as well as if cut the green grass on time although the quality is not maintained as same as winter. If we cut the grass in a huge amount, then it will be more wastage so better that harvest feed in moderate amounts so the grass will be rejected less at the consumer side.

Through the summer season green grass needs more water so farmers have to plan before summer to cope up with this situation which benefits all of them like farmers, suppliers, and consumers. If green grass is rejected by consumers, then it will affect the transportation, hard work of labors as well as the animal feed wastage.

There is a lack of a farming management system in India according to seasonal bases. There is a lack of supply chain management system (SCM) in between farmers and stakeholders as well as stakeholders and consumers. Manpower cost is high due to needing more physical work as well as due time work. When the rejected quantity is higher than the accepted quantity, then there is a huge financial loss. Lack of knowledge of Pre harvesting, Harvesting and Post Harvesting.

4 The Use of Technologies for Smart Farming

Regardless of the sum up population, which is predicted to reach 9.6 billion in 2050, the agriculture sector has to fulfill the demand, despite climate changes like unfavorable weather conditions. To meet the requirement of rising population, the agriculture sector will have to adopt new technologies to increase the much needed edge. Latest agriculture applications in smart farming through new technologies will enable the industry to gain operational efficiency, lower costs, reduce waste, and improve the quality of their yields.

4.1 Image Processing

Image processing is significantly used in the agriculture sector as well. The main benefits of using image processing are that it is non-harmful; it means that it can provide insight information about crops without coming into contact. There are four major applications of image processing in the agriculture sector. The first one is crop management, especially with respect to pest and disease detection and irrigation methods. The second one is likely to evaluate crop leaves and skin conditions, in this the identification of nutrition deficiency and plant content. As well as in thirdly, fruit quality inspections and sorting are possible. Just as importantly, in crop and land estimation, image processing is utilized for the Geographic Information Center and color and texture segmentation of crops [9].

4.2 Internet of Things (IoT)

In IoT-based smart farming, system is made up for monitoring the crop field with the help of sensors (light, humidity, temperature, and soil quality) and automating the irrigation system. Farmers can keep an eye on field conditions anywhere in the world. IoT-based smart farming is well structured compared to traditional approaches. In terms of environmental problems, IoT smart farming can supply significant advantages including more water usage or developments of inputs and hospitality [10].

IoT-based smart farming has numerous applications in the agriculture sector like precision farming, which is the most important application in the IoT-based agriculture sector. These techniques first come to mind when farming practice is more controlled and precise when it comes to growing livestock and crops. Secondly agricultural drones, which are used in the farming sector in order to modify the number of agriculture practices [11]. The main benefits of agricultural drones are crop health imaging, integrated GIS monitoring, ease of use, save time, and potential to raise yields. Thirdly livestock monitoring, which helps owners to identify the location, health of their cattle and well-being.

4.3 Machine Learning

Machine learning in agriculture gives real time to deliver straight gains to farmers. Agriculture is the main and basic need of any country. To meet the demand of a growing population, agriculture should be in the form of utilizing trendy technologies like machine learning. Machine learning can help to expand capability and reliability in decision making while at the same time reduce the probability and costs associated with agriculture operations [12].

By evaluating real-time sensor data and historical trends, machine learning can motivate farming decisions [13]. Machine learning can be used in a number of features of agriculture management, from assessing climate information to anticipating the incident of disease. Machine learning has various applications in agriculture industry like crop management, precision spraying, insect detection, field conditions management, yield mapping, livestock management, price forecasting for crops, automatic weeding, and automatic harvesting robots.

5 Related Work

These technologies have been used in various applications related to monitoring of animal feed as well as tracking of the agriculture sector. Few number of applications [4, 5, 7, 9, 12–18] are mentioned below.

V. Bloch et al. have described how image processing can be used to create the platform for measuring the cow feed efficiency for commercial dairies [14]. Furthermore the cost and maintenance time is important for a research system; therefore, design a system with convenience and low investment. So, a new system design consists of three important parts: (1) A hanging weighing system (2) A visual cow identification system (3) An automatic cleaning system.

Bin hu et al. proposed an algorithm for animal feed and animal waste detection based on computer vision using image processing [4]. The algorithm uses color and canny's edge features. Furthermore, the morphological process is used for identifying the animal feed detection area. For the animal waste detection, the polluted area is measured by using a median filter together with Hough's straight line transformation.

David Raba et al. have created IoT solutions to optimize the livestock feed supply chain [15]. The animal feed supply chain farm of framers and livestock farmers currently faces problems in outdated supply chain management. Stakeholders currently face issues in restocking their feeds as well as directly affecting cost and labor efficiency. There are some kind of solution provided for feedstock in bin for predicting stock level (i) Using (weight) load cells which are installed in the bin structure (ii) used level sensor which is based on radar, ultrasonic, and guided wave technology. The sensor RGB D which along with the camera is a new technology, and it enables color images as well as depth level pixels in real time. There is a cleaning system that is also included for clean reading, removing sticked feed as well as dust removal on the head of the bin near sensors. although sensor installation is not done for more than 20 min.

Luis Nobrega et al. proposed animal monitoring behavior with the use of IoT local network and cloud platform [13]. The SheepIT project proposed for controlling, monitoring, and managing IoT solutions for sheep herds. Machine learning is used for taking the best decision of sheep which is in the vineyards or not so it works as a supervised algorithm.

Lefteris Benos et al. studied that machine learning can work on crop management, water management, soil management, and livestock management in agriculture [16].

Here, along with machine learning, artificial neural networks are also used. Crop quality consisted mainly of RGB images, while X-ray images were also utilized (for seed germination monitoring). Additionally, quality parameters, such as color, mass, and flesh firmness, were used.

Nurashikin Ahmad Fuad et al. proposed that bananas are classified according to size, shape, color, texture, and taste [17]. So here ripeness of bananas is done using image processing with the help of MATLAB. Edge detection extracts edges from an image so For edge detection, canny's edge detection method is used for good results.

Vishal Meshram et al. mentioned that agriculture is very important for the economy as well as employment of developing countries like India [12]. So in this paper they described three different major areas like pre harvesting, harvesting, and post harvesting. Machine learning is helpful for farmers to minimize the losses as well as allow more efficient and precise farming with less manpower with high quality production.

Illan Halachmi et al. stated current trends and precision LiveStock Farming (PLF) using IoT and Data Science Tools [18]. PLF is a real-time monitoring system so the cow feed efficiency is measured by artificial intelligence and 3D camera. Cow feed efficiency is measured for future use as well with the help of mathematical models, camera, and sensors. While cow body temperature is measured by sensors to identify level of heat stress. Cooling system is also activated when needed for better efficiency.

6 System Flowchart Diagram

In this paper proposed an architecture for animal feed monitoring system in supply chain management which uses IoT and image processing technologies for solutions (Fig. 2).

First of all farmers harvest the green grass early in the morning around 3 to 4 a.m. and below 20 * c temperature because farmers have to serve fresh green grass to the stakeholder. Farmers need more labor for harvesting so it is a time-consuming process. It is the primary responsibility of farmers to cut grasses in predefined time for this adequate manpower to be deployed. Farmers are also advised to not to use pesticides in grasses as it can directly affect the health of animals to whom these grasses will be served. It is observed that farmers are used to adding pesticides to yield high production of grasses by using harmful pesticides which would directly affect the health of animals. The poisoner's chemicals present in pesticides may cause harmful disease or can even kill the animals. It is the responsibility of farmers to ensure that all the grass reaches warehouses on time. In order to sort the green grass, where manually sorting is done in which dead leaves and waste material from grass are removed and best quality as per The norm is to proceed further where grasses are tied in bundles. Before loading grasses on the vehicle, all vehicles are properly washed by water and unwanted mud and dust is removed. Post cleaning vehicle is ready to load different types of grasses on it.

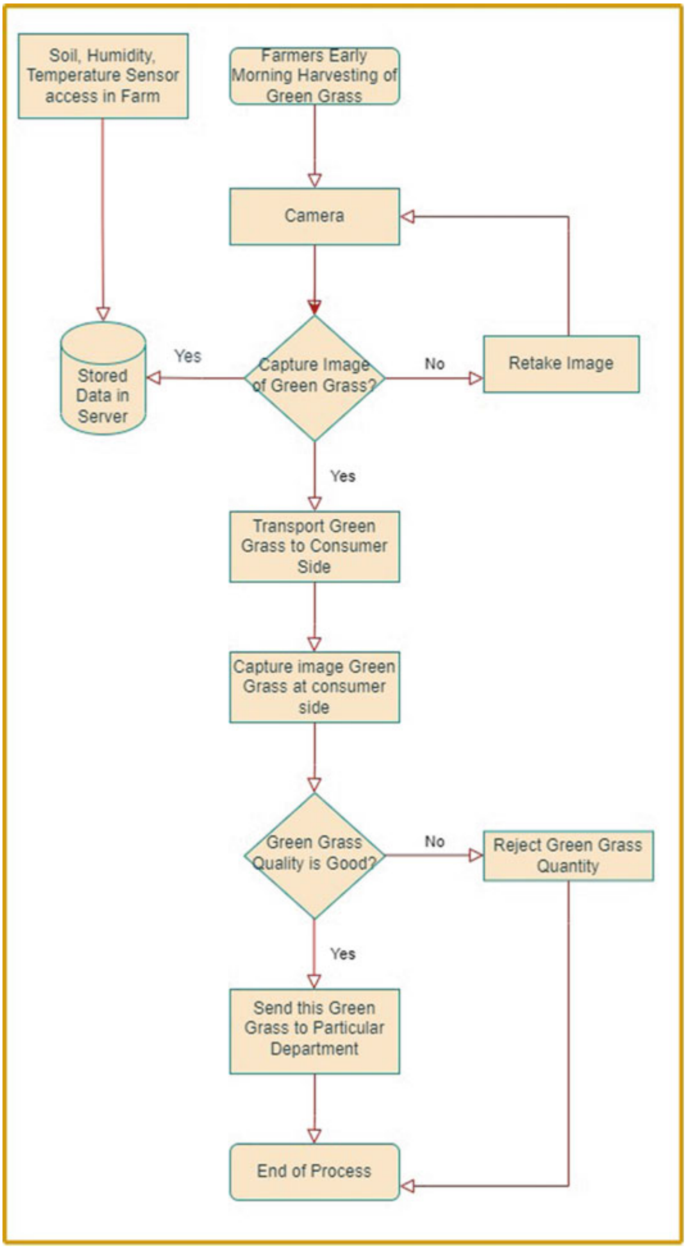


Fig. 2 System flowchart diagram of animal feed using supply chain

Using this architecture expect to monitor the quality of animal feed at the farmer side before loading green grass. To get precise details of the quality of animal feed, we need to evaluate the temperature, humidity, and soil moisture data using environment sensors and soil moisture sensors. With the help of camera take the picture of green grass and using image processing techniques evaluate the quality of green grass. To capture the location of green grass along with calculation of quantity of Green with respect to transportation. Lastly, there are tracking sensors which are used to track the transportation of green grass, which means where to go and when to go with an exact idea. At the consumer side when animal feed is delivered between 8:30 and 9 a.m. then first of all green grass weight is calculated on weighbridge. Once a vehicle reaches the desired site/factory, then it is supervised by a quality check supervisor who will ensure that adequate quantity is distributed among all areas, and also, quality is as per requirement. To illustrate, if a vehicle is to deliver elephant grass in an elephant area post observation of a quality check supervisor, the same grass is distributed at different cells as per the need of the elephant.

On the other side also take necessary steps to monitor the system. Due to high temperatures in summer, crops need more water so have to maintain and settle the irrigation system. In summer quality is maintained appropriately so sorting is done correctly and perfectly then winter. Before the supply of animal feed, it is necessary to coordinate between three of them so time as well as quality is maintained. Nowadays manpower costs are high so take necessary steps to reduce manpower work and take the idea of taking work from labor in a short period of time because this is a time consuming process.

Before sowing the crops, farmers should have knowledge of pre harvesting, harvesting and post harvesting so the quality of crops is maintained. Take appropriate steps of the rejected quantity of animal feed so wastage of material is no more waste. Use natural fertilizer instead of pesticides in land so crops will be more organic as well as quality will be good and long lasting. With the help of technical resources and techniques, the quality, temperature, and humidity will be measured. Resorting architecture will not waste time on dispatch, loading, and unloading. Individual green grasses have differently so get information before sowing of each green grass because different crops have different harvesting time.

After harvesting green grass, take the image of green grass leaves and check the necessary quality of leaves using image processing techniques. Moreover, using environment sensors, we get the information about temperature and humidity with respect to weather as well as accessing soil moisture sensors to know the data about how much quantity of moisture stay in soil. Furthermore, load the animal feed and store it at stakeholder warehouse. After completing weighing of grass and making a bundle of grass taken for loading to the factory. Here using the IoT system, track the loading vehicle from warehouse to factory. At the consumer side/factory side, they fetch the data regarding quality of green grass using IoT technology. Moreover, the rejection chances of green grass are less compared to standard technology, and the quality of animal feed is best compared to prior technology which helps to maintain the mandatory supply chain and quality of animal feed which directly benefits the three of them as well as environment. It is also expected that all the members involved in this procedure are fully professional because if any of the members is incompetent at their job then it can directly affect the rate at which animal feed is supplied. Farmers are also expected to hire labors who can do more dedicated and physical work early in the morning. For this a competent staff is employed by stakeholders as well as consumers.

7 Conclusion

In a nutshell, the proposed system will try to help to maintain the quality of animal feed along with maintaining supply chain management using image processing and Internet of things. By using an image processing system, we identify the quality of animal feed whether the grass is fresh or not as well as with help of environment sensors like temperature and humidity sensors get to know that atmosphere have how much temperature and humidity as well as soil moisture data using soil moisture sensors. On the other hand, using an IoT system, they get to know the data about real-time quality of green grass. This overall process is done within 5 h in the morning. Thus, the process of harvesting is time consuming and requires physical hard work by laborers. This proposed system gives improvement in the agriculture sector more transparency. Using this system the health of animals will be good in terms of eating the best quality food and having all sufficient nutrients from the green grass. In short this system indirectly helps animal health as well as animal products which are advantageous to people's health who are using animal products.

References

1. IMARC impactful insights, Indian Animal Feed Market: Industry Trends, share, size, growth, opportunity and forecast 2022–2027. https://www.imarcgroup.com/indian-animal-feed-market
2. Food and Agriculture organization of the United Nations (2017) Environmental performance of animal feeds supply chain 2017
3. Ravindra S (2020) IoT applications in agriculture 2020. https://www.iotforall.com/iot-applications-in-agriculture
4. Bin Hu QT (2014) Research on animal feed and animal waste detection in computer vision. IEEE, pp 206–208. ISBN:978-1-4799-6058-3 https://doi.org/10.1109/soli.2014.6960721
5. Daniel Dooyum Uyeh YH, Animal feed formulation: rapid and non destructive measurement of components from waste by products. ELSVIER, vol 274, pp 1–17. https://doi.org/10.1016/j.anifeedsci.2021.114848
6. Arora H (2017) Doctoral Scholar in Economics & Senior Research Fellow, CDS, Jawaharlal Nehru University, Supply Chain Management in Indian Agriculture 2017. https://www.civilsdaily.com/supply-chain-management-in-indian-agriculture/
7. Gomera PM, Mafini C (2020) Supply chain management enablers, barriers and disruptions in the animal feed industry in the Western Cape Province of South Africa. J Transp Supply Chain Manag 14:1–12, ISSN-2310 8789. https://doi.org/10.4102/jtscm.v14i0.510
8. Supply Chain Management—Agriculture Notes 2002. https://prepp.in/news/e-492-supply-chain-management-agriculture-notes

9. Xu N (2021) Image processing technology in agriculture. In: The 2nd International conference on computing and data science (CONF-CDS 2021), Stanford, United States, vol 1881, pp 1–7. https://doi.org/10.1088/1742-6596/1881/3/032097
10. Cropin, Internet of Things in agriculture: what is an IoT and How is it implemented in agriculture? https://www.cropin.com/iot-in-agriculture
11. Dutta PK, Mitra S (2021) Application of agricultural drones and IoT to understand food supply chain during post Covid-19. Agric Inform, pp 67–87. https://doi.org/10.1002/9781119769231.ch4
12. Meshram V, Patil K, Meshram V, Handchate D (2021) Machine learning in agriculture domain: a state-of-art-survey. ELSVIER 1:1–11. https://doi.org/10.1016/j.ailsci.2021.100010
13. Nobrega L, Tavares A, Cardoso A, Gonclaves P (2018) Animal monitoring based on IoT technology. IEEE Xplore (2018), pp 1–5, ISBN-978-1-5386-6930-3. https://doi.org/10.1109/IOT-TUSCANY.2018.8373045
14. Bloch V, HI (2021) Design a system for individual cow feed intake in commercial dairies. Int J Ani Biosci 15(27):1–7. https://doi.org/10.1016/j.animal.2021.100277
15. Raba D, Gurt S, Vila O, Farres E (2020) An Internet of Things (IoT) solution to optimize the Livestock feed supply chain. Comput Sci Inf Technol (CS & IT), pp 103–118. https://doi.org/10.5121/csit.2020.100409
16. Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D, Bochtis D (2021) Machine learning in agriculture: a comprehensive updated review. MDPI 21(11):1–55. https://doi.org/10.3390/s21113758
17. Mustafa NB, Fuad NA, Ahmed SK (2008) Image processing of an agriculture produce: determination of size and ripeness of a banana. IEEE Xplore, pp 1–7, ISSN:2155-8973. https://doi.org/10.1109/itsim.2008.4631636
18. Halachmi I, Levit H, Bloch V (2019) Current trends and perspective of precision livestock farming (PLF) with relation to IoT and data science tools. Food and Fertilizer Technology Center (2019), pp 1–10
19. Wedpathak G, Pawar V, Kadam S, Thigale S (2019) IoT based farming with live monitoring of temperature, soil moisture, weather conditions through image processing techniques. Int J Recent Trends Eng Res, pp 36–44. https://doi.org/10.23883/ijrter.2019.5022.tmfnd
20. Prache S, Adamiec C, Astruc T (2022) Review: quality of animal source food. ELSVIER, 1–16

Test Automation Framework with Help of Synonym-Based Identification Algorithm

Prabhat Padhy, Pratima Nayak, and Siddhesh Vaidya

Abstract The aim of this research paper is to propose a test automation framework that uses a synonym-based locator identification algorithm to identify locators available on web page through scraping the data and representing it using Python flask framework (PFF). The proposed technique also aims to simplify and streamline the process of test automation for web applications, making it more efficient and effective by using suitable data structure. This research also depicts benefits such as reliability, finding similar components on the web page.

Keywords Synonym based locator identification algorithm · Scraping · Python flask framework (PFF) · Test automation

1 Introduction

Test automation is an essential part of software development, especially when it comes to web applications. Automated testing helps to identify defects early in the development process, reduce the time and cost of testing, and improve the overall quality of the software. However, creating and maintaining automated tests can be challenging, especially when it comes to identifying web elements on a web page. Traditionally, web element locators have been identified using attributes such as ID, name, class, or CSS selector. However, these locators can be brittle and unreliable, especially when web pages change frequently. A test automation framework is a set of guidelines and coding standards that are used to create automated tests. The purpose of a test automation framework is to provide a structure for creating automated tests

P. Padhy (✉) · P. Nayak · S. Vaidya
ApMoSys Technologies Pvt. Ltd., Navi Mumbai, India
e-mail: prabhat.padhy@apmosys.com

P. Nayak
e-mail: pratima.b.nayak@apmosys.com

S. Vaidya
e-mail: siddhesh.vaidya@apmosys.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_48

that are maintainable, scalable, and reusable. A test automation framework typically includes the following components:

1. **Test scripts**: These are the actual automated tests that are run by the framework.
2. **Test data**: This is the data that is used by the test scripts to simulate real-world scenarios.
3. **Test runners**: These are the tools that execute the test scripts and report the results.
4. **Reporting**: This is the mechanism used to report the results of the tests.
5. **Configuration management**: This is the mechanism used to manage the configuration of the framework and the test environment.
6. **Logging**: This is the mechanism used to log the results of the tests.
7. **Error handling**: This is the mechanism used to handle errors that occur during test execution.
8. **Test environment management**: This is the mechanism used to manage the environment in which the tests are run.

There are many different test automation frameworks available, including keyword-driven, data-driven, and hybrid frameworks. The choice of framework depends on the specific needs of the project, the team's expertise, and the tools available.

2 Review of Literature

The identification of web elements using synonym-based locators [1–3] is an important area of research in the field of software testing and test automation. Synonym-based locators are used to identify web elements based on their associated labels or other text content, in addition to standard HTML attributes.

Several studies have explored the use of synonym-based locators for web element identification. For example, Kumar et al. proposed a technique for web element identification that uses a combination of traditional locators and synonym-based locators. They found that their technique improved the accuracy and efficiency of web element identification compared to traditional techniques that rely solely on standard HTML attributes.

In another study, Ma et al. proposed a machine learning-based technique for web element identification that uses a combination of image features and text features. They found that their technique was effective at identifying web elements using both traditional locators and synonym-based locators.

Other studies have explored the use of natural language processing (NLP) techniques for synonym-based locator identification. For example, in their study on web page testing, Li et al. proposed a technique that uses NLP techniques to identify synonym-based locators based on their associated labels or other text content.

In terms of test automation frameworks, several popular frameworks such as Selenium and Appium provide support for synonym-based locators. For example, Selenium provides support for locating web elements based on their associated text content using methods such as "By.linkText" and "By.partialLinkText". Appium also provides support for synonym-based locators using methods such as "By.AccessibilityId" and "By.IosUIAutomation".

Overall, the literature suggests that the identification of web elements using synonym-based locators is an important area of research in the field of software testing and test automation. There are various techniques and frameworks available for supporting this task, including machine learning-based techniques, NLP-based techniques, and support for synonym-based locators in popular test automation frameworks such as Selenium and Appium [4].

3 Problem Statement

To address this problem, this paper proposes a test automation framework that uses a synonym-based locator identification algorithm to identify web elements on a web page using a suitable data structure. The algorithm uses a database of synonyms for each attribute type to identify web elements according to the type. The proposed framework also uses a "BeautifulSoup" package to elevate the particular data and avail result for the intended request.

4 Synonym-Based Locator Identification Algorithm

The synonym-based locator identification algorithm works by creating a database of synonyms for each attribute type. For example, the database might contain a list of synonyms for the ID attribute, such as "identifier," "name," or "tag." When a test script needs to interact with a web element, it first looks up the synonyms for the desired attribute type and searches for the element using each synonym in turn. If the element is found, the script can interact with it using the appropriate attribute.

There are several algorithms that can be used for synonym-based locator identification in test automation. Here, we discuss some commonly used algorithms:

1. **String Matching Algorithm**: This algorithm compares the text content of web elements with the expected text values in test scripts. This approach can be effective for identifying web elements based on their associated text content, such as labels or headings [5].
2. **Natural Language Processing (NLP) Algorithm**: NLP algorithms can be used to identify synonym-based locators by analyzing the natural language patterns and relationships between web elements and their associated text content. This approach can be effective for identifying web elements based on their associated labels or other descriptive text [6].

3. **Machine Learning (ML) Algorithm**: ML algorithms can be used to learn the relationships between web elements and their associated text content, and then apply this knowledge to identify synonym-based locators. This approach can be effective for identifying complex relationships between web elements and their associated text content [7].
4. **Hybrid Algorithm**: A hybrid algorithm combines multiple techniques, such as string matching, NLP, and ML, to identify synonym-based locators. This approach can be effective for identifying a wide range of synonym-based locators with high accuracy [8].

5 Carried Research Work

HTML Form Element Matching Algorithm: This algorithm compares the web elements and finds out the similar elements based on the attributes, and result is retrieved on the basis of HTML locators. In this study, we have made use of HTML form element matching algorithm.

6 Implementation of the Proposed Work

Step 1: START
Step 2: GET request = URL
Step 3: GET HTML response.content = HML.parser
Step 4: Find form element → soup.find('form')
Step 5: Create → Class Locator() with locators function
Step 6: app.py & index.html {% if inputs %} and {% for inputs in inputs %} Displays the web element
Step 7: In app.py, add parameters using render_template → html, URL, inputs
Step 8: END.

7 Proposed System Design

Here, in Fig. 1, the user enters the URL with the help of user interface, and then, the data within that URL is scraped which further is rendered, and the intended output is displayed back to the user.

Note: We have considered an registration form as an entity for processing the research and retrieved the results.

Step 1: Get the egistration form URL and store it in a variable.
Step 2: To make a GetRequest from the URL, we have to use request library to get a response of a URL.

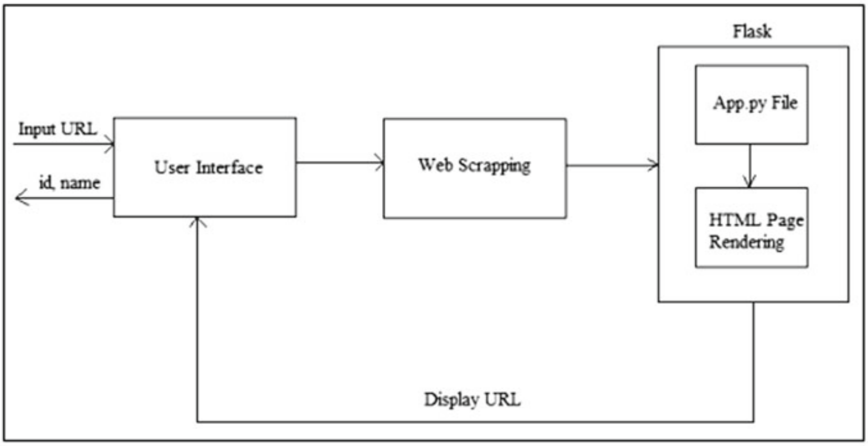


Fig. 1 Block representation of a system

Step 3: To scrap the data of an HTML page, we have to make use of BeautifulSoup Library.
Step 4: Find element of a registration form and store it in a variable.
Step 5: Find all input and select fields in the form.
Step 6: Generate HTML locators for each input and select field.

8 Flowchart of Synonym-Based Locator Search

The above-mentioned flowchart in Fig. 2 represents the working flow of a research work which is taken into consideration.

9 Results of Implementation

Figure 3 represents the application in which URL is to be inserted for retrieving the locators of the HTML web elements.

Figure 4 represents the submission of an URL and the results, i.e., locators are been displayed for synonym elements on the user interface.

Figure 5 represents the submission of an URL through a code and retrieving the synonym locators in the console.

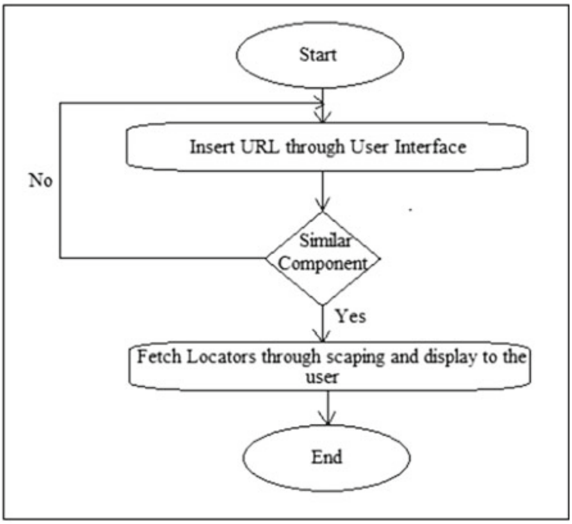


Fig. 2 Sequential flow of SBLS

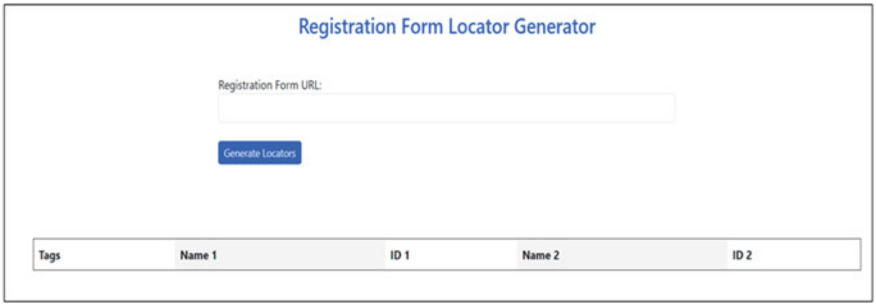


Fig. 3 View of an application for inserting URL to retrieve locators

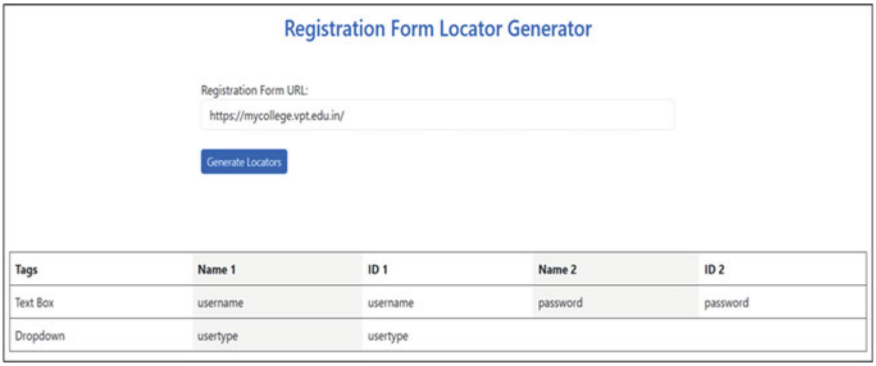


Fig. 4 Retrieving of results after submitting URL

Fig. 5 Console-based result retrieving

10 Comparison of Synonym-Based Locator Identification Algorithm

Algorithm	Description	Pros	Cons
String matching	Compares text content of web elements with expected values	Simple and easy to implement	May not be effective for identifying complex relationships
Natural language processing (NLP)	Analyzes natural language patterns and relationships between web elements and text content	Effective for identifying complex relationships	May require extensive knowledge of NLP
Machine learning (ML)	Learns relationships between web elements and text content	Effective for identifying complex relationships	May require significant data and expertise to train models
Hybrid	Combines multiple techniques to identify synonym-based locators	Can achieve high accuracy and efficiency	May require more resources and expertise to implement
HTML form element matching	Searches for the HTML locators using HTML form element	Achieve accurate locators for the URL passed for an HTML form element	No drawbacks rectified

11 Merits and Demerits for Test Automation

Merits

1. **Improved efficiency**: Automated testing can execute tests much faster and more reliably than manual testing, which can help improve the speed and efficiency of software development.
2. **Increased test coverage**: Automated testing can help increase the scope and depth of testing, enabling developers to test more scenarios and uncover more bugs than would be possible through manual testing alone.
3. **Consistency**: Automated tests are executed in a consistent and repeatable manner, reducing the risk of errors caused by human oversight or variation.
4. **Cost savings**: While there is an upfront cost to developing and maintaining automated tests, over time it can lead to significant cost savings as less time and effort is required for testing.
5. **Scalability**: Automated testing can be easily scaled to accommodate larger and more complex applications, making it an ideal solution for testing large-scale software systems.

De-Merits

1. **Initial investment**: Developing and maintaining automated tests can require a significant initial investment of time and resources.
2. **False sense of security**: Automated tests may give a false sense of security, leading developers to overlook manual testing and potentially miss critical bugs.
3. **Maintenance overhead**: As software changes, automated tests must be updated to reflect those changes, which can result in a maintenance overhead over time.
4. **Limited scope**: Not all tests can be automated, and there are some scenarios where manual testing is still required to ensure comprehensive testing.
5. **Technical expertise**: Developing and maintaining automated tests requires technical expertise, which may not be available in all development teams.

Overall, test automation can be a powerful tool for improving the efficiency and effectiveness of software testing, but it should be used judiciously and in conjunction with other testing methodologies to ensure comprehensive and reliable testing.

12 Conclusion

In this study, web elements are identified by using synonym-based locator identification algorithm. As compared to various techniques and algorithms, this algorithm helps to identify web elements using suitable data structure, i.e., list. The results are retrieved on the basis of Flask framework and BeautifulSoup.

Acknowledgements This research is carried out in ApMoSys Technologies Pvt. Ltd., Navi Mumbai, Maharashtra, India. We are sincerely thankful to Mr. Bibhu Padhy and Ms. Sangeeta Padhy for providing the setup to perform research.

References

1. Padhy P, Nayak P, Vaidya S, Web performance engineering—finding best data structure for automatically collected HTML locators. In: 2nd IEEE conference on computational systems and communications
2. Kuzu M, Islam MS, Kantarcioglu M (2012) Efficient similarity search over encrypted data. In: 2012 IEEE 28th International conference on data engineering. https://doi.org/10.1109/ICDE.2012.23, Arlington, VA, USA
3. Vaidya SA, Bhosale V (2016) Search over encrypted data. In: 2016 International conference on inventive computation technologies (ICICT). https://doi.org/10.1109/INVENTIVE.2016.7830194, Coimbatore, India
4. https://www.yogiti.com/2023/04/14/advantages-of-usingtest-automation-frameworks-for-quality-assurance/
5. Hakak SI, Kamsin A, Shivakumara P, Gilkar GA, Exact string-matching algorithms: survey, issues, and future research directions. IEEE Access. https://doi.org/10.1109/ACCE
6. Gelbukh A, Natural language processing. In: Fifth International conference on hybrid intelligent systems (HIS'05), https://doi.org/10.1109/ICHIS.2005.79, Rio de Janeiro, Brazil
7. Jha KK, Jha R, Jha AK, Al Mahedi Hassan M, Yadav SK (2021) A brief comparison on machine learning algorithms based on various applications: a comprehensive study. In: 2021 IEEE International conference on computation system and information technology for sustainable solutions (CSITSS). https://doi.org/10.1109/CSITSS54238.2021.9683524, Bangalore, India
8. Durairaj AK, Chinnalagu A (2021) Sentiment analysis on mixed—languages customer reviews: a hybrid deep learning approach. In: 2021 10th International conference on system modeling & advancement in research trends (SMART). https://doi.org/10.1109/SMART52563.2021.9676277

Generation of Smart Power with the Aid of Piezoelectric Sensors

T. P. Kausalya Nandan, B. Koushika Varsha, B. Aravind Reddy, and C. H. Sri Harsha

Abstract Numerous methods exist for generating power using natural resources such as water, wind, thermal energy, and solar energy. Sunlight is a major source of energy. Despite this, the world's population is growing rapidly. There may be an increase in resource demand across the globe due to overpopulation. As a result, energy demand will also increase, but the resources used for this will decrease over time. Thus, saving energy or finding an alternative way to generate power conventionally is crucial. Nuclear, coal, and wind are the most common conventional methods. Alternative energy sources have become more popular for heat and electricity production due to environmental problems. Alternative energy sources include biomass, geothermal, wind, and solar energy used for producing electricity. Despite this, the power plants emit very harmful pollution, which is why there is a need to find another pollution-free power source. This proposed model can be used to produce power using a piezoelectric sensor in an eco-friendly manner. The piezoelectric sensor works by producing a potential difference of the same magnitude of force when pressure is applied to it. This potential difference is then converted into electrical energy. Therefore, for generating power, piezoelectric sensors are placed beneath the road. Power is derived from the mechanical stress applied to the sensors whenever vehicles or people pass. Thus, the power can be used for various applications like lighting streetlights, charging mobile phones, security cameras, etc.

Keywords Power generation · Piezoelectric sensor · Charger module · Battery · Arduino · NI LabVIEW

T. P. Kausalya Nandan (✉) · B. Koushika Varsha · B. Aravind Reddy · C. H. Sri Harsha
B V Raju Institute of Technology, Electronics and Communication Engineering, Narsapur, Medak, India
e-mail: kausalya.nandan@bvrit.ac.in

B. Koushika Varsha
e-mail: 19211a0431@bvrit.ac.in

B. Aravind Reddy
e-mail: 19211a0437@bvrit.ac.in

C. H. Sri Harsha
e-mail: 19211a0444@bvrit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_49

1 Introduction

Humans need the energy to fulfill their basic needs such as water, food, health, and education. Energy is fundamental to fulfilling these needs. Any nation or region's growth depends heavily on its access to energy. Furthermore, energy is vital to meeting core social and human needs including lighting, transportation, the provision of water, food, health care, and education. Energy is a key factor in determining each country's economic and social growth because all of these services serve as the benchmarks by which a country's success and development are evaluated [1]. The energy can be generated from both conventional and non-conventional sources. Coal, fossil fuels, thermal power, nuclear power, and hydroelectric electricity are examples of conventional sources. Solar energy, wind energy, and tidal energy are some of the non-traditional sources. The primary resource for generating electricity is water. Hydropower, commonly referred to as hydro energy, is a renewable energy source that uses water that is stored in dams and flows through rivers to produce electricity. Sources of thermal energy include solar heat, electric heat pumps, geothermal heat, and fossil fuels including natural gas, coal, and oil. A power station that converts heat energy into electrical energy is referred to as a thermal power station. Solar energy generates power by harnessing solar energy. Solar panels collect energy from the sun's rays as they hit them through their PV cells. A cell's internal electric field causes this energy to generate electrical charges that move. In reaction to an internal electric field within the cell, this energy generates electrical charges that move, which results in the flow of electricity [2]. A piezoelectric sensor is used to produce power from the movement of the feet. A piezoelectric sensor is a device that uses the material in its construction to transform shear load into electrical energy. As a result, the quartz material transforms the stress that is imparted to the sensor from a footstep into electricity, which produces power [3]. Although there is a risk that conventional energy sources may become less abundant, there is also a chance that their costs will rise over time, and as the population grows, there will be a greater need for electricity. Furthermore, this results in more energy being wasted in many ways. Hence, the hardest challenge will be to solve this issue. So, the goal is to create a model that makes it simple to generate power utilizing piezoelectric sensors. Power will be generated because of adding pressure to the sensors. A strong current is produced by connecting several piezoelectric sensors in parallel. A bridge rectifier, which changes AC to DC, rectifies the current produced. A rectifier output is provided into the charger module since the system needs 3.7 V of power, although that voltage might not always be enough. A charger module is used to address this problem, increasing voltage while also shielding the battery from short circuits. Each of the sensors, the voltage sensor, and the current sensor receives additional rectifier output. The charger module's output is coupled to a rechargeable battery. Therefore, the battery can serve a variety of functions after being charged.

2 Literature Survey

In modern society, power conservation and generation play a crucial role, which results in an increase in power demand. The use of different technologies allows different methods of generating power to be employed. Several methods exist for generating power, and one is by using piezoelectric sensors. Researchers have worked on improving these methods and proposed many different solutions. Shahbaz Ahmad et al. [4] provided an IOT-based solution for smart streetlights is being developed. Across the sidewalk, piezoelectric sensors are positioned and linked to the main battery. Hakpyeong Kim et al. [5] presented a knowledge-based methodological strategy for future research on the integration of energy-saving systems in smart cities. Avdhoot Sunil Kulkarni et al. [6] their thesis models tiny hydropower plants using MATLAB and Simulink. When supplying the regional electricity demand, the establishment of a microhydropower plant will significantly lower the strain without having a detrimental impact on the environment. Weidong Chen et al. [7]. This study intends to give an in-depth analysis of green power healing from thermal treatment. Xiaokang Peng et al. [8]. The paper provides an overview of wind power generation using multiphase energy conversion and introduces recent technological advances in this field, such as wind turbine generator design, converter topologies, model design, and wind turbine control. Vineet Kumar Singh et al. [9] discussed the operational aspects of hydropower plants for generating energy with minimal environmental impact and operating costs. Kamboj, Akshat, et al. [11] discussed the production of electricity via footstep using the piezoelectric sensor. Frede Blaabjerg et al. [10] discussed some trends in the power electronics technologies used for wind power generation. Abdullah Al Mamun Momen et al. [12] proposed a hybrid solar-wind system to charge a train's batteries for only electric light loads. Minal Ghute et al. [13] propose a piezoelectric generator that uses walking vibrations and pressures to create electrical power. Dr. Meena et al. [14] discussed the technique for extracting the energy from human mobility using a piezoelectric sensor, and it also describes how to use the stored energy to reliably charge a mobile phone using RFID. Anis Maisarah Mohd Asry et al. [15]. The generation of electricity using piezoelectric transducers with bending mechanisms is demonstrated in this work. 3D printing is employed for the bending mechanism. S. Suresh Kumar et. Al. [16] discussed using a piezoelectric sensor array to generate power for highways. The system contains an Internet of Things (Arduino) with sensors, which are all controlled by the microcontroller known as Arduino. Zhipan Zhang et al. [17]. In this work, paper is employed to produce power because of moisture influx. Si-Bum Cho et al. [18] article addressed how to affix a piezoelectric sensor to the placement of a coastline structure relative to the sea. Hailu YangIn et al. [19]. In this paper, a Piezoelectric Energy Harvester (PEH) of the staggered matrix type is constructed with a safety package, which can enhance the PEH's efficiency and operational duration. Sarala T et al. [24] explained piezoelectricity use in smart street lighting. Researchers in [20] developed a paper-based piezoelectric touch sensor using screen printing. Kafi Mohammad Ullah et al. [21] explained how to produce electrical power through speed breaker. M. Nitashree et al. [22] delve into piezoelectric sensor-based footstep power generation. Anurag Pandey et al. [23] discussed a versatile work for open space that generates energy by using the road and unconventional energies. It illustrates how energy can be generated by roads and transportation in diverse manners and using various storage technologies. Although researchers have used various power generation methods for decades, it is still necessary to generate more power abundantly for the growing population. Furthermore, power should be conserved since most of them waste it. In order to prevent power shortages in the long run, conserving energy is important and crucial.

3 Proposed Model

Using this proposed model as described in Fig. 1, one can generate power using a piezoelectric sensor in an eco-friendly manner. This Model generates electricity by putting piezoelectric sensors underneath the road. Whenever vehicles or people pass, power is generated due to the mechanical stress applied to the sensors. Thus, the generated power can be used for various applications. This generated electricity is stored in a battery and can be discharged when we need electricity. In this application, piezoelectric sensors are used which work on electromechanical coupling phenomena. Where a material, namely quartz, barium, or titanate will produce an electrical voltage when exposed to mechanical stress. The emitted power is stored in a battery and with the help of Arduino and LabVIEW, we can monitor the voltage and current values.

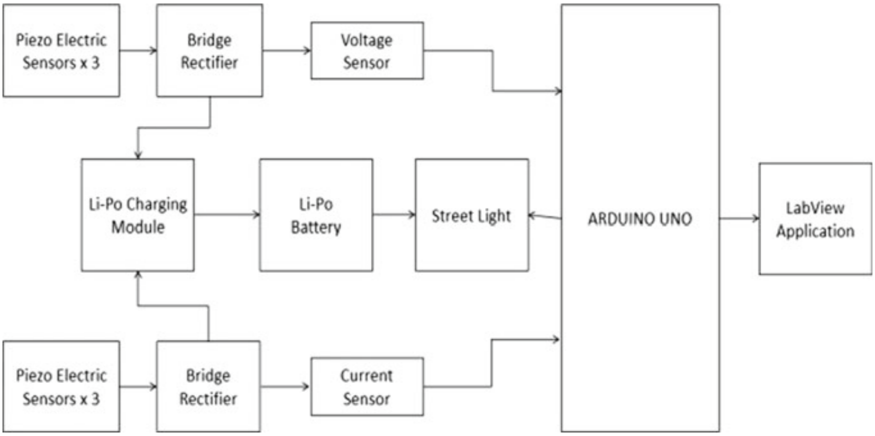


Fig. 1 Smart power generator

4 Experimental Setup

The experimental set has been designed with the help of Arduino Uno, piezoelectric sensor, bridge rectifier, current sensor, voltage sensor, Li-Po charging module, Li-Po battery, and LED as shown in Fig. 2. Three sensors from each set of two parallel-connected sets of piezoelectric sensors help to generate a high current in the suggested approach. Due to the electrochemical coupling phenomenon of the piezoelectric sensor, pressure on the road caused by a moving vehicle generates power. The purpose of a bridge rectifier is to correct the current generated. While the suggested system only requires 3.7 V of power, a rectifier output is sent into the charger module; nevertheless, that voltage could not be enough in some circumstances. However, that voltage might not be adequate for all applications. As a result, a charger module is employed that boosts voltage while also shielding the battery from short circuits. To measure voltage and current levels, voltage and current sensors are employed. As a result, a battery is employed to store the energy generated. Arduino serves as the control unit of the proposed model, and NI LabVIEW monitors the voltage and current values. Through LabVIEW's VISA module, the interface of Arduino is done, which obtains information from Arduino. The whole setup is connected to Arduino UNO, with 2 analog pins to piezoelectric sensors (A2 and A3), 2 analog pins (A0 and A1) for voltage and current sensor, an A4 pin for output, and 2 pins for ground and supply. It is necessary to implement the same serial communication in LabVIEW that is employed by Arduino. The data is read from the COM port via VISA read VI and displayed as string output. This string output is the combined data of both voltage and current. For splitting the data, we used split VI. Therefore, to get the rounded voltage and current values we used split vi so that the values can be displayed in the front panel. For example, if the current value is 0.80 mA, then the output would be rounded to 1 mA.

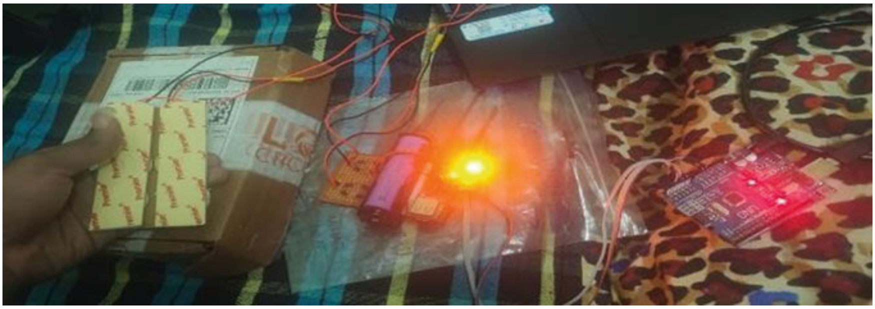


Fig. 2 When pressure is exerted on the piezoelectric sensor, then the LED will be illuminated

5 Flowchart

The flow of the work is explained with the help of a flow chart as described in Fig. 3. The Arduino voltage and current sensor are initially initialized, then when a car or human goes by the pavement or roads because of the pressure enforced on the sensors, the sensor mechanism generates power, which then leads the Li-Po charging module to power up the battery and enables the voltage and current values to be read by the sensors. This procedure continues, and the output can be seen in LabVIEW simultaneously every time a vehicle moves.

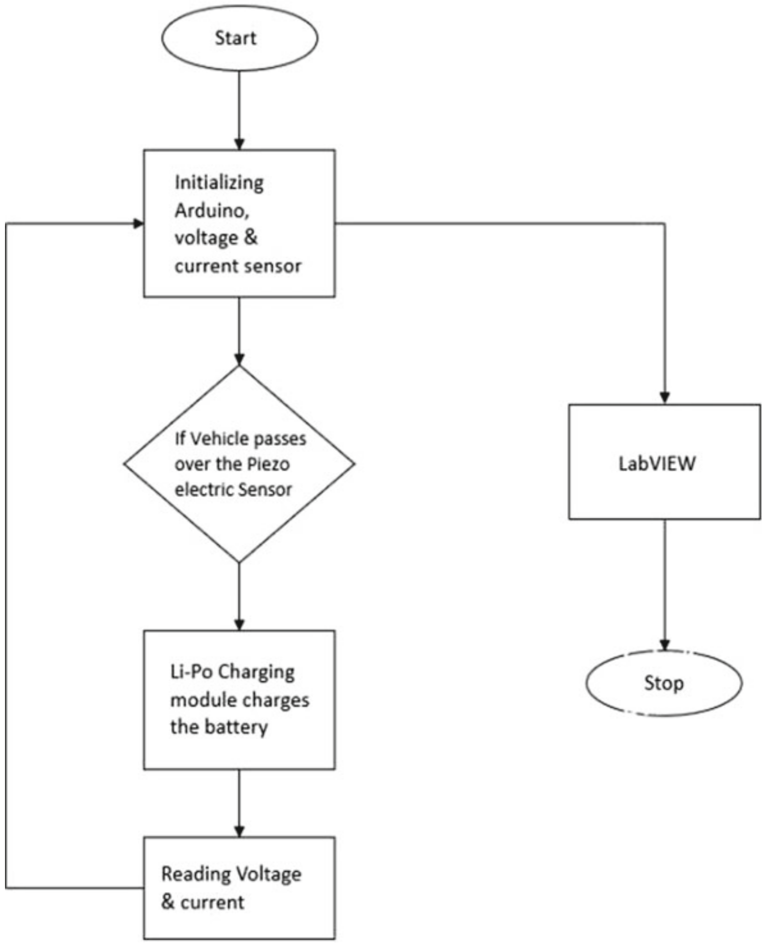


Fig. 3 Flowchart

6 Results

When pressure is not applied to the piezoelectric sensor, then the values of voltage 3.11 and current 0 are shown in Fig. 4.The values of current and voltage are displayed in the LabVIEW.

As illustrated in Fig. 5, when pressure is put on to the piezoelectric sensor, the voltage value is 3.91 V and the current value is rounded up to 1. Therefore, when pressure is applied to the sensors, current and voltage readings rise in this way. In this manner, pressure is taken from the sensors whenever vehicles or people move on the roads or sidewalks, boosting the current value. Hence, the battery retains the power generated, allowing it to serve a wide range of applications.

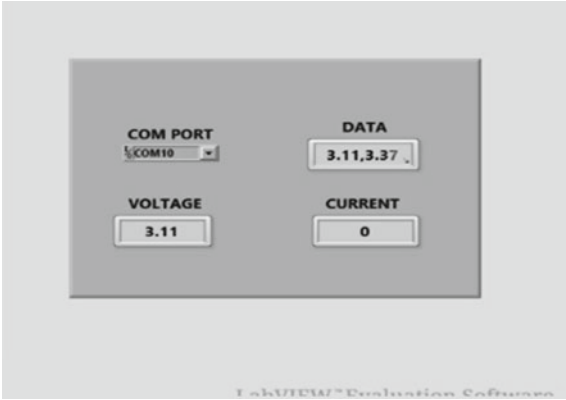


Fig. 4 Output of sensor when no pressure applied

Fig. 5 Output of sensor when pressure applied

7 Future Scope and Limitations

There are several potential areas for study and use for the power production employing piezoelectric sensors, which has a bright future. Here are some potential applications for this technology in the future:

IoT device growth needs reliable and independent power sources. Wireless sensor nodes, Internet of Things (IoT) gadgets, and low-power electronics can all be powered by piezoelectric energy harvesting. Piezoelectric sensors can offer a constant and renewable power supply by using environmental vibrations or human interactions, obviating the requirement for regular battery changes.

Research is increasingly being done on the creation of self-powered sensors and devices. Piezoelectric sensors can make it possible to build autonomous, self-powered devices that don't require any external power sources or regular battery replacement. This has major effects on a variety of industries, including remote monitoring, environmental sensing, and smart infrastructure.

While promising, piezoelectric sensor-based power production does have certain limitations.

Piezoelectric sensors often produce modest quantities of power, especially when exposed to little vibrations or movements. The vibration's amplitude and frequency directly affect the power output. It can be difficult to produce a large quantity of power using piezoelectric materials alone, and it might not be adequate for high-power applications.

Environmental variables may have an impact on the performance of piezoelectric sensors. For example, temperature variations and harsh outdoor conditions might affect the sensor's work. The performance or longevity of the piezoelectric materials may also be impacted by moisture exposure or contact with corrosive chemicals.

Using piezoelectric sensors, the conversion of mechanical energy into electrical energy is not 100% efficient. Energy conversion involves certain intrinsic losses, including internal resistance, electrical losses, and mechanical losses. The field continues to face difficulties in maximizing efficiency and reducing these losses.

8 Conclusion

This proposed model helps to generate power easily with the help of a sensor. This allows the generated power to be used for everyday purposes. For example, power can be used for street lighting. It is found that existing technology wastes lots of energy, as some streetlights are still on during daylight, while others do not go on until the dark is falling. This proposed model can be used to solve this problem. Therefore, this application would generate electricity in an eco-friendly manner.

References

1. Roga S, Bardhan S, Kumar Y, Dubey SK (2022) 52. https://doi.org/10.1016/j.seta.2022.102239
2. Shaikh MRS, Waghmare SB, Labade SS, Fuke PV, Tekale A, ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887 5(IX)
3. Helonde NJ, Suryawanshi P, Bhagwatkar A, Wagh A, Vetal P, India-Volume 9 Issue XII Dec 2021, ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
4. Ahmad S, Siddique A, Iqbal K, Hussain A, Ijaz-H A (2021) 10(01)
5. Kim H, Choi H, Kang H, An J, Yeom S, Hong T (2021) A systematic review of the smart energy conservation system: From smart homes to sustainable smartcities. 140:110755. https://doi.org/10.1016/j.rser.2021.110755
6. Kulkarni AS, Ambekar RS (2021) Power generation from water in pipeline though hydro generator. XIII(05)
7. Chen W, Huang Z, Chua KJ (2021) https://doi.org/10.1186/s13705-021-00304-6
8. Peng X, Liu Z, Jiang D (2021) A review of multiphase energy conversion in wind power generation 147:111172, https://doi.org/10.1016/j.rser.2021.111172
9. Singh VK, Singal SK (2017) Operation of hydro power plants-a review. Renew Sustain Energy Rev 69:610–619
10. Blaabjerg F, Ma K (2017). 105(11). https://doi.org/10.1109/JPROC.2017.2695485
11. Kamboj A et al (2017) Design of footstep power generator using piezoelectric sensors. In: 2017 International conference on innovations in information, embedded and communication systems (ICIIECS). IEEE
12. Al Mamun Momen A, Nath SK, Md. Golam Mostafa SK (2019) Design and implementation of hybrid power generation system for electric loads of train. In: 2019 5th International conference on advances in electrical engineering (ICAEE). https://doi.org/10.1109/ICAEE48663.2019.8975657
13. Ghute M, Korde M, Kamble K, Khade S, Niranjane V (2021) Power generation using renewable energy sources. In: 2021 Fifth international conference on I-SMAC. https://doi.org/10.1109/I-SMAC52330.2021.9640802
14. Chavan M et al (2020) Footstep power generation using piezoelectric sensor and distribution using RFID. Int Res J Eng Technol 7(09):1416–1420
15. Asry AMM, Mustafa F, Ishak M, Ahmad A (2019) Faculty of Engineering Technology, University Tun Hussein Onn Malaysia, Malaysia, Vol. 10, No. 1, March 2019, pp. 562–567, ISSN: 2088-8694. https://doi.org/10.11591/ijpeds.v10.i1
16. Suresh Kumar S, Kaviyaraj R, JeniNarayanan LA Saleekha (2019) https://doi.org/10.1109/ICACCS.2019.8728367
17. Zhang Z, Gao X, Xu T, Shao C (2019) Electric power generation by paper materials. J Mater Chem A7(36). https://doi.org/10.1039/C9TA08264F
18. Cho S-B, Kim H-D, Shim K-T (2018) 2018, Article ID 7986438
19. Yang H et al (2018) A preliminary study on the highway piezoelectric power supply system. Int J Pavement Res Technol 11(2):168–175. Singh VK, Singal SK (2017) Operation of hydro power plants-a review 69. https://doi.org/10.1016/j.rser.2016.11.169
20. Emamian S, Narakathu BB, Chlaihawi AA, Atashbar MZ, Fabrication and characterization of piezoelectric paper based device for touch and force sensing applications. In: Conference. https://doi.org/10.1016/j.proeng.2016.11.248,2016
21. Ullah KM, Ahsan-uz-Zaman KM, Hosen S, Khan RH, Parvin S (2016) Electrical power generation through speed breaker. In: IEEE 2016 9th International conference on electrical and computer engineering (ICECE)—Dhaka, Bangladesh (2016.12.20–2016.12.22), 2016 9th international conference on electrical and computer engineering (ICECE), 30–33. https://doi.org/10.1109/ICECE.2016.7853848
22. Nita Shree M et al (2015) Footstep power generation using piezoelectric material. Int J Adv Res Electron Commun Eng 4:2503–2506

23. Pandey A, Sharma JP (2015) Power Generation & automation system designing for highway. In: IEEE 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)—Coimbatore, India (2015.3.5–2015.3.7), 2015 (ICECCT) https://doi.org/10.1109/ICECCT.2015.722597
24. Sarala T, Poornima M, Lekhana HD (2020) Generation and utilization of electricity using footsteps as a source of energy. In: 2020 International conference on recent trends on electronics, information, communication & technology (RTEICT). IEEE

Prediction of Higher Education Student's Programming Skills Using Item Response Theory

Vinayak Hegde and A. V. Vismaya

Abstract The current review plans to dissect Object-oriented programming multiple-choice questions got from a preliminary testing led in different degree colleges. Earlier in classical test theory, responses were modeled based on the total number of correct answers. If candidate A provides more correct answers than candidate B, candidate A is thought to have an ability level than B. Item Response Theory is a theory, that describes how to model a candidate's response in such a way that we can examine each response separately rather than as a group. In this proposed system, three different models are used to design the IRT-1PL,2PL, and 3PL model which uses certain parameters to evaluate the test than just the number of correct answers. Initially, the proposed system calculates the discrimination index and difficulty index of the questions answered by the students and then it uses the log-likelihood function and Newton Raphson method to get the IRT scores and also uses Bayesian methods to provide the feedback to each student. The accuracy of the current research is 92.4%. IRT can be used to determine scaled scores on a new type of test without collecting data from those who take it. If the test is given by computer, IRT can send a scaled score to each test taker immediately after the test. Along with the score the system also provides feedback to the student and the student will get to know about his skill set.

Keywords Item response theory · Log likelihood function · Newton Raphson method · Bayesian model · Educational data mining

---

V. Hegde (✉) · A. V. Vismaya
Department of Computer Science, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Mysuru Campus, Mysuru, India
e-mail: vinayakhegde@my.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_50

1 Introduction

Formulation of test items with actual test scoring is the main concern of the item response theory. To configure various abilities like math ability and behavioral abilities, we design the test items accordingly. Responses that the learner gives can be binary or ordinal.

The data which is in psychological evaluation, and educational testing such types of data are traditionally analyzed by the IRT models. Over the last few decades, educational assessment has increasingly relied on IRT-based techniques to develop tests. Item response theory is now used to raise all prime educational examination, including SAT and GREs exams, since the methodology can greatly increase calculation precision and reliable Ness while potentially reducing evaluation time and attempt, particularly through computerized adaptive testing.

One-parameter model also called the Rasch model is the basic form of IRT model. It posits that a test taker's likelihood of successfully answering an item is determined by only one feature of the question—its difficulty. That assumption, of course, is not always valid. Nonetheless, this version of IRT is utilized on many tests and works well in general. Another popular variant of IRT takes into account two elements of the item: its complexity and discrimination. Discrimination refers to how far a question divides test takers who are higher on the ability scale from those who are lower. So, what is the point? The difficulty of the item determines this. The more difficult the object, the higher the ability value where it discriminates most effectively. The "two-parameter model" refers to the IRT variant that employs these two item attributes. It necessitates more information than the one-parameter model, but it frequently yields scores that are more accurate in representing each test taker's ability. There is also a "three-parameter model" available, which considers three components of the item: difficulty, discrimination, and how easy it is for a low-ability test taker to predict the correct answer. This variant of IRT requires even more data than the two-parameter model, yet the results are typically extremely comparable.

2 Literature Survey

To understand more about the item response theory, there are a few works that helped to get a better idea about how IRT works and its various types of models. Hoang et al. [1] Rasch model is mainly encouraged in this work which also supports the communication between the student and the item. Lee et al. [2] This work looks into a systematic method of estimating students' ability to solve issues in a computer-based learning environment is the main subject of this research. Wang et al. [3] The logistic model is mainly used in this research as a sample, and this article presents the basic models and parameter estimation methods of IRT, particularly the artificial intelligence-based IRT parameter estimation algorithms. Chen et al. [4] This paper discusses how to adapt traditional item response theory (IRT) to handle non-crisp fuzzy responses and estimate learner ability using a revised estimating function of learner ability. Based on the courseware modeling requirement, this paper also put forward a program modeling operation for determining courseware difficulty parameters and designing program content for personalized courseware recommendation services. Vega et al. [5] This work describes IRT and how it may be used in an online course test scenario to generate adaptive assessments within a programming course from online repositories. Tian et al. [6] This work discloses the measurement attributes of quality analysis which is done by using item response theory. Furthermore, this work finds out the several framework relations of the questions. Hirose et al. [7] This work deals with adaptive testing and is frequently included in IRT test systems. This type of testing method automatically chooses the most applicable items for test takers. Noguchi et al. [8] This work talks about the one critical issue is the computational complexity of IRT parameter estimation. To address this issue, they proposed an IRT parameter estimation method based on the Holonomic approach. Moraes et al. [9] A new IRT model is proposed in this work which suits for designing the absolute errors of regression algorithms, and nonnegative unbounded responses are also considered. Mohamad et al. [10] To check the strength and weaknesses of CTT and IRT is the main aim of this work and to create recommendations for when the theories should be used. Jianhua et al. [11] In this research, a few latest associate criteria, such as the logistic regression method and the symmetric relative entropy method, are observed. Lakshman et al. [12] A new IRT model is executed in this work which is used in Web-based training, and clustering probability is calculated from the k means algorithm. Li Chen et al. [13] In this paper, they are mainly trying to iredict consumers decision making which is significant for improving retail performance. Eye movement data is collected to reflect the shopping process. IRT is used to examine the bond between eye-tracking data and decision making. Rungrat et al. [13] This paper's contribution is to put forward an algorithm for implementing IRT to assess systems. As part of the "KidArn" project, the algorithm is verified in view of its uses for real-time issues. Hirose et al. [14] This system has proposed a new method for estimating difficulties with the problem and students' potential, EM algorithm, and IRT is used. Ueno et al. [15]. This work is an altered graded item response model with assessment-criterion parameters from an overseer. The advantage of this idea is the potential to assess learners' abilities on the same scale even if the overseer uses them, unlike limitations.

Arnold et al. [16] The idea of quantitative penetration testing based on particular metrics to identify the difficulty rather than the likelihood of assaults is proposed in this work. Li et al. [17] The plan of a new form of test paper composition algorithm is used which is based on IRT is put forward, with the eventual aim of precisely calculating the candidates' potential. Hegde et al. [18] The main of this paper was to provide a better education system for online education. This work has used the k-nn algorithm to provide automated personalized feedback. Hegde et al. [19] This system mainly works based on the Rasch model, used for the evaluation of the test scores of the aptitude of the learner.

Table 1 Pattern of the quiz conducted

Topics covered in quiz	Number of questions
OOPS	5
Conditional statements and loops	5
Logical operators	2
Arrays	5
Functions and case study	5
Basic concepts	3

3 Methodology

3.1 Data Collection

The current research builds on the first stage of administering a quiz with Java questions to help students identify their skill set. To predict the skill set of the students, the way of modeling the questions, the way of conducting exams, and the student's mindset everything plays an important role. Hence, the common quiz is designed so that it assists experimenters and upcoming educators in learning about the students' programming abilities.

For further analysis, the information was collected with the help of Google forms, and around 150 learners filled out the form. The data was collected from various classes and colleges. The questionnaire was constructed in such a way that it covers practically all types of Java programming language questions. The Java programming language was used to create the quiz. Subjects such as object-oriented programming, conditional and looping statements, classes, objects, arrays, and case study kinds of questions were included. The initial step is the preprocessing of the assessment data in the proposed idea, points for each segment are retrieved, and applied maximum likelihood function and Newton Raphson method to find the ability using the Bayesian model the feedback is provided to each student for the responses received from the quiz.

For this, an interface has been developed that will allow the IRT technique to be used to assess the learners' responses. The framework is built with Python's Flask, which retrieves and displays information from students' replies. An admin[teacher] is granted all of the above privileges to load the exam into the system (Table 1).

3.2 System Architecture Diagram

See Fig. 1.

The process starts with the student logging into the web site. The login interface is connected to the student database, which helps in logging. By having the student's responses in the .csv form the student can their ability by choosing 1PL, 2PL, or

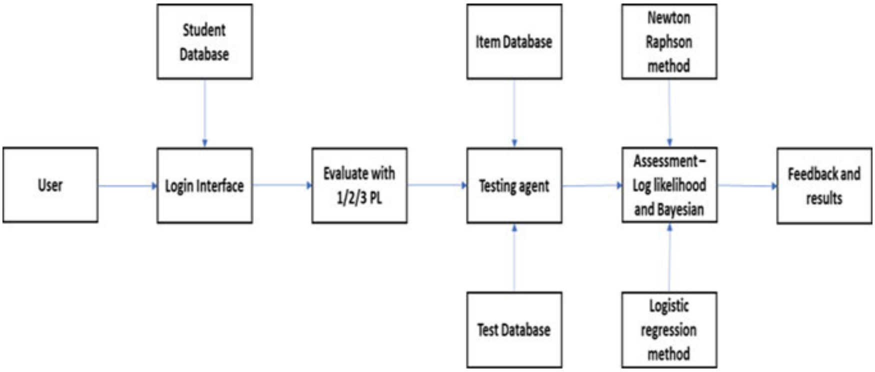


Fig. 1 System architecture diagram of the proposed work

3 PL. The testing agent manages the items and the test database. The assessment is done base on the methodology like the log-likelihood which is combined with Newton Raphson's method. The feedback is given by the Bayesian method.

3.3 Methodology

The system uses Python and Flask to implement the IRT with two modules—student and teacher module performing certain operations. Calibration is the process of determining response curves for a group of items testing the same ability in IRT. The one-parameter approach involves calculating a single number for each item's difficulty. It includes estimating two numbers: difficulty and discrimination, for each item in the two-parameter model. It entails estimating three numbers for each item in the three-parameter model: its difficulty, discrimination, and ease of predicting the correct answer. To find the difficulty index, we use the formula mentioned below.

$$\text{Difficulty index} = \frac{\text{PT} + \text{PB}}{n} \tag{1}$$

where PT is the marks scored by the higher group, PB is the marks scored by the lower group, and N is the total number of students who answered the question. The discrimination index formula is given below.

$$\text{Discrimination index} = \frac{\text{PT} - \text{PB}}{n} \tag{2}$$

where PT is the marks scored by the higher group, PB is the marks scored by the lower group, and n is half of the students who answered the question.

After finding the difficulty and discrimination index, to predict the student's ability we need to find the ability, i.e., (ϴ) value, so we need to calculate the one, two, and three parameters, respectively. We will first calculate the probability of a learner answering one item using the formula given below.

$$P_i(\theta) = c_i + \frac{1 - c_i}{1 + e^{-a_i(\theta - b_i)}} \tag{3}$$

$P_i(\theta)$ is the probability of getting the correct, ϴ is the ability, a_i is the discrimination index, b_i is the difficulty index, and c_i is the pseudo guessing parameter.

After finding the probability value, next we need to find the log-likelihood value of the student with multiple questions. So we use the formula given below.

$$\text{Log}(L(\theta)) = \sum_{i=1}^{n} xi \log(Pi) + (1 - xi)\log(1 - Pi) \tag{4}$$

P_i is the probability, and x_i is the responses from the learner. After getting the likelihood value, we will use the Newton Raphson method to find the root of the equation from which we find the ability value. By finding the ability value, we provide the respective feedback according to the performance done by each student.

4 Implementation and Results

There are two main modules in the proposed system, one is the student module, and the other one is the teacher module. The student module is mainly to assess the student's programming ability using item response theory where the student can log in to the system and upload the results CSV file and the student can choose whether the result should be in 1 parameter logic,2 parameter logic, or 3 parameter logic. And there is another option to take the live test in the system, and the student will get the result based on the test results. On the other hand, when it comes to the teacher module, the teacher can update the items in the teacher module, and the teacher can choose how many questions should be visible to students during the live test (Figs. 1, 2, 3, 4, 5, 6 and 7).

5 Discussion

The well-planned process of assessing learners' knowledge, skills, and capacities to discover finer ways to reframe instructing and studying can be termed an educational assessment. Anyway, in reality, nonetheless, instructive conditions vary among schools, courses, and course conveyance modes, making it trying to foster evaluations

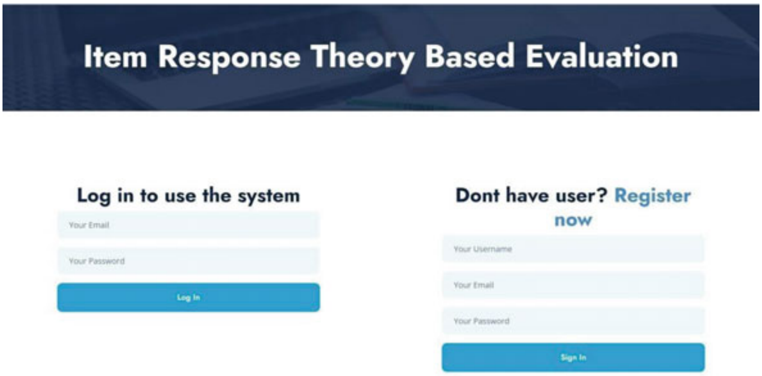


Fig. 2 Login page of the student module

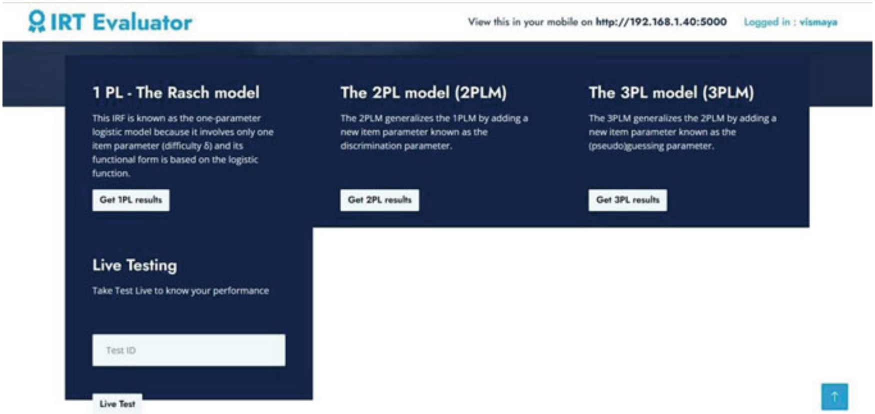


Fig. 3 Displaying the various parameter logics to get the result

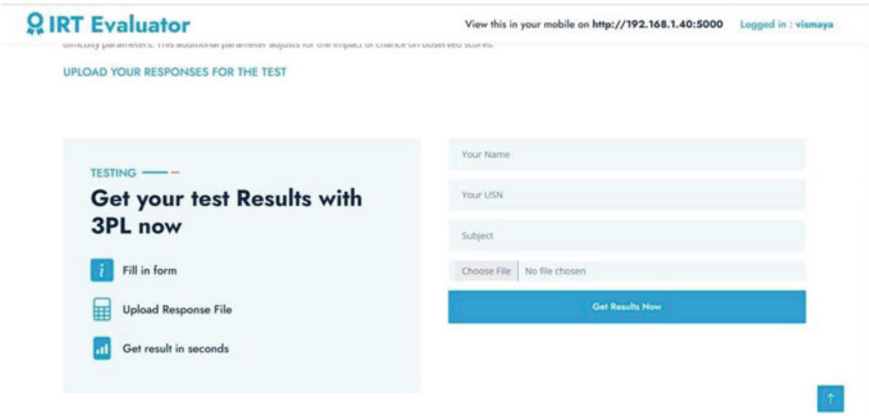


Fig. 4 Displaying the page where students can upload the result

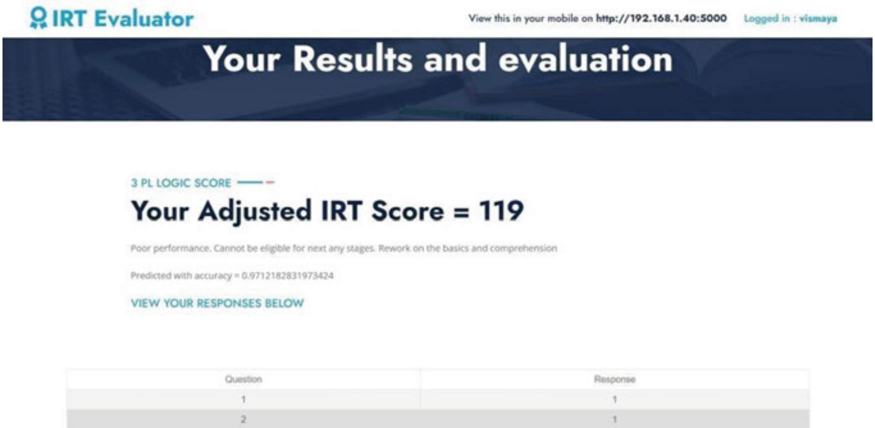


Fig. 5 Outcomes of the students with feedback are displayed

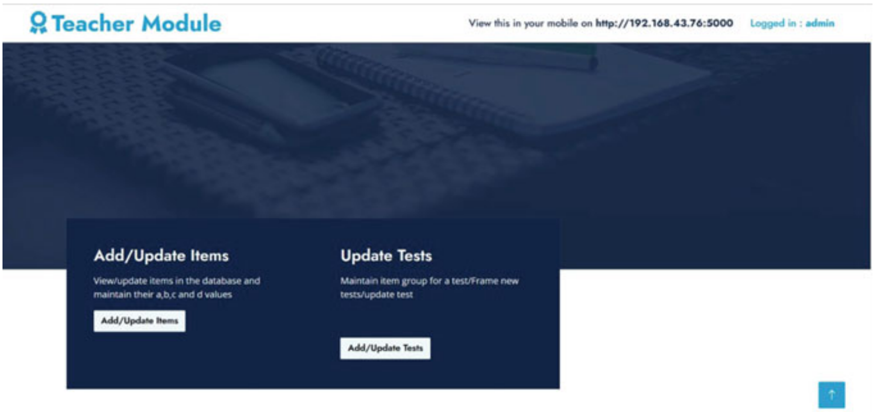


Fig. 6 Displaying the teacher module to upload the questions and tests

that mirror the complexities of educational material in test items. To put it another way, subject-based evaluation is required to precisely assess and forecast learners' skill results. The notion of IRT helps to deal with such scenarios by predicting the student's programming competence based on the responses he provided by taking into account the difficulty, discrimination index, and guessing parameter.

Fig. 7 Displaying teacher module to update the questions

6 Future Work

To increase the accuracy analysis of the given data, expanding the size of the dataset will enhance the current research. The language which is used for the assessment in the current work is Java. However, the assessment can also be done by giving another programming language. The current work only displays one student's reaction at a time; however, this work can be expanded by providing an interface that can load a large number of results from the teacher and distribute them to numerous students. The tests conducted here are primarily based on multiple-choice questions. Nevertheless, in the upcoming days, the proposed tool could be improvised so that the interface includes descriptive-type questions, which aid in obtaining better results when such questions are included.

7 Conclusion

According to the literature review, researchers attempted to discover several machine learning techniques to forecast the students' skills based on their responses. In the proposed work, we have tried to predict the student's programming skill based on mathematical models such as the maximum likelihood function and Newton Raphson method. The current work also tries to give accurate feedback to the students based on the responses using the Bayesian model. By making use of the logistic regression, model accuracy is calculated. There is a teacher module in the current work, where the teacher can update the tests and the number of questions that the student should answer in the live test platform. This work provides the flexibility of uploading the responses from the Google form as well as in the form of a live test platform. The accuracy calculated for the model is 92.4%. This model can be used in the field of education which enhances the evaluation process with less error, and it gives a brief idea of student's skills which will help the teacher in the assessment as well as the teaching process.

References

1. Binh HT, Duy BT (2016) Student ability estimation based on IRT. In: 2016 3rd National foundation for science and technology development conference on information and computer science (NICS). IEEE, 2016
2. Lee Y (2019) Estimating student ability and problem difficulty using item response theory (IRT) and TrueSkill. Inf Discov Deliv 47(2):67–75
3. Wang H, Ma C, Chen N (2010) A brief review on Item Response theory models-based parameter estimation methods. In: 2010 5th international conference on computer science & education. IEEE
4. Chen C-M, Duh L-J, Liu C-Y (2004) A personalized courseware recommendation system based on fuzzy item response theory. In: IEEE International conference on e-technology, e-commerce and E-service, 2004. EEE'04. 2004. IEEE
5. Vega YLP et al (2012) Application of item response theory (IRT) for the generation of adaptive assessments in an introductory course on object-oriented programming. In: 2012 Frontiers in education conference proceedings. IEEE, 2012. Rami
6. Haddad J, Kalaani Y (2015) Can computational thinking predict academic performance? In: 5th IEEE integrated STEM conference, 2015
7. Tian X et al (2017) Applying item response theory to analyzing and improving the item quality of an online Chinese reading assessment. In: 2017 6th IIAI International congress on advanced applied informatics (IIAI-AAI). IEEE
8. Hirose H, Aizawa Y (2014) Automatically growing dually adaptive online IRT testing. In: 2014 IEEE International conference on teaching, assessment and learning for engineering (TALE). IEEE
9. Noguchi K, Ito E (2018) Holonomic approach for item response theory parameter estimation. In: 2018 7th International congress on advanced applied informatics (IIAI-AAI). IEEE
10. Moraes JVC et al (2020) Item response theory for evaluating regression algorithms. In: 2020 International joint conference on neural networks (IJCNN). IEEE
11. Mohamad M, Omar AJ (2017) Measuring cognitive performance on programming knowledge: classical test theory versus item response theory. In: 2017 7th World ENGINEERING EDUCATION FORUM (WEEF). IEEE
12. Jianhua X, Shuliang D (2010) Model-based methods for test equating under item response theory. In: 2010 International conference on E-business and E-government. IEEE
13. Lakshman BV, Wijekulasooriya JV, Sandirigama M (2018) Modified item response theory (IRT) model and k means clustering for agent based E learning system. In: 2018 International conference on communication, computing and Internet of Things (IC3IoT). IEEE
14. Runzrat S, Harfield A, Charoensiriwath S (2019) Applying item response theory in adaptive tutoring systems for Thai language learners. In: 2019 11th International conference on knowledge and smart technology (KST). IEEE
15. Hirose H, Sakumura T (2012) Item response prediction for incomplete response matrix using the EM-type item response theory with application to adaptive online ability evaluation system. In: Proceedings of IEEE international conference on teaching, assessment, and learning for engineering (TALE) 2012. IEEE
16. Uto M, Ueno M (2015) Item response theory for peer assessment. IEEE Trans Learn Technol 9(2):157–170

17. Arnold F, Pieters W, Stoelinga M (2013) Quantitative penetration testing with item response theory. In: 2013 9th International conference on information assurance and security (IAS). IEEE
18. Hegde V, Pai AR, Shastry RJ (2022) Personalized formative feedbacks and recommendations based on learning analytics to enhance the learning of Java Programming. In: ICT Infrastructure and computing: proceedings of ICT4SD 2022. Springer Nature Singapore, Singapore, pp 655–666
19. Hegde V, Shushruth S (2022) Evaluation of student's performance in programming using item response theory. In: 2022 IEEE International conference on data science and information system (ICDSIS). IEEE

An Artificial Intelligence Integrated Technique for Screening Neurological Diseases with Sound as a Biomarker

Nysa Gaur and Reetu Jain

Abstract It has been discovered that neurological illnesses including Parkinson's disease (PD) and paralysis affect a person's voice. Common machine learning models like kNN and convolutional neural networks (CNN) have been used by researchers to analyse the variations in the pitch (in decibels) of sound produced by people with various diseases. For instance, those with Parkinson's disease (PD) have a pitch range of 55–65 dB, while people who are paralysed have a range of 45–55 dB, producing a deeper and thicker voice. Researchers recorded the speech of patients and healthy individuals, building a dataset, in order to construct a programme that can recognize the existence of neurological disorders (ND) like PD and paralysis in people with speech problems. They used Mel-frequency cepstral coefficients (MFCC) to extract features from the speech, and then, they used kNN and CNN to classify the data into PD, paralysis, and normal groups. With an overall accuracy rate of 60.1 and 56.1% for two laryngologists, the program's diagnostic ability was discovered to be on par with that of human professionals. Clinicians can correctly determine the stage of PD and suggest suitable therapies like therapy, medicine, or both by starting the script numerous times.

Keywords Neurological diseases · Sound · Biomarker · k-nearest neighbor · Convolutional neural network

1 Introduction

The term "neurological diseases" (ND), often known as "nervous system diseases," refers to illnesses that affect the brain, spinal cord, cranial nerves, peripheral nerves, nerve roots, autonomic nervous system, neuromuscular junction, and muscles. The

N. Gaur (✉) · R. Jain
On My Own Technology, Mumbai, India
e-mail: nysa.gaur@gmail.com

R. Jain
e-mail: reetu.jain@onmyowntechnology.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_51

ND is primarily detected through methods such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and single proton emission (SPECT) scans. However, the main drawback of the procedures is that they are expensive and overpriced. Nevertheless, the development of artificial intelligence (AI), machine learning (ML), and deep learning (DL) has made it possible to identify ND very quickly and at a lower cost.

Due to its susceptibility to even minor disturbances brought on by ND, speech is one of the numerous factors that exhibits the earliest symptoms in a person with ND. Breathiness, roughness, diminished loudness, and vocal tremor are frequently seen with ND speech impairment [1]. Seventy to ninety percent of those with ND experience speech impairment (SI), and about one-third of patients list SI as one of their primary disease-related impairments [2].

1.1 Motivation and Novelties

According to the reviewed research, a patient with ND may have a digital biomarker for the disease in their speech, which can be used to diagnose the condition. The goal of the study is to develop a smart system that can identify and diagnose ND through speech and voice. ND is marked by tremor, bradykinesia, muscle rigidity, and postural instability. Speech has been demonstrated to be impacted in addition to these distinctive motor indicators. Here, we offer a technique for ND detection that makes use of a patient's voice that has been edited from a raw audio stream. ND prodromal symptoms and unusual speech patterns have been connected. At the early phases of ND detection, our research provides a diagnosing method that is far less invasive than the methods now used. Verbal alterations are the prodromal symptom of ND that has been associated with it. ND is identified in this study based on the patient's pronunciation, pitch, and phase correctness. The combination of a patient voice recording and a machine learning model developed utilising information from the speech frequency features of patients and a database of healthy individuals is crucial. Using a model with a high accuracy rate, the programme will be able to determine whether the user has an ND or not. Because earlier therapy not only improves long-term clinical outcomes but also lessens acute symptoms, the significance of early ND identification cannot be emphasised. Pronunciation accuracy, pitch accuracy, and phase accuracy are used in conjunction with k-nearest neighbour (kNN), and convolutional neural network to accomplish the purpose of the article (CNN).

2 Literature Review

A SI is one among the initial symptoms to occur. To identify PD, speech analysis has been employed in numerous research. They uncovered a vocal disorder called hypokinetic dysarthria, which showed itself as a decrease in prosody, phonational abnormalities, and articulation issues. At mild to advanced stages of the illness, the categorization performances (accuracy rate) using voice analysis ranged from 65 to 99% [3–6]. There are research that concentrates on studying the SI to detect PD early [7, 8]. Recently, PD detections on healthy volunteers, neurologists' early-stage PD subjects, and healthy volunteers at any stage of the disease have all been carried out via telephone networks [9, 10].

Several classification methods have been researched to identify PD using voice analysis. The early studies, which are detailed in, averaged global metrics like the number of pauses, the number of varied words, the standard deviation (SD) of pitch and intensity with low-level disturbances like shimmer, jitter, voice onset time, signal-to-noise ratio, formants, or vowel space area [11]. The authors regularly used feature selection, preserving traits that were statistically significant and eliminating those that were not. Last but not least, classifiers like support vector machines were given a subset of features (SVM) [12, 13], kNN [14], decision trees [15], multilayer perceptrons [16], probabilistic neural networks [17], or Gaussian kernel density minimax classifiers [18].

In the last two decades, MFCCs [19] have appeared in the detection of vocal pathologies [20–22]. Initiated in, the use of MFCCs for PD detection [23]. Since then, MFCCs have been employed in numerous investigations to detect PD [24, 25] or PD monitoring [26].

On MFCC characteristics, a variety of statistical studies and classifiers can be used. For instance, if MFCC dispersion is minimal within classes, often due to a dearth of phonetic variety, one simply needs to consider the MFCC averages (in addition to other features). This is typically true for jobs requiring sustained vowels [27, 28] or when frames that are phonetically similar are chosen [29]. Authors frequently add other statistics to the means, such as the standard deviation, kurtosis (flattening measurement), and skewness, in order to gain a little more information (asymmetry measurement). These features are then added to classifiers like SVMs, multilayer perceptrons, or decision trees.

3 Preliminaries

The Model for Machine Learning On datasets for FY 2014–2018, training and data analysis for a few indicators are performed. Every dataset contains more than 200 financial indicators for a variety of businesses that are normally included in the 10-K files that every publicly traded company submits every year (on average, 4 k stocks are listed in each dataset). This dataset was produced using the Financial Modelling Prep API and the pandas DataReader. Each stock is classified in binary form in the last column, class, where class = 1 for each stock if the PRICE VAR [%] value is positive. Figure 1 shows which stocks a hypothetical trader should BUY at the beginning of the year and SELL for a profit at the end of the year in terms of trading.

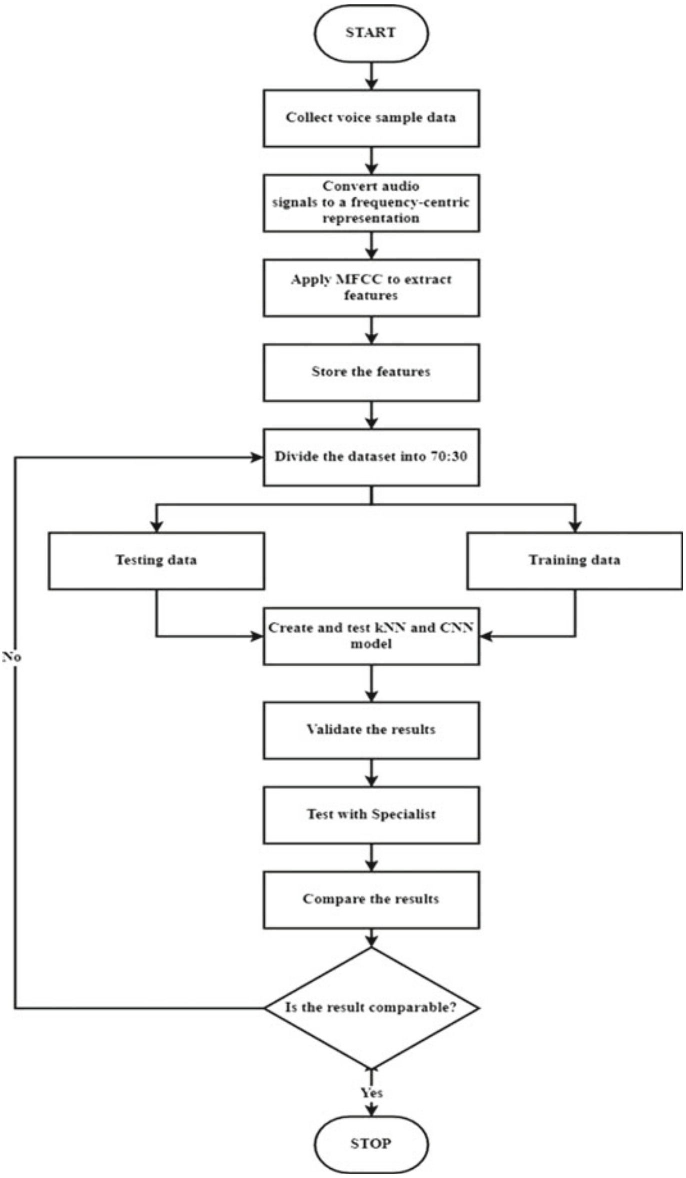


Fig. 1 Troposed methodology

If the PRICE VAR [%] value for any given stock is negative, class = 0. From a trading perspective, a fictitious trader should AVOID BUYING the stocks denoted by a zero since they would lose money if their value decreases.

4 Exploratory Data Analysis

4.1 Data Collection

The suggested approach gathers two distinct datasets from people with paralysis and Parkinson's disease (PD). The PD dataset includes 108 age-matched healthy volunteers (mean age: 60.211.0 years) and 115 patients (mean age: 68.29.2 years). In order to study the paralysis, we gathered 189 samples of normal voices and 552 samples from people who had voice problems, such as adductor spasmodic dysphonia ($n = 30$), unilateral vocal paralysis ($n = 50$), and vocal atrophy ($n = 224$). The 741 samples were split into two sets: 593 samples were used for testing, and 148 samples were used for training. To convert speech to text and text to speech, as well as to recognise the correct pronunciation, we used the speech recognition module from Google in Python to analyse the accuracy of the pronunciation. The type of ND is then determined by the patient's pitch. Patients with paralysis and Parkinson's disease have pitches that range from 45 to 55 dB and 55 to 65 dB, respectively. The category of paralysis can then be found and diagnosed using the phrase accuracy. Then, we made a comparison between the entire patient group and voice samples taken during the emission of vowels and sentences. According to this investigation, speech tasks performed diagnostically in a substantial and comparable manner. The collected dataset is trained and tested using kNN and CNN.

4.2 Mel-Frequency Cepstral Coefficients (MFCC)

The balance sheet consists of two parts, namely assets and liabilities. Both short-term and long-term assets are options under the asset section. Under the section on liabilities, there are also two options: Short-term obligations in comparison with long-term obligations. Examples of short-term assets include cash and cash equivalents, inventories, accounts receivable, securities, etc. Intangible assets that are long-term investments include property, plant, and equipment. Short-term commitments Deposits from customers, dividend and trade account payables, commitments for the near term, and the current amount of long-term debt Loans, delayed income, postponed wages, and other obligations are all considered long-term liabilities. Let us look at each P&G short-term asset individually: thinking about the following parameters currency and its equivalent inventory of cash and short-term investments and average investment in receivables. The cepstral representation of an audio sample is where Mel-frequency cepstral coefficients (MFCC) are derived (a nonlinear "spectrum-of-a-spectrum"). In contrast to the linearly-spaced frequency bands found in the typical spectrum, the Mel-frequency cepstrum (MFC) includes frequency bands that are uniformly spaced on the Mel scale, making it different from the cepstrum in that it better mimics the response of the human hearing system. Frequency warping can improve the way sound is represented. We use the subsequent procedure to obtain the MFCCs. A typical MFCC derivation is as follows:

- Perform a signal's Fourier transform on a windowed extract.
- To transfer the spectrum powers gained in step one onto the Mel scale, use triangle overlapping windows or, alternatively, cosine overlapping windows.
- Calculate the power logs for each of the Mel frequencies.
- Treat the discrete cosine transform of the list of Mel log powers as a signal.
- The amplitudes of the resultant spectrum are the MFCCs.

Procedure of extracting MFCC involves the following phases:

(1) Preemphasis,
(2) Windowing,
(3) Discrete Fourier transform (DFT),
(4) Bank Mel-filter,
(5) Application of log,
(6) IDFT, and
(7) Dynamic features.

The first stage of MFCC extraction is pre-emphasis. Increasing the energy for higher frequency sound during this phase enhances the model's performance. Windowing is the division of the audio into smaller, independent parts from which the features needed to be mapped for classification are retrieved. The signal is then converted from the time domain to the frequency domain using the DFT transform. It is simpler to analyse audio signals in the frequency domain than in the time domain. It has been demonstrated that imitating the human hearing property during the feature extraction step improves the model's performance. Hence, we will use the Mel frequency to convert the real frequency into a frequency that individuals can perceive.

On the basis of where the tongue and other articulators are placed, the vocal cavity selectively amplifies and dampens frequencies. Each sound will be produced using a certain tongue and other articulators. The cepstrum is the logarithm of the signal's magnitude inverted. The first- and second-order derivatives of the characteristics and features are taken into account by the MFCC approach. Derivatives are produced using these coefficients between audio signal samples to assist illustrate how the transition occurs. From each auditory signal sample, the MFCC technique will provide features that will be fed into a model that will determine whether the patient has a SI or not.

4.3 Machine Learning

ML, a branch of AI, is the ability of a machine to mimic intelligent human behaviour [30]. AI systems are used to complete challenging tasks, much to how people approach problems. In this study, the features collected by the MFCC are used to diagnose the patient with ND using kNN and CNN.

k-Nearest Neighbour (kNN): The most fundamental machine learning algorithm, the kNN, is also one of the most effective categorization techniques. The kNN algorithm can compete with the most accurate models because it generates exceptionally accurate predictions [31]. As a result, applications that require high accuracy but do not call for a model that can be read by humans can employ the kNN approach. The accuracy of the projections is influenced by the distance measurement. The accuracy of the projections is influenced by the distance measurement. So, when there is enough domain information, the kNN technique is appropriate. This data supports the selection of a suitable measure. The calculations used to generate the projections are as follows.

Convolutional Neural Network (CNN): Throughout the past 20 years, CNN has been employed for a number of challenging tasks, including pattern identification, voice recognition, and image categorization. A CNN is a deep learning network architecture that gains knowledge directly from data. Following are the procedures used to recognise and categorise the patient as having ND using the features collected by MFCC via CNN:

Step 1: Performing convolution operations
Step 2: Performing the activation operations
Step 3: Performing the pooling operations
Step 4: Layer stacking to reduce error in the model
Step 5: Classifying the outputs
Step 6: Testing the developed model
Step 7: Predicting the outputs.

5 Proposed Methodology

Voice recordings from healthy individuals and ND patients make up the study's dataset. The procedure of sampling is then used to transform the voice recordings into discrete numerical signals. Python uses Fourier Transforms (FT) to represent audio signals in a frequency-centric manner. The process of extracting features, which turns this frequency domain data into a useful feature vector, comes next. Moreover, the best attributes are employed to identify ND. After that, the MFCC arrays are saved in a dataframe using the pandas package. The next stage is to determine whether a person has ND or not using kNN and CNN.

We classified each dataset in the feature array according to its correctness for the pronunciation, pitch, and sentence. We also encoded the data using feature mapping. Following the scaling of the characteristics, the dataset was split into training and testing datasets. A total of 593 samples are used for training, and 148 samples are used for testing from the total amount of data collected. The ND was then diagnosed using kNN and CNN in the following phase.

6 Result and Discussion

6.1 Results

Both healthy individuals and ND patients provided the 743 samples of data. MFCC extracts the features from the samples and divides them into three groups: patients with PD, paralysed patients, and healthy patients. An approximate 70:30 split of the dataset is made between training and testing data. The ML models are tested using the remaining 30% of the data after they have been trained using 70% of the data. The programmes were created in Python 3 and were executed using Jupyter Notebook. Figure 2 displays the receiver operating characteristics (ROC).

The kNN model had accuracy, precision, and recall values of 0.846, 0.859, and 0.898, respectively. 0.88326 square metres cover the ROC curve. 60 out of 71 labels for paralysis, 57 out of 71 for Parkinson's disease, and 68 out of 81 for normal could all be predicted by the CNN model (Fig. 3).

After that, a CNN model is built using the training data. 50 iterations of the model are run. Figure 4 displays the accuracy and loss graph for training, whereas Fig. 5 displays the accuracy and loss graph for validation.

The CNN model's training accuracy rises from 0.6778 in the first epoch to 0.9778 in the 50th, while its training loss declines from 0.6952 in the first epoch to 0.0810 in the 50th. In contrast, the CNN model's testing accuracy increased from 0.6753 in the first epoch to 0.9984 in the 50th, while its training loss reduced from 0.6721 in the first epoch to 0.0234 in the 50th. 62 out of 71 PD labels, 64 out of 71 paralysis labels, and 71 out of 81 normal labels could be predicted by the CNN model (Fig. 6).

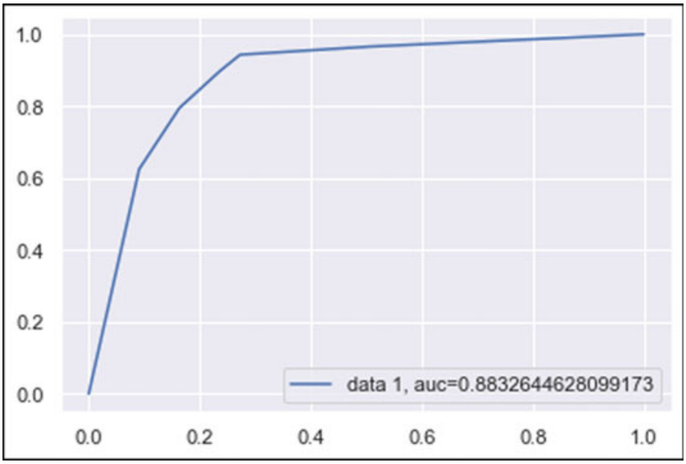


Fig. 2 ROC curve

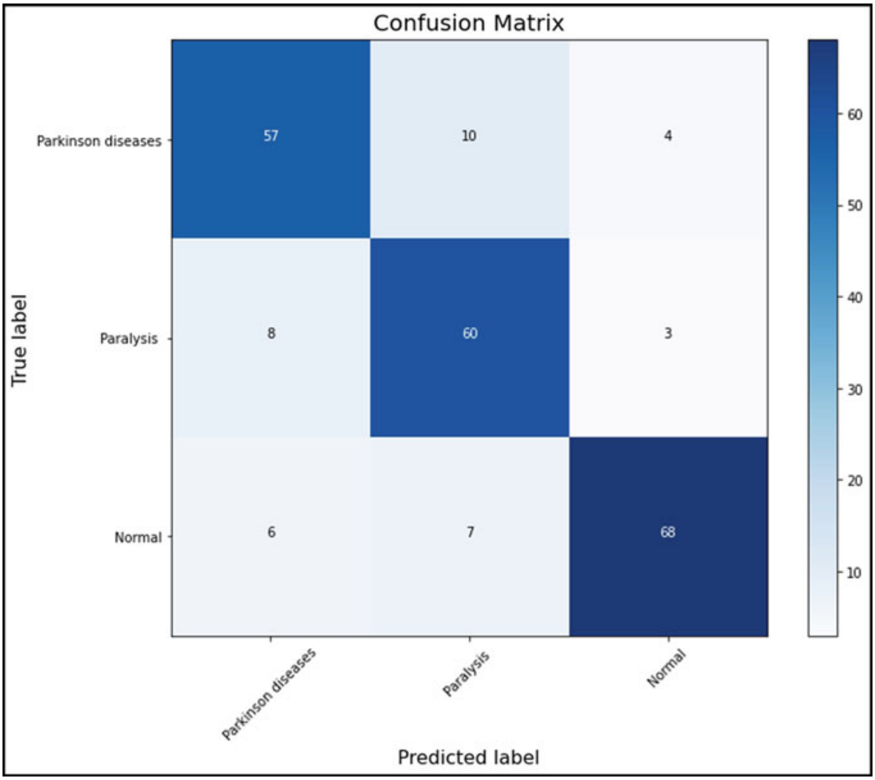


Fig. 3 Heatmap of the confusion matrix

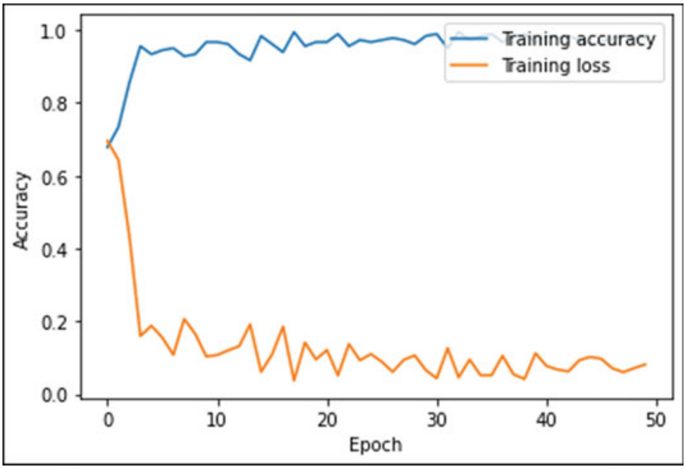


Fig. 4 Training accuracy and loss graph

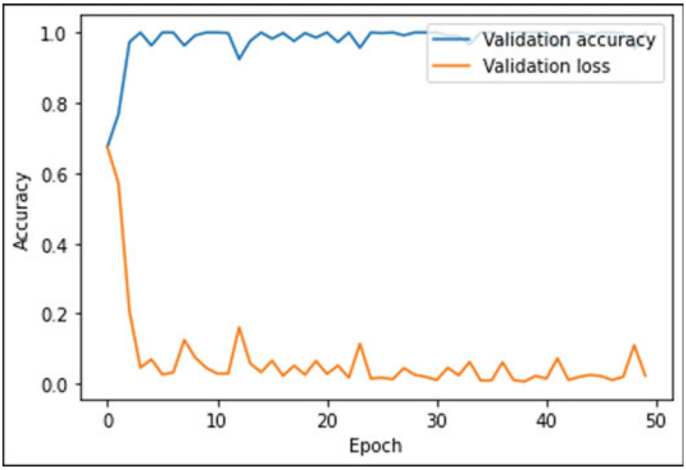


Fig. 5 Validation accuracy and loss graph

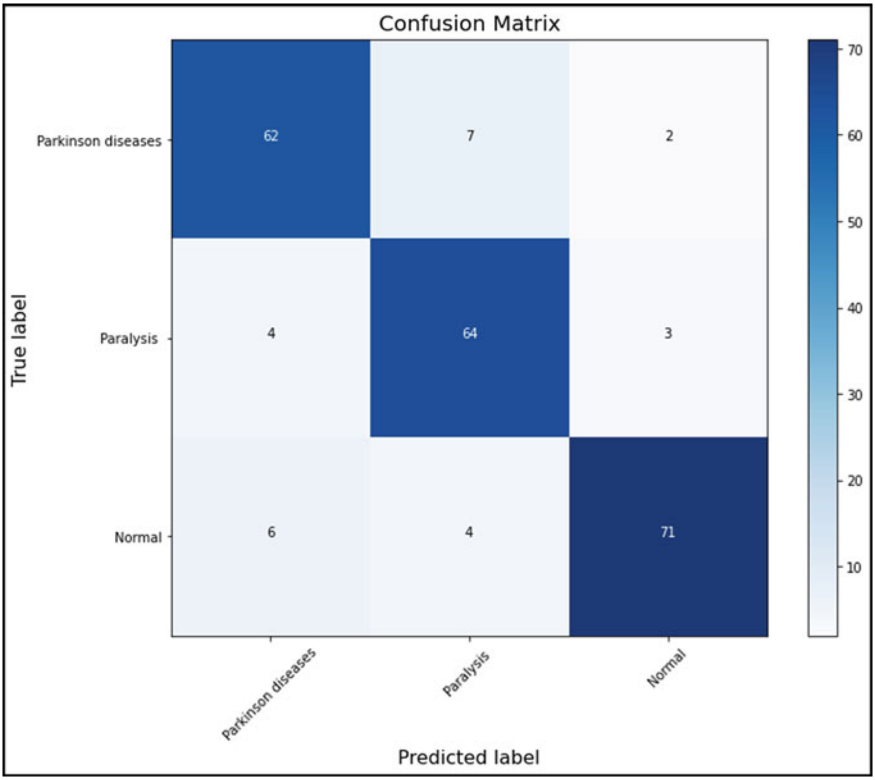


Fig. 6 Heatmap of the confusion matrix

6.2 Validation and Models

The accuracy of experts is compared with the model's output in order to validate it. By listening to the patients' voice records, experts can identify ND. The total accuracy rates for ENT physicians and laryngologists, respectively, were 51.4 and 43.2%. It can be proven that the voice can be utilised as a biomarker for identifying ND in patients by comparing the results from the ML models and the diagnoses made by the professionals. Moreover, integrating ML models will make it possible to recognise and detect the ND category.

7 Conclusion

The study's primary goal is to create a mechanism for identifying and classifying ND in patients. For this reason, we have combined ML models like kNN and CNN with voice as a biomarker. According to previously published studies, one of the major factors that the ND has an effect on is voice. A total of 743 voice samples from PD patients, paralysed people, and healthy individuals were gathered for the study. The MFCC, trained with kNN and CNN and mapped with the three categories, extracts the voice's features. The ratio between the two halves of the dataset is 70:30.The remaining 30% of the data are used to test the model once it has been trained using the 70% of the data. The accuracy, precision, and recall for the kNN model were 0.846, 0.859, and 0.898, respectively. The ROC curve's area under the curve is 0.8832. The CNN model, on the other hand, displayed training and testing accuracy of 0.9778 and 0.9984, respectively. The CNN model's training and testing losses are 0.0810 and 0.0234, respectively. The ML model's accuracy is compared against the experts' diagnostic accuracy in order to validate the model. The two laryngologists' respective total accuracy ratings were 60.1% and 56.1%. It would be more accurate to run this software on a person more than once, which would aid patients in determining what stage of PD they have and what should be their next step in treatment, such as therapy, medicine, or both. The voice can be employed as a biomarker for diagnosing ND in patients, according to the general discussion. Moreover, integrating ML models will make it possible to recognise and detect the ND category.

8 Ethical Standards

The authors would like to declare that the dataset collected for the study is done under the supervision of the specialists. The patient is not pressurised in any form either mentally or physically for collecting data.

References

1. Potter NL (2011) Voice disorders in children with classic galactosemia. J Inherit Metab Dis 34(2):377–385
2. Stachler RJ, Francis DO, Schwartz SR, Damask CC, Digoy GP, Krouse HJ, McCoy SJ, OuelletteDR, Patel RR, Reavis CCW, Smith LJ, Nnacheta LC (2018) Clinical practice guideline: hoarseness (dysphonia)(update). Otolaryngol–Head Neck Surg 158(1_suppl):S1–S42
3. Guo PF, Bhattacharya P, Kharma N (2010) Advances in detecting Parkinson's disease. In: International conference on medical biometrics. Springer, Berlin, Heidelberg, pp 306–314
4. Gómez-Vilda P, Mekyska J, Ferrández JM, Palacios-Alonso D, Gómez-Rodellar A, Rodellar-Biarge V, Rektorova I (2017) Parkinson disease detection from speech articulation neuromechanics. Front Neuroinformatics 11:56
5. Ali L, Zhu C, Zhang Z, Liu Y (2019) Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network. IEEE J Transl Eng Health Med 7:1–10
6. Avuçlu E, Elen A (2020) Evaluation of train and test performance of machine learning algorithms and Parkinson diagnosis with statistical measurements. Med Biol Eng Compu 58(11):2775–2788
7. Rusz J, Hlavnička J, Tykalová T, Bušková J, Ulmanová O, Růžička E, Šonka K (2016) Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder. Sleep Med 19:141–147
8. Orozco-Arroyave JR, Hönig F, Arias-Londoño JD, Vargas-Bonilla JF, Daqrouq K, Skodda S, Rusz J, Nöth E (2016) Automatic detection of Parkinson's disease in running speech spoken in three different languages. J Acoust Soc Am 139(1):481–500
9. Jeancolas L, Mangone G, Corvol JC, Vidailhet M, Lehéricy S, Benkelfat BE, Benali H, Petrovska-Delacrétaz D (2019) Comparison of telephone recordings and professional microphone recordings for early detection of Parkinson's disease, using mel-frequency cepstral coefficients with Gaussian mixture models. In: INTERSPEECH 2019: 20th annual conference of the international speech communication association. International Speech Communication Association (ISCA), pp 3033–3037
10. Arora S, Baghai-Ravary L, Tsanas A (2019) Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality voice. J Acoust Soc Am 145(5):2871–2884
11. Jeancolas L, Petrovska-Delacrétaz D, Lehéricy S, Benali H, Benkelfat BE (2016) L'analyse de la voix comme outil de diagnostic précoce de la maladie de Parkinson: état de l'art. In: CORESA, vol 2016, p 18e
12. Rusz J, Bonnet C, Klempíř J, Tykalová T, Baborová E, Novotný M, Rulseh A, Růžička E (2015) Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. J Neurol 262(4):992–1001
13. Erdogdu Sakar B, Serbes G, Sakar CO (2017) Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease. PLoS ONE 12(8):e0182428
14. Sakar BE, Isenkul ME, Sakar CO, Sertbas A, Gurgen F, Delil S, Apaydin H, Kursun O (2013) Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. IEEE J Biomed Health Inf 17(4):828–834
15. Mucha J, Galaz Z, Mekyska J, Kiska T, Zvoncak V, Smekal Z, Eliasova I, Mrackova M, Kostalova M, Rektorova I, Alonso-Hernandez, JB (2017) Identification of hypokinetic dysarthria using acoustic analysis of poem recitation. In: 2017 40th International conference on telecommunications and signal processing (TSP). IEEE, pp 739–742
16. Gil D, Johnson M (2009) Diagnosing Parkinson by using artificial neural networks and support vector machines. Glob J Comput Sci Technol 9:63–71
17. Ene M (2008) Neural network-based approach to discriminate healthy people from those with Parkinson's disease. Ann Univ Craiova Math Comput Sci 35:112–116

18. Rusz J, Cmejla R, Tykalova T, Ruzickova H, Klempir J, Majerova V et al (2013) Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task. J Acoust Soc Am 134:2171–2181. https://doi.org/10.1121/1.4816541
19. Frédéric B, Jean-François B, Corinne FD, Guillaume G, Ivan MC, Syl-vain M, Dijana PD, Reynolds DA (2004) A tutorial on text-independent speaker verification. EURASIP J Appl Sig Process 430–451
20. Dibazar AA, Narayanan S, Berger TW (2002) Feature analysis for automatic detection of pathological speech. In: Proceedings of the second joint 24th annual conference and the annual fall meeting of the biomedical engineering society, engineering in medicine and biology, vol 1, no Houston, TX, 182–183. https://doi.org/10.1109/IEMBS.2002.1134447
21. Godino-Llorente J, Gómez-Vilda P (2004) Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. IEEE Trans Biomed Eng 51:380–384. https://doi.org/10.1109/TBME.2003.820386
22. Malyska N, Quatieri TF, Sturim D (2005) Automatic dysphonia recognition using biologically-inspired amplitude-modulation features. In: IEEE international conference on, acoustics, speech, and signal processing, 2005, proceedings (ICASSP'05), vol 1, Philadelphia, PA, pp 1–873. https://doi.org/10.1109/ICASSP.2005.1415253
23. Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO (2012) Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. IEEE Trans Biomed Eng 59:1264–1271. https://doi.org/10.1109/TBME.2012.2183367
24. Drissi TB, Zayrit S, Nsiri B, Ammoummou A (2019) Diagnosis of Parkinson's disease based on wavelet transform and mel frequency cepstral coefficients. Int J Adv Comput Sci Appl 10:125–132. https://doi.org/10.14569/IJACSA.2019.0100315
25. Fang H, Gong C, Zhang C, Sui Y, Li L (2020) Parkinsonian Chinese speech analysis towards automatic classification of Parkinson's disease. In: Proceedings of machine learning research 2020, p 12
26. Orozco-Arroyave JR, Vàsquez-Correa JC, Honig F, Arias-Londono JD, Vargas-Bonilla JF, Skodda S et al (2016b) Towards an automatic monitoring of the neurological state of Parkinson's patients from speech. In: Conference: 41st ICASSP (Shanghai), pp 6490–6494. https://doi.org/10.1109/ICASSP.2016.7472927
27. Benba A, Jilbab A, Hammouch A (2016) Discriminating between patients with Parkinson's and neurological diseases using cepstral analysis. IEEE Trans Neural Syst Rehabil Eng 24(10):1100–1108
28. Hemmerling D, Orozco-Arroyave JR, Skalski A, Gajda J, Nöth E (2016) Automatic detection of Parkinson's disease based on modulated vowels. In: INTERSPEECH (San Francisco, CA), pp 1190–1194. https://doi.org/10.21437/Interspeech.2016-1062
29. Orozco-Arroyave JR, Hönig F, Arias-Londoño JD, Vargas-Bonilla JF, Daqrouq K, Skodda S et al (2016) Automatic detection of Parkinson's disease in running speech spoken in three different languages. J Acoust Soc Am 139:481–500. https://doi.org/10.1121/1.4939739
30. Shan T, Tay FR, Gu L (2021) Application of artificial intelligence in dentistry. J Dent Res 100(3):232–244
31. Guo G, Wang H, Bell D, Bi Y, Greer K (2003) KNN model-based approach in classification. In: OTM Confederated international conferences on the move to meaningful internet systems. Springer, Berlin, Heidelberg, pp 986–996

Designing an Electronic Ticketing System for Local Commuter Transportation

Yash Ayare, Atharva Dhoble Hemakshi Lohiya, Shrawani Gulhane, and Santosh Kumar

Abstract The rapid urbanization witnessed in cities has resulted in an increased reliance on public buses for transportation. Consequently, it becomes imperative to modernize the ticket booking system for public transportation within a specific city. This paper proposes the development of an Online Bus Ticket Reservation System, a web-based application designed to facilitate various tasks such as checking bus schedules, ticket availability, purchasing bus tickets, and making online payments for bus fares. Traditionally, bus conductors have had to manually collect fares from each passenger on board and provide them with physical tickets. However, this proposed system aims to alleviate the challenges faced by bus conductors by streamlining fare collection and introducing a transparent financial process. By utilizing a database to display and store financial transactions, the system enhances transparency, reduces risks associated with physical fare collection, and simplifies financial management for public transportation.

Keywords Online ticketing system · Machine learning · Local transport

1 Introduction

The increasing urbanization and population growth in cities have brought about a significant surge in the demand for efficient and reliable public transportation systems [12]. Local commuter transportation plays a vital role in catering to the daily travel needs of a large number of residents, commuters, and visitors. However, the conventional paper-based ticketing systems employed by many local commuter transportation services often suffer from numerous drawbacks, including ticket fraud, revenue leakage, and inconvenience to passengers.

Y. Ayare (✉) · A. D. H. Lohiya · S. Gulhane · S. Kumar
Department of AI&DS, Vishwakarma Institute of Information Technology, Pune, India
e-mail: yash.22011123@viit.ac.in

A. D. H. Lohiya
e-mail: hemakshi.22011041@viit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_52

To address these challenges and enhance the efficiency of local commuter transportation, the design and implementation of an electronic ticketing system have become imperative. An electronic ticketing system leverages modern technology to streamline ticketing operations, improve passenger experience, and ensure accurate revenue management.

This paper aims to present the design and development of an electronic ticketing system tailored specifically for local commuter transportation. The system will be web-based, allowing passengers to conveniently access and utilize its features through various devices such as smartphones, tablets, or computers. By adopting an electronic ticketing system, local commuter transportation services can transition from manual, paper-based processes to a technologically advanced and efficient approach. This paper aims to accomplish the following objectives.

1. To design a user-friendly interface for passengers to access the electronic ticketing system.
2. Developing a secure and robust platform for ticket purchasing, fare payment, and ticket validation.
3. Implementing a centralized database to store ticketing data and facilitate seamless transaction processing.
4. To integrate the electronic ticketing system [11] with existing local commuter transportation infrastructure, such as ticketing machines and fare gates.
5. Evaluating the effectiveness and efficiency of the electronic ticketing system through performance analysis and user feedback.

Through the implementation of an electronic ticketing system, local commuter transportation services can bring forth a multitude of advantages for passengers, operators, and authorities alike. Passengers will enjoy a heightened level of convenience and efficiency in the ticketing process, leading to reduced waiting times and an enhanced overall travel experience. Operators stand to gain from improved revenue management, a decrease in ticket fraud instances, and increased operational efficiency. Additionally, authorities can harness the system's data analytics capabilities to extract valuable insights regarding passenger patterns, allowing them to optimize routes and effectively plan for future transportation infrastructure.

1.1 Scope of Online Bus Ticket Reservation

The specific goals of developing the Online Bus Ticket Reservation System are as follows:

Improved Ticket Reservation Process: The primary goal of the system is to enhance the ticket reservation process for passengers by providing them with a convenient and user-friendly platform to book bus tickets online. It aims to simplify the ticket booking process, eliminate the need for physical ticket counters, and enable passengers to reserve their seats in advance.

Streamlined Operations: The system aims to streamline the operations of bus conductors by automating the ticket reservation and management process.

Real-Time Seat Availability: The system provides real-time information about seat availability on buses, allowing passengers to make informed decisions while booking their tickets.

Efficient Payment Management: It includes a secure and integrated payment gateway, enabling passengers to make online payments for their ticket reservations. This eliminates to handle cash transactions, reduces the risk of errors or discrepancies, and ensures a seamless payment process.

Passenger Convenience and Satisfaction: By offering an online ticket reservation system, the goal is to enhance passenger convenience and satisfaction. Passengers can easily access the system from anywhere, choose their preferred seats, and receive e-tickets on their mobile devices.

By addressing these specific goals, the Online Bus Ticket Reservation System aims to overcome the challenges faced by bus conductors, such as manual ticketing processes, limited seat allocation visibility, cash handling, and inefficient reservation management.

The forthcoming sections of this paper will thoroughly examine the intricate design, execution, and assessment of the electronic ticketing system, elucidating its diverse elements, functionalities, and benefits. Furthermore, we will delve into the technical aspects, security measures, and integration complexities associated with the deployment of this system in the context of local commuter transportation.

2 Literature Review

The electronic ticketing system [5, 8] has gained immense popularity in recent times, owing to its multitude of advantages over traditional paper-based ticketing methods. It has brought about a revolutionary change in the way ticketing and reservations are managed. By shifting the entire ticketing system online, it significantly reduces the reliance on paper and offers users the convenience of purchasing tickets from anywhere in the world. Moreover, it streamlines the ticket booking process, reducing the workload of staff members. In this section, we will present an overview of the electronic ticketing system, exploring its various types, benefits, and challenges. Various forms of electronic ticketing systems exist.

- Smart Card Ticketing System [1, 9]: This system utilizes a smart card to store ticket data and validate the information contained within, ensuring its authenticity. One prominent example of such implementation is the EZ-Link card in Singapore, widely utilized for public transportation purposes. Similarly, in Japan, the Suica smart card is employed for seamless travel across various public transportation modes. In Germany, the VDV-KA smart card is extensively used for day-to-day public transportation requirements.

- Barcode Ticketing System [10]: Our primary focus is to develop a model for the local bus transportation ticket booking system, which serves as the core of our research. This system employs barcode technology for ticket authentication and validation. An example of such a system is the Clipper card barcode ticketing system, utilized for public transport in the San Francisco Bay Area, USA. In Australia, the Opal Card barcode system is implemented for customer convenience in public transportation across New South Wales. Furthermore, in India, the Indian Railways e-ticketing system solely relies on this type of ticketing mechanism.
- Mobile Ticketing System [2, 3]: This system allows user to purchase and receive tickets via their mobile application. This kind of system is increasing in todays world. In France, SNCF mobile app is used to purchase train tickets and store them on their mobile devices for scanning at train station. In India, mobile ticketing systems are used for transportation for Mumbai's suburban railway system. In China, WeChat mobile app allows users to buy and store tickets for upcoming events on their mobile devices.
- RFID Ticketing system [4]: This system uses radio-frequency identification technology to validate tickets. In South Korea, T-Money card uses this RFID technology. This card is widely used for public transportation over there. In UK, Oyster card which based on this system is used in London and other several parts of UK for ticketing.

2.1 Challenges of Local Commuter Transportation System

– Tackling Technical issues: This issue includes system crash, connectivity problem, or hardware malfunctions. So electronic ticketing system must be prone to these kinds of issues. These issues can lead to delays and inconvenience to customer which eventually leads to losing revenue.
– Accessibility: To access the electronic ticketing system user needs to have access of mobile phone, desktop, or any other device. This main barrier for users who do not have access any of these devices.
– Resistance to change/upgrade: Some users may go with the traditional ticketing system instead of this system, but to attract these people system must be user-friendly easy to use and most importantly convenient and error-free. This can limit the adaptation of electronic ticketing system.
– Security concerns: This system must be prone to security breaches such as hacking or identity theft. This can lead to threat to sensitive information about customer and reduces customer trust which leads to financial loss for the ticketing system company.

2.2 Case Studies of Electronic Ticketing Systems

Singapore's EZ-link Card System Singapore's EZ-Link card system is among the greatest illustrations of an electronic ticketing system for local commuter transit [6]. With this system, users may pay for both their bus and rail rides with a single contactless smart card. Efficiency, convenience, and cost savings are just a few advantages of the EZ-Link card system. The technology has reduced the demand for paper tickets, improved the ticketing process, and reduced the environmental impact of the transport sector. The system's real-time information on commuter behavior allows transportation authorities to optimize pricing and improve the caliber of their services.

London's Oyster Card System The Oyster card system in London is another successful example of an electronic ticketing system for local commuter travel. The system allows commuters to pay for their bus and train rides using a contactless smart card or a smartphone app. Numerous advantages of the Oyster card system include its effectiveness, practicality, and compatibility with different modes of transit. The technique has improved ticketing accuracy and speed while reducing the requirement for paper tickets. The system's real-time information on commuter activities allows transport authorities to confidently optimize services and determine rates.

Japan's Suica Card System A popular electronic ticketing system for local commuter transit is Japan's Suica card system. The system enables travelers to use a contactless smart card to pay for their bus and rail rides. Numerous benefits, including speed, cost savings, and convenience, come with the Suica card system. The technology has decreased the demand for paper tickets, increasing the effectiveness of the ticketing procedure and lowering the transportation industry's environmental impact. As a result of the system's real-time data on commuter behavior, transport authorities may optimize pricing and raise the standard of their services.

3 System Architecture

The system architecture for designing an electronic ticketing system [7] for local commuter transportation typically involves various components and their interactions. Here's an example of a system architecture for such a system (Fig. 1):

In this architecture, the system consists of the following components:

Passengers: The passengers interact with the system through a web or mobile interface to purchase tickets, check schedules, and make reservations.

Ticketing Machines: These machines are physical devices located at stations or vehicles. They are used to issue tickets and validate them when passengers board the vehicles.

Web/Mobile Interface: This component provides the user interface for passengers to interact with the system. It allows them to view schedules, select tickets, make payments, and receive digital tickets.

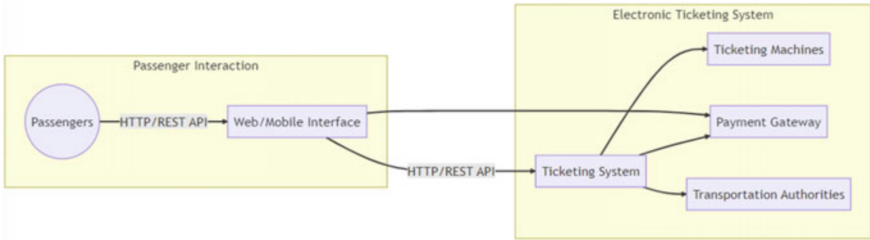


Fig. 1 System architecture of electronic ticketing system

Ticketing System: This is the core component of the electronic ticketing system. It handles ticket purchases and validations and manages ticket-related data such as ticket types, passenger information, and payment records. It exposes an HTTP/REST API for communication with the web/mobile interface and ticketing machines.

Payment Gateway: External payment systems or gateways handle online transactions for ticket purchases. This component interfaces with the ticketing system to process payment information securely.

Transportation Authorities: These are the governing bodies responsible for overseeing and managing the local commuter transportation system. They interact with the ticketing system to access data, generate reports, and perform administrative tasks.

The interactions between these components are as follows:

Passengers interact with the web/mobile interface to browse schedules, select tickets, and make payments. The interface communicates with the ticketing system via an HTTP/REST API.

The ticketing system processes the passenger's requests, handles ticket purchases, and stores ticket data in a ticket database. It also communicates with the payment gateway to process payments securely.

Ticketing machines interact with the ticketing system to validate tickets. They communicate with the system to verify the authenticity of the ticket when a passenger boards a vehicle.

The ticketing system interacts with transportation authorities to provide access to data, generate reports, and perform administrative tasks related to the electronic ticketing system.

This system architecture provides a high-level overview of the components and their interactions in designing an electronic ticketing system for local commuter transportation. It can be further expanded and refined based on specific requirements and additional functionalities.

3.1 Mathematical Model

Designing a comprehensive mathematical model for an electronic ticketing system for local commuter transportation involves considering various factors and components of the system. Here's an example of a mathematical model that captures some key elements:

Ticket Demand Model
Let $D(t)$ represent the ticket demand at time t. D_0 be the initial demand. λ be the demand rate or arrival rate of passengers. Then, the ticket demand can be modeled using a Poisson process:

$$D(t) = D_0 + \lambda t$$

Ticket Sales Model
Let $S(t)$ represent the ticket sales at time t. $R(t)$ be the ticket reservation rate. $V(t)$ be the ticket validation rate. Then, the ticket sales can be calculated as:

$$S(t) = R(t) + V(t)$$

Revenue Model
Let $R(t)$ represent the revenue generated at time t. P be the ticket price. Then, the revenue can be calculated as:

$$R(t) = P \cdot S(t)$$

Queueing Model
Let $Q(t)$ represent the number of passengers in the ticketing queue at time t. $A(t)$ be the passenger arrival rate. $D(t)$ be the ticket departure rate. Then, the queue length can be modeled using a queueing system:

$$Q(t) = A(t) - D(t)$$

Service Time Model
Let ST be the average service time per passenger. Then, the ticket departure rate can be calculated as:

$$D(t) = \frac{Q(t)}{ST}$$

Resource Allocation Model
Let N be the number of ticketing agents or machines. Let C be the capacity of each ticketing agent or machine. Then, the total service capacity can be calculated as:

$$Capacity = N \cdot C$$

System Performance Model
Let $P(t)$ represent the system performance or efficiency at time t. Let $T(t)$ be the total time for ticket processing. Then, the system performance can be measured as:

$$P(t) = \frac{S(t)}{T(t)}$$

The mathematical models capture some aspects of an electronic ticketing system for local commuter transportation. The actual mathematical model may include additional factors such as ticket cancelation, refunds, system reliability, and passenger behavior. It's important to analyze the specific requirements and characteristics of the system to develop a more comprehensive and accurate mathematical model.

4 Methodology

The methodology for designing an electronic ticketing system for local commuter transportation involves a structured approach to create an efficient and user-friendly ticketing solution. It outlines the systematic process followed to develop the electronic ticketing system, ensuring its effectiveness, efficiency, and alignment with the specific requirements of the local commuter transportation context.

4.1 Research Design and Approach

There are mainly two approaches used to form research design, first one quantitative and second is qualitative. These two approaches are used to assess the effectiveness and scalability of online local bus ticket booking system. In quantitative approach, it includes collection and study of numerical data such as number of tickets sold each day, view on web pages and bounce rates, number of successful transactions taking place every day. Using different statistical analysis techniques to identify trend patterns in the data that has been generated through system. It includes peak usage time and the most popular destination routes that most people are traveling every day. The result of this approach is to find improvement in the system which will be again tested in qualitative approach phase. In qualitative approach, it includes directly contacting or interviewing user about their experience with this system. Also gaining valuable insights into their experience with this system. Also testing will be part of it to identify any issue that the user is facing with the system. The result of this approach will be used to identify areas improvements in system design for better user satisfaction and make system more user-friendly to decrease user issues.

4.2 Data Collection Methods and Sources

User surveys: These surveys will be in-personal or through online mode to get an idea about user satisfaction regarding the online local bus ticket booking system. Also, assessing the area of improvements and suggestions through these surveys to make system more user-friendly and reduce issues within the system user are facing.

Usability testing: In this method, we will be assessing the user experience with the system as they use the system and their valuable feedback will be recorded for further need of improvement in the system. This data collection methods comes under the qualitative approach.

System logs: This method will be assessing and gather the data on user behavior such as number of tickets purchased, number of successful transactions rates, page views, bounce rates and ticket sales, number of buses traveling through these routes, their arrival and endpoint arrival time. This comes under quantitative approach.

Monitoring Social Method Platform: It will be used for monitoring feedback from user within system through social media post comments, likes, and all reviews, comments will be monitored to get idea about common themes and areas of improvements.

Market Research survey: Industry reports and academic publications will be reviewed to gain additional idea on local bus transportation industry. These surveys will be best practice for ticketing system and user behavior trends.

4.3 Data Analysis Techniques

Comparative analysis: It will be used to compare system's performance according to industry benchmarks and assesses them to get about areas of improvements Sentimental analysis: It will be based on content analysis. It will be used to analyze qualitative data such as user's feedback from surveys and monitoring social media platform. This data is coded and categorized into theme, patterns, clusters to identify system issues and improve it.

Descriptive statistics: It will be using different descriptive statistical techniques like mean, mode, median, and standard deviation. These techniques will be used to summarize the numerical data such as number of user visit web-page, number of tickets purchase (ticket sales), number of successful transactions, bouncing rates, number of buses traveling in local routes.

Inferential statistics: This statistics technique will use techniques like t-test and ANOVA. Regression analysis will be used to identify trends, patterns, and relationship between data points such as user usage time and number of tickets purchase, impact of user demographics on system usage system design and development process requirement acquiring. This step involves identifying needs and expectation of users, and other relevant authorities. These all requirements can be acquired using surveys through online or offline mode, interviews of users, and other methods.

Analysis of system: After completing step one of acquiring requirements, next step is to analyze the system. It aims to have system goals, getting its functional and non-functional requirements and with the same time identifying system constraints and systems limitations.

System design: After the above two steps, now there is the developing system design stage. In this stage, we make system architecture, design UI/UX, and design systems components with modules. Modules include login/registration, entering starting and destination point, making payment through UPI, etc., generating QR code of ticket to scan after successful payment for ticket fare.

Implementation: After finalizing the system design, the implementation stage takes place. In this stage coding of system, components and modules take place, then testing code, tracking and solving errors and bugs and integrating all the modules with each other for better user interface interaction.

System testing: After implementation stage is done, system testing stage takes place which ensures system satisfies requirements and all functions. This stage includes performance and usability testing. Deployment: After system testing stage, now in deployment stage system gets deploy in production environment. This includes installing system on servers, configuring system settings, and making system available for user to use and uses every function in the system.

Support and maintenance: After successfully deployment of system on servers, now maintenance of servers and support process begins. This involves monitoring system errors bugs, system updates and enhancement, and providing support 24/7 to user of online local bus ticket booking system.

5 System Requirements and Design

As for system requirements, the functional requirements are that the application should load within 3 s of time period, email verification mail should be sent to the user, mobile number verification OTP should be sent to user. The non-functional requirements that include the email verification must be done within time span of 10 min, the application should be able to handle traffic of 500 users at a time, the registered complaints must not take more than 12 h of time, the support system should process the requests of the user within a time period of 24 h.

The use case diagram depicted in Fig. 2 illustrates the interactions between various actors, including users and external systems, and the electronic ticketing system for local commuter transportation. This diagram serves the purpose of identifying the system's functionalities and features as perceived by the users. It assists in recognizing the primary characteristics and behaviors of the system as perceived by users, thereby ensuring that the system aligns with the passengers' and transportation authority's requirements and expectations. User has use cases like login/signup, view buses, select starting point, enter destination point, select number of tickets, type of bus, bus availability ,to pay for the ticket in order to generate ticket. Booking(admin)

Fig. 2 Use case of system

staff has use cases which includes updating system, provide list of available buses, verify the payment, reserve ticket of user, providing all necessary info.

The depicted activity diagram in Fig. 3 showcases the progression of activities and actions within the electronic ticketing system for local commuter transportation. It provides a visual representation of the sequential and parallel activities encompassed in the ticketing process. Sequential activities includes from login/sign-up to verify details, if verified then user can view buses, set it active location, then select destination, after that ticket will be displayed to user, then user have to pay for ticket in order to generate QR code for own ticket, and at last after payment is done successfully user can view own ticket and travel with it easily.

A data flow diagram (DFD) for designing an electronic ticketing system for local commuter transportation illustrates the flow of data within the system and between its various components. It helps to understand how information is processed and transferred between different entities involved in the ticketing process. Here it is a discussion of the key components and data flows in the DFD:

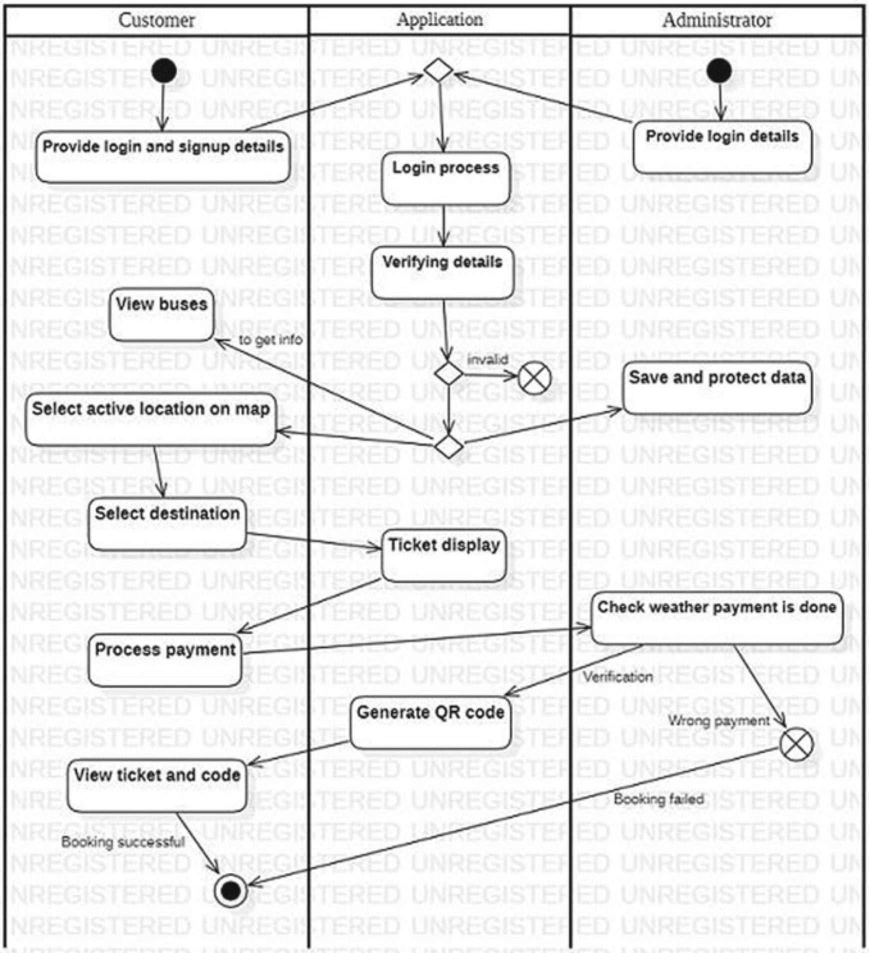


Fig. 3 Activity diagram of system

External Entities

- Passengers: Passengers interact with the system to purchase tickets, check schedules, and make reservations.
- Ticketing Machines: These machines are used to issue tickets and validate them when passengers board the vehicles.
- Payment Gateway: External payment systems or gateways handle online transactions for ticket purchases.
- Transportation Authorities: The governing bodies responsible for overseeing and managing the local commuter transportation system.

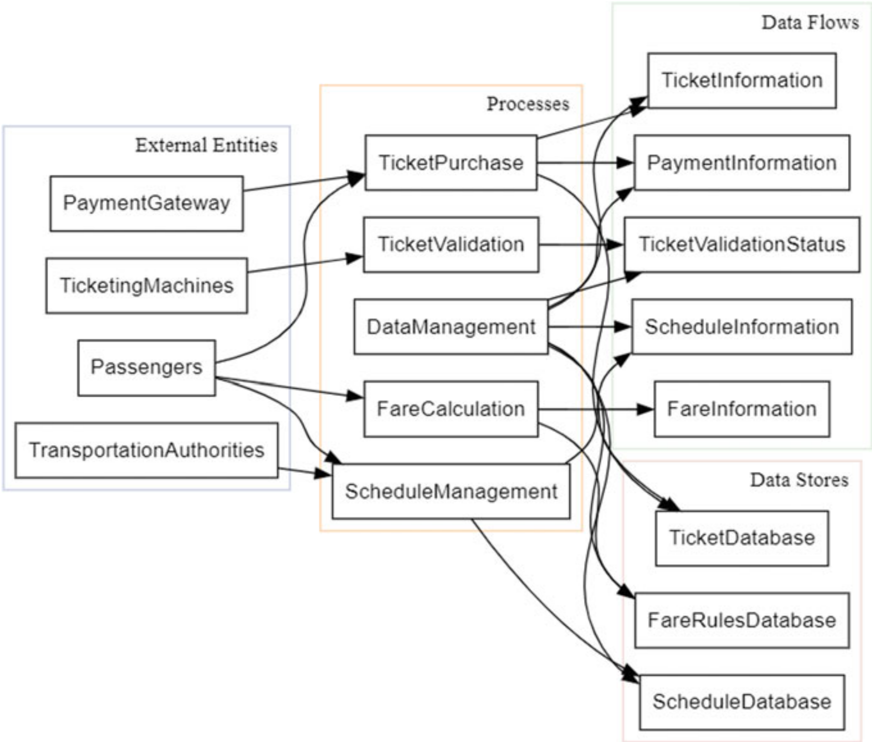


Fig. 4 Data flow diagram of system

Processes

– Ticket Purchase: This process involves a passenger selecting the desired ticket, providing payment information, and generating the ticket.
– Ticket Validation: When a passenger boards a vehicle, the ticket is validated using a ticketing machine to ensure its authenticity.
– Schedule Management: This process manages the schedules of different routes and vehicles, ensuring accurate information is available to passengers.
– Fare Calculation: Fare calculation process determines the cost of the ticket based on various factors like distance, fare rules, and discounts.
– Data Management: This process is responsible for storing and managing the data related to tickets, schedules, passenger information, and payments (Fig. 4).

The data flow diagram provides a visual representation of the flow of data within the electronic ticketing system for local commuter transportation. It helps in understanding the interactions between different entities, processes, and data stores involved in the ticketing process, ensuring smooth and efficient ticket transactions for passengers and accurate record-keeping for transportation authorities.

6 Results and Discussion

The evaluation of the design system is a thorough evaluation of the benefits of several system designs to select the best option which can be considering several options like functionality, user experience, performance, scalability, maintenance, and support. In the analysis of the system's performance, there are many factors for considering the usability. Analyzing the response for the website that the booking system is working on, error rate of the overall system, navigation and the basic layout of the system, compatibility for the users. While comparing the system with other electronic ticketing systems, our system proposes a plan which allows the commuters to book the ticket for the local bus which is currently unavailable for now. A user-friendly system is for booking a ticket online and securing his seat with ease.

6.1 Advantages

- Convenience: Electronic ticketing system is used to purchase tickets online through mobile/Web-based apps or at self-service. This eliminates the need to wait in long queue to purchase ticket for transportation. Saving time for both staff and user.
- Convenience: Electronic ticketing system is used to purchase tickets online through mobile/Web-based apps or at self-service. This eliminates the need to wait in long queue to purchase ticket for transportation. Saving time for both staff and user.
- Cost saving: This electronic system reduces need for physical tickets which uses loads of paper to print and distribute. This system saves money for ticketing company and reduce environmental impact of ticketing.
- Cashless purchasing: This system helps to make purchasing of ticket through UPI/cashless way. So there no need to handle cash. This system breaks the barrier of cash to cashless in ticketing system.
- Efficiency: This electronic ticketing system enables quick scanning and validation of tickets, reduces need for manual ticket checking, and reduces ticket fraud. It saves time for both staff and customer.

Limitations: Since there are many areas, some listed with multiple names, there is possibility that user is unable to find his/her desired destination if they don't enter a specific stop. Growing competition can be one of the limitations since anyone can create their own system and not every platform is equal.

7 Conclusion

In conclusion, our electronic ticketing system offers numerous benefits that enhance the overall ticket processing flow and provide greater convenience to passengers. By enabling customers to easily access their tickets on their devices through a unique

QR code, the system streamlines the ticketing process and offers flexibility in city travel. Moreover, the online ticket booking system contributes to environmental sustainability by eliminating paper waste and supporting a cleaner planet.

The system's user-friendly approach ensures a comfortable and hassle-free experience for passengers, eliminating concerns about carrying or misplacing physical tickets. This not only enhances customer satisfaction but also reduces the workload of employees, minimizes reservation time, and increases overall efficiency. Additionally, the system mitigates the risk of scams or misbehavior between conductors and passengers, promoting a safer and more reliable public transportation experience.

An added advantage of the application is its ability to automatically update records in secure files, alleviating employees' concerns about data security. Importantly, our system does not aim to exclude walk-in passengers from the local bus system, ensuring accessibility for all while still providing the benefits of online ticketing.

In summary, our user-friendly ticket booking system improves the travel experience, increases employment opportunities, and simplifies the conductor's work. By combining efficiency, convenience, and security, our system serves as a valuable solution for anyone seeking a stress-free and reliable mode of transportation.

To make the system even more accessible and user-friendly, we aim to introduce additional functionalities that align with changing needs and technological advancements. This includes exploring the possibility of integrating chat or voice assistants, specifically designed to assist elderly individuals in booking their tickets effortlessly. By doing so, we strive to ensure that our system caters to the needs of a diverse range of users, promoting inclusivity, and convenience for all. Continuously improving our system's accessibility and incorporating user feedback will remain key focal points as we move forward.

Acknowledgements We would like to express our sincere gratitude and appreciation to the Department of AI&DS (Artificial Intelligence and Data Science) for their invaluable support and contribution to the development of this paper on "Designing an Electronic Ticketing System for Local Commuter Transportation." We would also like to extend our gratitude to the faculty members and researchers who have provided valuable insights and feedback during the process.

References

1. Huang Y, Huang Y (2021) Design and implementation of an electronic ticketing system for bus transportation based on cloud computing. In: Proceedings of the 2021 10th international conference on computer science and network technology (ICCSNT), pp 31–34
2. Mukhtar HM, Tanwir S (2020) Design and development of a mobile-based electronic ticketing system for public transport. IEEE Access 8:209383–209393
3. Zhang Y, Yu H, Luo X (2019) Design and implementation of an electronic ticketing system based on mobile communication technology. In: Proceedings of the 5th international conference on transportation information and safety (ICTIS 2019), pp 635–643
4. Gopalan R, Suresh P, Ramachandran R (2018) Design and development of an electronic ticketing system using near field communication (NFC) technology. Int J Eng Sci Comput 8(1):16591–16595

5. Amaral AR, Vasconcelos LA (2017) Design and implementation of an electronic ticketing system for urban public transportation. Int J Interact Multimedia and Artif Intell 4(4):59–64
6. Ali I, Afzal H (2016) Design and implementation of electronic ticketing system for public transport. Int J Eng Res General Sci 4(2):101–108
7. Sharma R, Kumar A (2015) Design and development of a web-based electronic ticketing system for local commuter transportation. Int J Transport Syst 3(2):89–103
8. Wang Y, Li Q (2014) Designing a mobile-based electronic ticketing system for local commuter transportation. Transport Res Record: J Transport Res Board 2442:30–39
9. Li X, Wu J, Zhang S (2013) Design and implementation of an electronic ticketing system using contactless smart card technology for local commuter transportation. In: Transportation research part C: emerging technologies, vol 36. pp 1–15
10. Chen H, Liu S (2012) A framework for the design of an electronic ticketing system in local commuter transportation. J Public Transport 15(3):95–110
11. Johnson R, Smith A (2011) Design considerations for an electronic ticketing system in local commuter transportation. In: Proceedings of the 14th international conference on urban transport and planning (UTP 2011), pp 145–150
12. Cosmas NI, Etus C, Chukwuemeka, Ajere I, Agomuo U (2015) Online bus ticket reservation system. Stat Comput 1

"A Novel Software for Fruit Price Prediction Using Machine Learning"

Omkar Khade, Yash Kadam, Parikshit N. Mahalle, Yashwant S. Ingle, Mandar P. Diwakar, and Monali R. Borade

Abstract This research paper presents a novel approach for fruit price prediction system using machine learning. The aim of this study is to develop an accurate and efficient method for forecasting fruit prices based on historical data. The proposed methodology involves several key steps. First, the fruit dataset is loaded and preprocessed. Then, train–test split is performed to create separate datasets for model training and evaluation. Two regression models, random forest regressor (RFR) and decision tree regressor, were evaluated, and the results showed that the RFR performed better. Next, a pipeline is constructed, incorporating column transformer for feature transformation and RFR for price prediction. The pipeline is trained on the training dataset and evaluated on the testing dataset using metrics such as R-squared score and RMSE. The trained model is then utilized to predict fruit prices for a specified center and month, with the top six fruits having the highest predicted prices displayed. Visualizations in the form of bar charts are generated to compare the predicted prices with the actual prices. Additionally, a Web interface is provided through Flask application, enabling users to interact with the trained model and obtain real-time fruit price predictions. The application integrates the trained pipeline, dataset, and visualization capabilities to deliver an intuitive user experience. Experimental results demonstrate the effectiveness of the proposed machine learning

O. Khade (✉) · Y. Kadam · P. N. Mahalle · Y. S. Ingle · M. P. Diwakar · M. R. Borade
Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India
e-mail: omkar.22110174@viit.ac.in

Y. Kadam
e-mail: yash.22220009@viit.ac.in

P. N. Mahalle
e-mail: parikshit.mahalle@viit.ac.in

Y. S. Ingle
e-mail: yashwant.ingle@viit.ac.in

M. P. Diwakar
e-mail: mandar.diwakar@viit.ac.in

M. R. Borade
e-mail: monali.borade@viit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_53

approach for fruit price prediction, showing promising accuracy and usability. The results of this work can be used in variety of fields, including agricultural markets, supply chain management, and consumer decision-making.

Keywords Fruit price prediction · Machine learning · Pipeline-based approach · Scikit-learn · Regression analysis · Feature engineering · Data preprocessing

1 Introduction

This study presents a novel fruit price prediction system that leverages machine learning algorithms to overcome the limitations of traditional techniques, enabling precise predictions for informed decision-making and operational optimization in the agricultural industry. Our research aims to harness the potential of machine learning in fruit price forecasting by leveraging its pattern recognition capabilities, ultimately developing a system that overcomes challenges and enhances accuracy and efficiency for stakeholders in the agricultural industry.

The core components of our system involve preprocessing the collected dataset, including handling missing values, detecting outliers, and normalizing the data. Additionally, we accounted for seasonality and temporal trends that may impact fruit prices. By collecting a comprehensive dataset from diverse markets and regions, we aimed to build a robust predictive model for fruit price forecasting.

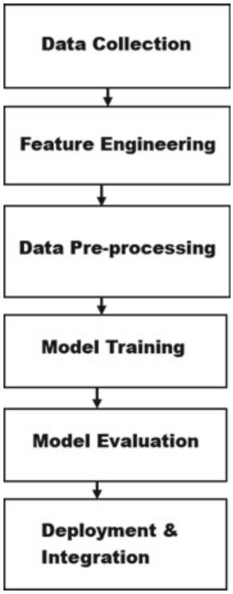
To enhance the accuracy of fruit price prediction, we employed feature selection methods to identify the key attributes influencing fruit price fluctuations. This process aimed to improve the model's ability to capture the underlying patterns and dynamics of fruit pricing. Furthermore, we performed feature engineering techniques, including one-hot encoding and scaling, to effectively preprocess the data and facilitate training of the machine learning models.

Our system employs optimized random forest regression and decision tree regression models on a preprocessed dataset, yielding promising results in fruit price prediction with strong correlations between predicted and actual prices. Comparative analyses highlight the superior performance and efficiency of our machine learning-based system compared to traditional forecasting methods. Our innovative fruit price prediction system leverages machine learning to revolutionize decision-making in the agricultural industry, enhancing accuracy, efficiency, and market outcomes (Fig. 1).

2 Fruits Price Datasest

The study utilized a dataset containing monthly price information for various fruits across different cities in India. The dataset was collected from agricultural markets and government reports for the period spanning January to December. The fruits included in the dataset were Ber (Indian jujube), Kinnow, Lime, Mandarin,

Fig. 1 Block diagram for fruit price prediction



Mosambi, Grape_t_seedless, Guava, Litchi, Bennet, Gulabkhas, Kesar, various types of Mangoes, Parkleman, Aonla, various types of Apples, Papaya, Pineapple, Pomegranate, Sapota, and Banana.

The dataset provided monthly average prices in INR for each fruit and city combination, encompassing Indian cities like Ahmedabad, Bangalore, Delhi, Kolkata, Mumbai, and more. This comprehensive dataset facilitated the analysis of price trends and variations, offering insights into the Indian fruit market. Data collection involved multiple sources such as market reports, government publications, and surveys to ensure accuracy and reliability (Fig. 2).

Month	Centername	Ber	Kinnow	Lime	Mandarin	Mosambi	Grape_t_seedless	Guava	Litchi	...	Mango_totapari
January	Raipur	2400	0	4575	1650	4350	0	2600	0	...	0
October	Nasik	0	0	1000	0	0	0	0	0	...	0
June	Hyderabad	0	0	2000	0	2100	0	1660	0	...	0
December	Nasik	0	0	1400	0	0	0	0	0	...	0
June	Srinagar	0	0	6200	0	0	0	0	0	...	4800
October	Chennai	0	0	5800	0	3300	0	0	0	...	0
March	Hyderabad	0	800	2300	1600	1900	1750	1650	0	...	0

Fig. 2 Dataset of fruits price prediction

3 Related Work

Fruit price prediction has been a subject of interest in agricultural economics and machine learning research. Several studies have explored the application of machine learning techniques to forecast fruit prices, aiming to assist farmers, traders, and policymakers in making informed decisions. In this section, we review some of the relevant works in this domain.

1. Femling et al. "Fruit and Vegetable Identification Using Machine Learning for Retail Applications" [1] presents a system that uses a video camera attached to a Raspberry Pi to identify fruits and vegetables in the retail market. The system allows customers to label the desired items with a price based on their weight, reducing the need for human–computer interactions. By employing convolutional neural networks, different objects can be classified accurately. Usability testing through a heuristic evaluation demonstrates that the system improves user-friendliness compared to existing manual systems.
2. Hossain et al. "Automatic Fruit Classification Using Deep Learning for Industrial Applications" [2] This paper presents an efficient deep learning-based framework for fruit classification, with applications in supermarkets and dietary decision-making. The framework utilizes two different deep learning architectures and achieves high classification accuracies on both clear and challenging fruit image datasets. The proposed models demonstrate accuracies of 99.49% and 99.75% on dataset 1, and 85.43% and 96.75% on dataset 2, respectively.
3. Fang et al. "Fruit Market Trend Forecast Using K means-based Deep Learning Models" [3] This paper proposes an image clustering-based deep learning framework for precise forecasting of fruit market trends. Fruit price data is transformed into fixed-length curve images, segmented into input, and output curves, and labeled using K-means clustering. Convolutional neural network (CNN), long short-term memory (LSTM), and hybrid models are trained, with CNN showing superior capability in fruit market prediction. However, the prediction accuracy of the trained models may not be sufficiently high.
4. Apolo-Apolo et al. "Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV" [4] This study presents an automated image processing methodology using deep learning to estimate citrus fruit yields accurately and early. A Faster R-CNN model detects and counts oranges with a 6.59% error rate compared to visual counting. A long short-term memory (LSTM) model achieves approximately 4.53% error in yield estimation per tree and demonstrates potential for total yield estimation.
5. Gill et al. "Fruit Image Classification using Deep Learning" [5] This study proposes a deep learning-based method utilizing CNN, RNN, and LSTM models for fruit classification, demonstrating effective accuracy and quantitative analysis. The proposed scheme's computational efficiency paves the way for real-time fruit classification applications in the future.

6. Ma et al. "Characteristic mango price forecasting using combined deep-learning optimization model" [6] This study proposes a deep-learning combination forecasting model for mango price prediction, addressing the challenges posed by the volatility of mango prices. The model, based on a back-propagation long short-term memory (LSTM) neural network, achieves high accuracy with low error rates and a strong determination coefficient. The combined model outperforms separate BP and LSTM models, demonstrating superior prediction capabilities and generalizability.
7. Sharma et al. "Price Prediction Model of fruits, Vegetables and Pulses according to Weather" [7] This paper addresses the need for profit-oriented agriculture in India, where climatic conditions impact crop yields and prices, affecting farmer's livelihoods. The proposed method utilizes decision tree regression to anticipate the prices of fruits, vegetables, and pulses based on historical data and various factors. Additionally, the method provides a time series analysis of price and profit for future years compared to previous years.

This research paper presents a novel machine learning approach for accurate and efficient fruit price prediction using historical data. The methodology utilizes regression models like random forest regressor and incorporates feature transformation and prediction pipelines. The system includes a Web interface for real-time price predictions, enhancing usability. Related work in the field showcases the effectiveness of machine learning techniques in various domains, including fruit and vegetable identification, market trend forecasting, and agricultural production prediction, contributing to improved decision-making and market analysis.

4 Proposed Methodology

4.1 Data Collection

The first step in the project was to collect a dataset that would serve as the basis for fruit price prediction. The dataset was carefully curated and included historical data on fruit prices for various cities. It contained relevant features such as the month and the center name, along with the corresponding prices of different fruits. This dataset provided the necessary information to train and evaluate the machine learning models.

The collected data should be accurate, comprehensive, and representative of the target market or markets. It is crucial to make sure the data is accurate and gathered from credible sources. This may involve collaborating with industry experts, agricultural organizations, or government agencies responsible for monitoring and reporting fruit prices.

4.2 Data Preprocessing

The acquired data should be cleaned and preprocessed to deal with missing values, outliers, and inconsistencies. Normalize or scale the data if necessary to ensure fair comparisons and accurate model training.

4.3 Feature Selection/Engineering

Identify the most informative features for fruit price prediction. The independent features considered for prediction were the month and center name. On the other hand, the remaining features, such as specific fruit types and corresponding prices, were treated as dependent features since they were influenced by the month and center name.

4.4 Model Selection

Depending on the features of the data and the goals of the research, choose the best machine learning algorithms. Regression models, time-series analysis techniques, neural networks (such as RNNs and LSTMs), or ensemble methods could be suitable options to consider.

4.5 Machine Learning Model Training

To predict the price of a fruit, we trained two different regression models:

- Random forest regressor (RFR): A widely used machine learning algorithm for regression tasks is the random forest regressor. It works by constructing multiple decision trees at training time and outputting the average prediction from the individual trees as the final prediction. For predicting fruit prices, we can use the random forest regressor to create a model that, based on the inputs, can forecast the prices of various fruits.
- Decision tree regressor (DTR): The decision tree regressor is a versatile machine learning algorithm for fruit price prediction, utilizing recursive partitioning of the dataset based on input feature values to capture nonlinear relationships and handle categorical and continuous data. It offers a nonparametric approach by reaching a maximum depth and effectively predicting the target variable.

In order to prepare the data for training, a column transformer was employed to apply different transformers to different columns of the dataset. The transformers used in the pipeline are:

- Standard Scaler: Used to standardize the numerical data.
- One Hot Encoder: Used to encode categorical data.

Two pipelines were constructed, one for each regression model, utilizing the column transformers and respective regression models. Training on the training data, predictions were made on the test data, and model performance was evaluated using R-squared score and RMSE metrics, offering insights into the models' ability to capture fruit price variability.

4.6 Model Integration

This project utilizes a Flask application to enable user interaction with a machine learning model, allowing customized input parameters, and providing fruit price predictions, with the model loaded using the pickle library for seamless integration.

The Flask app consists of home.html and result.html templates, facilitating user input of parameters and displaying predicted prices, with the home.html form capturing input and the result.html template presenting the user-friendly predicted price.

This Flask Web app allows users to input a month and center, predicting fruit prices and displaying the top six predictions in a bar chart. It also generates separate bar charts for the top six actual prices and a combined chart of predicted and actual prices, all created using Matplotlib and embedded in the Flask-rendered HTML template in base64 format.

To run the Flask app, use the "app.run(debug = True)" command to launch the development server with debug mode, and access the home page (http://localhost:5000/) to input parameters and view predicted prices, making it deployable locally or on a server for public use.

5 Results

After training and testing the two regression models, we found the following results (Table 1).

Table 1 R-square and RMSE

Model	R-square	RMSE
Random forest regressor (RFR)	0.22	1119.69
Decision tree regressor (DTR)	0.06	1380.21

Comparing the R-square and RMSE values, the random forest regressor (RFR) outperforms the decision tree regressor (DTR) in predicting the target variable. The RFR exhibits an R-square of 0.22, explaining 22% of the variance, with a relatively low RMSE of 1119.69. In contrast, the DTR has an R-square of 0.06, explaining only 6% of the variance, with a higher RMSE of 1380.21. Based on these results, the random forest regressor (RFR) proves to be the superior algorithm for this problem (Figs. 3, 4, 5, 6 and 7).

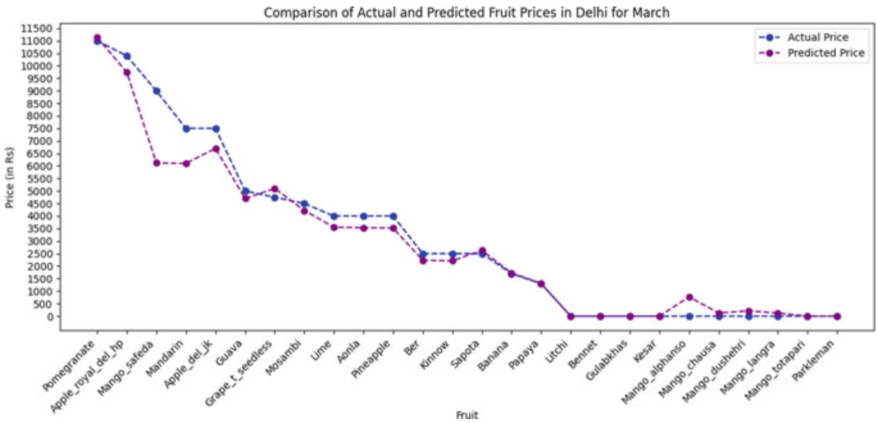


Fig. 3 Line graph: actual versus predicted fruit prices in Delhi for March

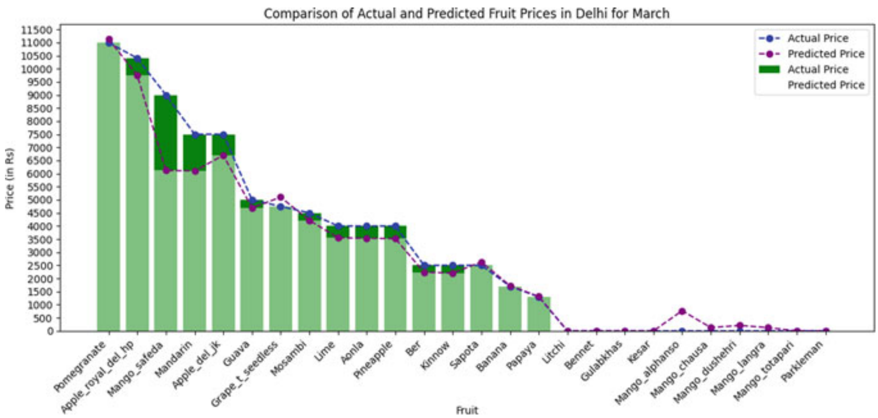


Fig. 4 Graph: actual versus predicted fruit prices in Delhi for March (All fruits)

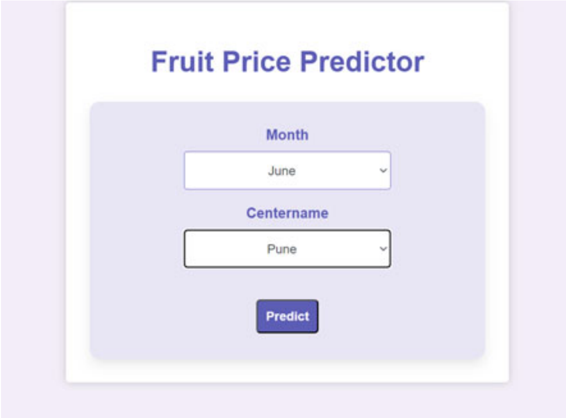


Fig. 5 Flask web interface—home page

6 Discussion

Machine learning models, particularly random forest regression (RFR), have shown promising results for fruit price prediction compared to traditional methods. RFR effectively handles multiple input features, nonlinear relationships, overfitting, and missing data. The feature importance rankings provided by RFR offer valuable insights into the factors influencing fruit prices, aiding market analysis and decision-making. Although decision tree regression is interpretable, it demonstrated slightly lower performance due to its susceptibility to overfitting. Further research can explore regularization techniques and ensemble methods to enhance decision tree performance. Additionally, models like support vector regression, gradient boosting regression, and neural networks (e.g., LSTM) can be considered based on dataset characteristics and interpretability requirements. Careful model selection, hyperparameter tuning, and evaluation techniques like cross-validation are important for balancing accuracy, resources, and interpretability. Despite relatively low R-squared values, this study contributes to fruit price prediction research and its implications span agricultural markets, supply chain management, and consumer decision-making, empowering stakeholders to make informed decisions in production, purchasing, and pricing strategies.

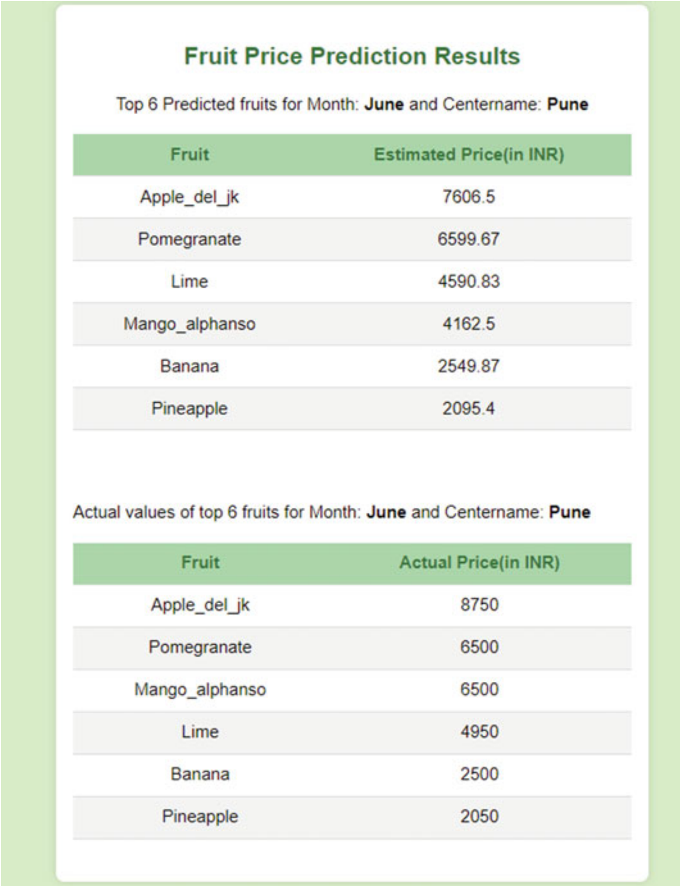


Fig. 6 Flask web interface—result page

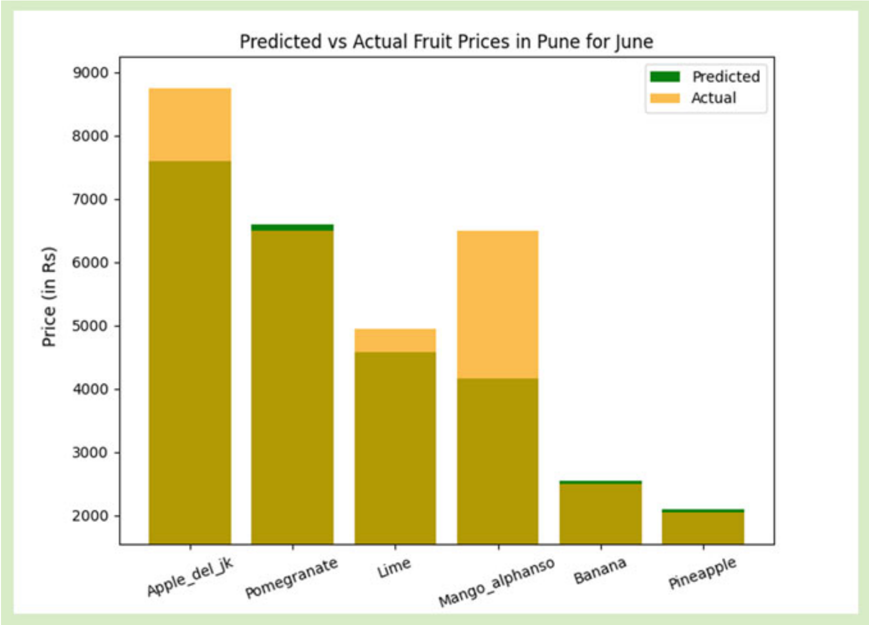


Fig. 7 Flask web interface graph

7 Conclusion

In conclusion, this research paper presented a novel approach for fruit price prediction using machine learning techniques. By leveraging historical data and employing a random forest regression (RFR) model, accurate and efficient predictions of fruit prices were achieved. The RFR model demonstrated superior performance compared to the decision tree regression model, showcasing its ability to capture complex relationships and handle missing data effectively.

The findings of this study contribute to the field of fruit price prediction, providing valuable insights for agricultural markets, supply chain management, and consumer decision-making. The developed machine learning-based prediction system offers real-time predictions and visualizations, enabling stakeholders to make informed decisions regarding production, pricing, and purchasing strategies.

Model selection is a critical aspect of fruit price prediction, and the RFR model proved to be a suitable choice due to its ability to handle many features and capture nonlinear relationships. However, other models such as support vector regression, gradient boosting regression, and neural networks, including LSTM, can also be considered based on the dataset characteristics and specific requirements of the problem.

Future research can focus on exploring ensemble methods, improving model interpretability, and integrating additional data sources to further enhance the accuracy and usability of fruit price prediction systems. Additionally, the deployment of the developed system into existing platforms or as a standalone tool can provide stakeholders with real-time predictions and valuable insights, enabling them to optimize their decision-making processes.

Overall, the results of this study demonstrate the efficacy and applicability of machine learning techniques in predicting fruit prices. This research contributes to the advancement of agricultural markets and decision-making processes, opening opportunities for improving profitability and sustainability in the fruit industry.

8 Limitations and Future Scope

This project has a few limitations that should be acknowledged. Firstly, the reliance on a specific dataset of fruit prices in India may limit the model's accuracy and generalizability due to potential gaps in coverage and representativeness. Secondly, the evaluated regression models achieved relatively low R-squared values, indicating that they explain only a small portion of the variance in fruit prices, suggesting the presence of other influential factors not captured by the dataset. To address these limitations, future directions for this project could involve incorporating additional features such as weather data, transportation costs, and market demand to enhance the predictive power of the model. Exploring advanced regression models like ensemble methods or neural networks could also improve the accuracy of price predictions. Additionally, developing dynamic pricing models that consider real-time data and market fluctuations, as well as continuously updating the model with new data, would ensure its relevance and accuracy over time.

References

1. Femling F, Olsson A, Alonso-Fernandez F (2018) Fruit and vegetable identification using machine learning for retail applications. In: 2018 14th international conference on signal-image technology and internet-based systems (SITIS), Las Palmas de Gran Canaria, Spain, 2018, pp 9–15. https://doi.org/10.1109/SITIS.2018.00013
2. Hossain MS, Al-Hammadi M, Muhammad G (2019) Automatic fruit classification using deep learning for industrial applications. IEEE Trans Industr Inf 15(2):1027–1034. https://doi.org/10.1109/TII.2018.2875149
3. Fang Y, Min X, Zheng L, Zhang D (2019) Fruit market trend forecast using k means-based deep learning models. In: 2019 IEEE 10th international conference on software engineering and service science (ICSESS), Beijing, China, pp 628–633. https://doi.org/10.1109/ICSESS47205.2019.9040838
4. Apolo-Apolo OE et al (2020) Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. Eur J Agron 115:126030. https://doi.org/10.1016/j.eja.2020.126030
5. Gill HS, Khehra BS (2022) Fruit image classification using deep learning, 07 November 2022, PREPRINT (Version 1) available at Research Square https://doi.org/10.21203/rs.3.rs-574901/v1

6. Ma X, Tong J, Huang W, Lin H (2023) Characteristic mango price forecasting using combined deep-learning optimization model. PLoS ONE 18(4):e0283584. https://doi.org/10.1371/journal.pone.0283584
7. Sharma C, Misra R, Bhatia M, Manani P (2023) Price prediction model of fruits, vegetables and pulses according to weather. In: 2023 13th International conference on cloud computing, data science and engineering (Confluence), Noida, India, pp 347–351. https://doi.org/10.1109/Confluence56041.2023.10048880

Quantum Key Distribution and Blockchain Based Secure Authentication in Medical Cyber-Physical Systems

Piyush P. Gawali, Parikshit N. Mahalle, Gitanjali R. Shinde, Nilesh P. Sable, Dattatray G. Takale, and Janki Barot

Abstract Medical cyber-physical systems (MCPSs) enable flexible patient–medical system interaction as the foundation of application in the field of smart health care, realizing all-encompassing three-dimensional medical service. The security barrier must be built using a secure and trustworthy device identity authentication mechanism in order to meet the controlled and believable requirements of MCPS. This work, therefore, develops an efficient authentication and encryption system (EAES) to improve the security and authentication of MCPSs. This research suggests a secure authentication method to safeguard sensitive healthcare data. The suggested modified "elliptic-curve Diffie–Hellman" (ECDH) encryption is first applied to the medical data that is acquired from the patients and stored in the cloud in encrypted form. This is done in order to guarantee secure access to the stored data for use in medical data analysis. The key in the ECFH is controlled via a quantum key distribution system. Additionally, blockchain is considered for the mutual authentication

P. P. Gawali (✉) · G. R. Shinde · D. G. Takale
Department of Computer Engineering, Vishwakarma Institute of Information Technology, SPPU, Pune, India
e-mail: piyush.gawali@viit.ac.in

G. R. Shinde
e-mail: gitanjali.shinde@viit.ac.in

D. G. Takale
e-mail: dattatray.takale@viit.ac.in

P. N. Mahalle
Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India
e-mail: parikshit.mahalle@viit.ac.in

N. P. Sable
Department of Information Technology, Vishwakarma Institute of Information Technology, SPPU Pune , Pune, India
e-mail: nilesh.sable@viit.ac.in

J. Barot
Silver Oak University, Ahmedabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_54

of users in the MCPSs. The proposed methodology is implemented and evaluated by using performance measures. Additionally, it is compared with the conventional techniques.

Keywords Quantum key distribution · An encryption technique · Efficient authentication encryption technique · Elliptic-curve Diffie–Hellman

1 Introduction

The market for medical equipment is evolving swiftly as it embraces the promise of embedded software and network connectivity. Rather than standalone devices that can be developed, approved, and regularly used to treat patients independently of one another, researchers will soon have to deal with distributed systems that simultaneously govern a number of different aspects of the patient's physiology [1]. The presence of "embedded software that controls the devices, networking capabilities," and the complex physiological dynamics that patient bodies show all contribute to the fact that current medical device systems, which are also often referred to as MCPS, are their own unique class of cyber-physical systems [2].

The creation of safe and effective MCPS will necessitate the use of new design, verification, and validation approaches due to the increase in size and complexity. The MCPS design should make better use of model-based technology [3]. Models should include tools and how they communicate with one another, but patients and carers should also be included. Also, to approve their usage in patient care, MCPS will institute a new regulatory system [4]. On the one hand, the U.S.'s current traditional process-based regulatory structure [5].

To build MCPSs, it will be necessary to overcome technological barriers in the development of the system's structural components, such as cloud computing architectures, quick Internet and mobile connectivity, and sensors [6]. To ensure the privacy of patient's medical records during transit from the sensor networks to the cloud and from the cloud to physician's mobile devices, an MCPS will need a sophisticated cryptographic architecture [7]. New encryption technologies, in contrast to the standard cryptographic methods utilized in this design, provide flexibility in data sharing and computer security [8]. The intelligent healthcare system is linked to the healthcare industry's 4.0 initiative. "The data mining, artificial intelligence, machine learning, blockchain, cloud computing, and fog computing are just few of the technologies used in Industry 4.0".

The following sections of the paper are already set up as follows: Sect. 2 gives the associated works of the MCPS security enhancement, and Sect. 3 offers the suggested architecture of the blockchain technology and secure encryption method. In Sect. 4, the outcome assessment of the suggested strategy is laid out. Section 5 of the essay contains its conclusion.

2 Related Work

This section reviews works that are linked to the security of medical cyber-physical systems.

Kocabas et al. in [11], the typical architecture of an MCPS consists of four levels: data collection, data aggregation, cloud processing, and action. These layers are in order from the bottom up. To protect the confidentiality of the data stored inside each layer, it is necessary to make use of distinct encryption strategies. This is necessitated by the fact that the hardware and communication capabilities of each layer are distinct. In this section, we will evaluate both well-established and recently developed cryptographic algorithms based on how well they enable safe data exchange, secure processing, and secure storage.

Qiu et al. [12], the method of secure data storage and sharing was developed by making use of a "selective encryption algorithm in conjunction with fragmentation and dispersion to maintain data security and privacy even in circumstances in which both the transmission medium (such as cloud servers) and the keys are compromised. This is possible because the algorithm uses fragmentation and dispersion to distribute the data across multiple locations". This method, which is based on a design approach, gives people access control for the sharing of data while simultaneously maintaining the data on a reliable device, such as the user's smartphone. In order to demonstrate that the algorithm is effective, you need also to analyze how something works when it is implemented in a mobile application.

Kumar et al. [13], a blockchain solution that immediately protects patient data security and privacy was demonstrated. Unauthorized data access by unauthorized users is controlled via the blockchain modules. According to the performance analysis of the suggested framework, data sharing occurs with low latency, a high transaction rate, and high throughput as the number of network users rises.

Rajawat et al. [14], in this research, find out the order to safeguard intelligent roads and vehicles as part of ITS. In order to provide a sufficient Internet of Vehicles (IoV) service while mining an acceptable degree of safety assurance, the recommended architecture takes into consideration a semi-distributed paradigm for blockchain implementation. This is done in order to fulfil the requirements of both objectives. According to the findings of the tests conducted on intelligent highways and smart parking management, the model that was developed is capable of achieving comparably excellent data delivery while simultaneously reducing latency. This paves the way for innovative applications of blockchain technology in IoV, which may be used to create smart transport systems that are stable and trustworthy.

Li et al. [15], the subject of fixed-time secure cooperative control for a certain class of nonlinear "cyber-physical systems (CPSs) under denial-of-service (DoS)" attacks has been brought. A fix-time secure leader-following consensus solution for connective-preserved DoS attacks was initially devised utilizing an event-based data transfer technique. In order to stop Zeno behavior, a label measurement error was also made. But hostile rivals may unavoidably disrupt the fundamental framework of communication. In order to overcome this difficulty, a brand-new fixed-time secure observer is developed. A self-triggered control mechanism was also released, which eliminates continual communication and creates a secure leader-following consensus that is susceptible to DoS assaults that retain or destroy connective tissue.

Alzubi et al. [21], a brand new model known as PPEDL-MDTC, which stands for privacy-preserving encryption with DL-based medical data transmission and categorization, has been developed. “The sailfish optimizations and the multiple key-based holomorphic encryption (MHE) approach are included into the MHE-SFO algorithm-based encryption process, which is created with the help of the suggested model. In addition to this, the cross-entropy based artificial butterfly optimization-based approach as well as the best deep neural network (ODNN)-based classification were used. P block cipher-based area of Interest Medical Image Encryption with multiple maps was successfully shown by Kiran et al. [22]. [Cipher] After the regions of interest (ROI) are extracted with the help of a Laplacian edge detection operator, the image is then segmented into wanted (ROI) and unwanted (RONI or ROB) portions using the edges in the blocks”. This occurs after the regions of interest (ROI).

A cyclic permutation of the critical ROI zones is then performed using Arnold’s cat map and angle value. The ROI part that had been permuted was encrypted using the shuffling process, while unimportant regions remained untouched. In order to successfully manage EMRs, Li et al. [23] developed the on-chain ledger and off-chain storage (OLOS) paradigm, which entails uploading EMR indexes to a public blockchain and storing the actual EMR data on local servers. This model can preserve and free up public ledger space. In the meanwhile, data security can be maintained by merely refraining from acts on real EMR data. In order to improve system security, traditional procedures are used. Therefore, accurate precision is not achieved effectively. The best method should therefore be created in order to accomplish the highest level of security.

Nevertheless, there are significant constraints on the growth of MCPS. The main goal of MCPS research is to decrease the likelihood of unsafe system circumstances by applying software engineering techniques and to improve understanding of and support for MCPS development efforts in safety analysis. The primary goal is to minimize dangerous states, failed control loops, component interactions, and inadequate requirements for the design and development of MCPS. The four methods are used for empower the security of the MCPS. These methods are lacking in some limitations such as high delay, high encryption and decryption time, and improper key distribution. Additionally, it does not achieve efficient accuracy. So, the optimal technique should be developed for attaining optimal.

3 Proposed System

The proposed system model includes a key authority, ring network, three nodes, and quantum channel for quantum key distribution along with patients, public medicine, home infusion, home health, physician practice, health information exchange, health plan, hospitals, and pharmacies. It also includes a healthcare user group, a healthcare

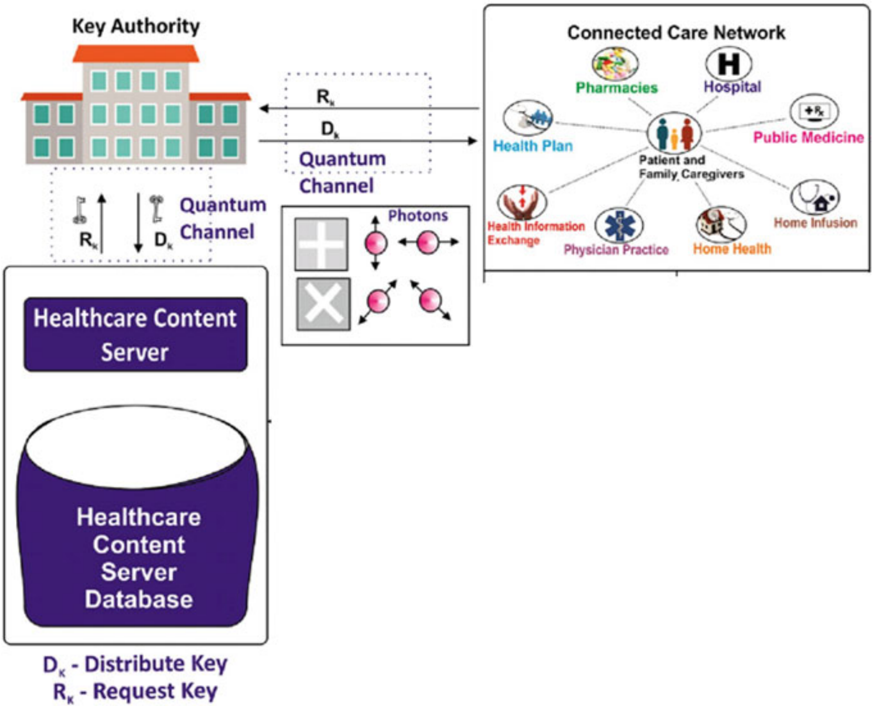


Fig. 1 System architecture

content server, and a healthcare content server. Figure 1 shows the structure of the suggested architecture.

Initially, MCPSs is designed with patients and data. After that, the data is stored in the system which secured by using modified elliptical curve Diffie–Hellmann encryption. In this encryption, the key is optimally distributed by considering quantum key distribution. Additionally, to enhance user authentication process, the blockchain technology is developed. The section that follows provides an explanation of the suggested methodology's whole design.

3.1 Blockchain Technology Based Authentication Process

The utilization of blockchain technique in mutual authentication consists of different advantages when compared with different methodologies. "It is a general technique to avoid the middleman attack as it contains additional overhead expenses. Moreover, this entity is highly useful, and it is only path different parties can achieve an agreement on a set of MCPS. Blockchain design in this case not only manages the middleman attack but also gives an attack removal of novel opportunities of mutual authentication of MCPS. Hence, this system contains various regulations on a key set of information without an intermediate party contains to account for complete transactions" [16] (Fig. 2).

The primary goal of a blockchain establishes which transaction phase must be improved for all active nodes taking part in the architecture. With the help of this method, users can now handle data that has moved or changed states or locations without it moving simultaneously through two phases. The simple design of blockchain is a distributed decentralized ledger. The block generation time has a direct impact on network speed and is simple to create. A distributed, decentralized ledger is the foundation of the simple measure blockchain. The medical information is then added to the ledger and is not removed. The traits are therefore described as write once, read mostly (WORM) characteristics. A network of peer-to-peer nodes shares the blockchain copy, which is arranged into blocks with all the data that has been saved. A consensus on the medical data must be reached by the majority of participating nodes because any other attempts to control the information are banned. In contrast to centralized databases, the unified design is unaffected by single points of failure. The generalized blockchain architecture with multiple nodes is illustrated in Fig. 1. This is a simple structure of the blockchain [17]. This block is divided into two phases, one is a block header which consists of difficulty target, Merkle root, pervious blocked hash, nonce, timestamp, and some other required data. The remaining section is the block body, and it contains the medical data.

$\mathrm{VirID_{DJ}}$ is a "virtual id, $\mathrm{SK_{DJ}}$ is a secondary key, $\mathrm{PK_{I}}$ is defined as public key, M is defined as the message, K is a system parameter, TS is defined as timestamp, and n is server connected user".

Registration process

Before conducting any operations in this structure, each user must register with the registration center. The request was sent to the medical facility after computing the public key pairs of the users.

Mutual authentication process

Beginning with the message and the blockchain virtual id and time step parameters will complete the authentication process. Virtual ID is then verified. The communication also has a signature and is encrypted. The basis of signature is used to verify the identity last. The original message is finally converted via the reverse technique.

A variety of blockchain are available, and they can be categorized based on the architecture, node visibility, user connectivity to the network, and the privileges that come with that connectivity. Hence, private, permissioned, and public blockchain are considered. Every blockchain contains a various level of requirement on transparency, efficiency, immutability, and anonymity. Normally, the centralized data provides lower security stage contrasted with the blockchain because of the way information is saved, edited, and accessed. Any user with privilege can edit the information with the process of chaining a master copy. A single point of failure could compromise the security of a blockchain because it is entirely dependent on local architecture [18]. In the mutual authentication stage, the registered center and medical server (MCPS)

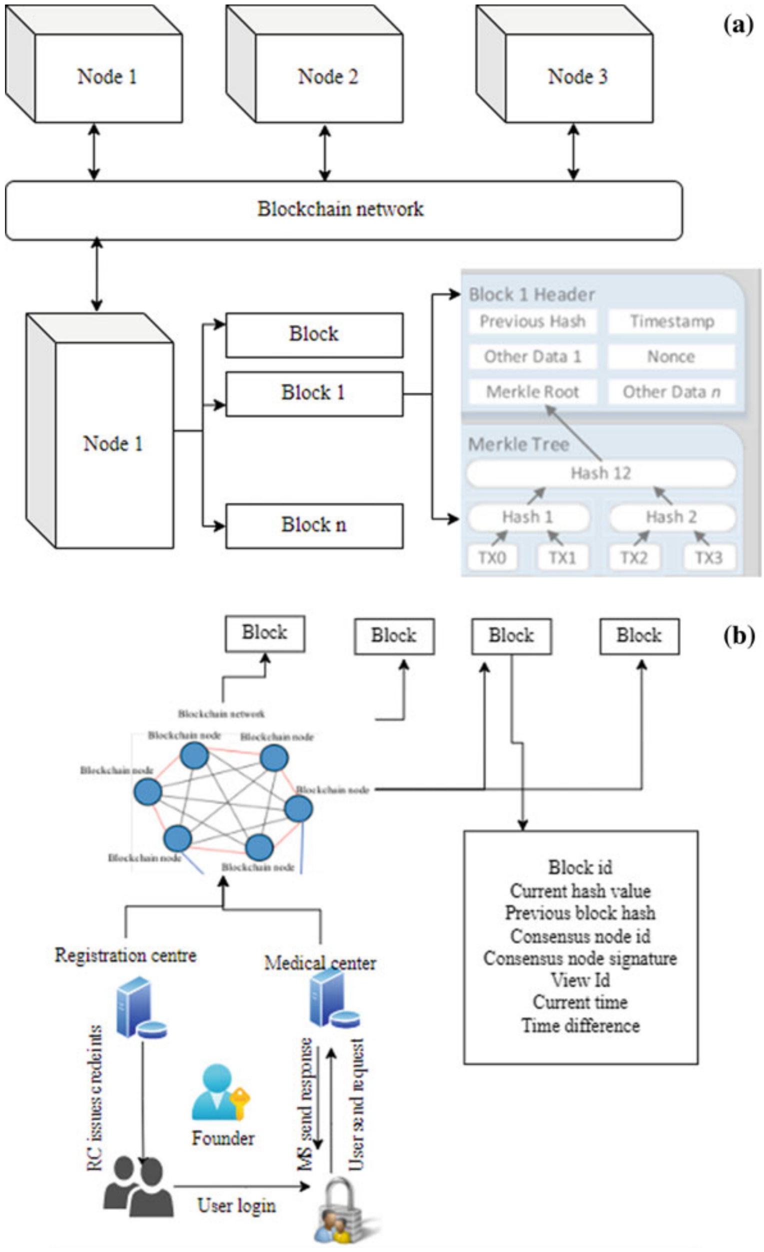


Fig. 2 **a** General architecture of the blockchain technology and **b** authentication stage

data is formulated. Based on the registered center, the user sent request by considering login process. The medical server grants the request. The user authentication mechanism is activated on the medical server.

3.2 Modified Elliptical-Curve Diffie–Hellmann Encryption

ECC is defined as a technique in the field of public key cryptography. Additionally, it creates utilization of elliptic curves in that coefficients and variables are parameters of the finite filed. It provided a same level of security with a reduced key size when compared to another technique. Hence, it helps to cut down on storage and transmission needs.

The elliptic curve

$$(\text{A1, A2, A3, A4 and A6}) \tag{1}$$

Point on the curve (x, y) is related with the field f_2^k. This field is made in public presented by a demonstrated third party. Here, a selects a private parameter,

$$N_a\left(0 \le N_a \le 2^k - 1\right) \tag{2}$$

and calculates $q_a = N_a(x, y)$ which sends q_a to party b. Additionally, party b selects a private parameter,

$$N_b\left(0 \le N_b \le 2^k - 1\right) \tag{3}$$

and calculates $q_b = N_b(x, y)$ and sends q_b to party a. Party a is calculated,

$$N_a q_b = N_a N_b(x, y) \tag{4}$$

and party b is calculated,

$$N_b q_a = N_a N_b(x, y) \tag{5}$$

Due to their similarity, combined parameters serve as the encryption's private keys. The suggested encryption offers greater security than the RSA technique. The suggested method provides effective forward secrecy but calls for key exchange between parties (users). If data needs to be modified between numerous parties, each party can distribute keys using a similar curve [19] and do away with the need for each party to have a unique set of private and public key pairs, or for a server to control the data exchange as in RSA. The party a, b, c will give their public keys q_c, q_a and q_b and after that, computation of N_c, q_a and q_b, N_b, q_a and q_c, N_a, q_a and q_c, respectively. The general private key is,

$$N_a, N_b, N_c(x, y) \tag{6}$$

In the encryption technique, the key is distributed by using quantum key distribution. The thorough explanation of the distribution of quantum keys. In the modified elliptical curve Diffie–Hellmann encryption technique, the key is distributed by using quantum key distribution. The detail explanation of quantum key distribution.

3.3 Quantum Key Distribution (QKD)

The quantum channel is used by the key authority to generate a series connection of 500 qubits and deliver it to a server that hosts group and healthcare content. It picks a diagonal or rectilinear polarization basis or qubit basis at random. Then, with -450 and + 450 shifted to the phases |+ > and |− > , it maps vertically and laterally using the qubit phases |0 > and |1 > . The main authority delivers 500 qubits with a 0.5 bias consideration to a server for groups or healthcare content. Moreover, the quantum channel can be used to eavesdrop on information at a rate of 0.1 and with a bias of 0.5. A listener is now managing in on the channel [20]. Figure 3 shows a sample architecture for quantum key distribution.

Sifting stage

The healthcare content server and clusters publish on a traditional public channel the qubits that achieved a measure too quickly. Then, healthcare clusters interact

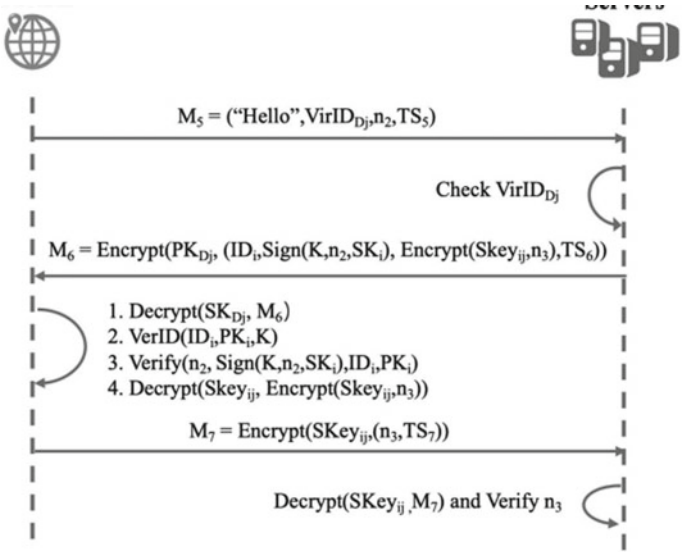


Fig. 3 Mutual authentication process

and disclose the data bases it utilized. The final step is to authenticate using three messages. As a result, the bases average a connected match with 50% of the time, and they all add their associated bit to their personal key. The two keys must match exactly for the eavesdropper to succeed in the absence of channel noise. The 500 transmitted qubits were used as input for the shifting stage, which was reduced to 257 bits. Using the universal hashing method of the linear feedback shift register (LFSR) and a pre-shared secret key for verification, the healthcare and key authority user clusters authenticate their related-on exchange messages.

$$|0 \geq 1-\rangle \tag{7}$$

$$|1 \geq 1|\rangle \tag{8}$$

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |\rangle \tag{9}$$

$$\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |\backslash\rangle \tag{10}$$

Four-state polarization is formulated as follows,

$$\Big|- > \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1\Big|x > +0|y\rangle \tag{11}$$

$$\Big|| > \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0\Big|x > +1|y\rangle \tag{12}$$

$$\Big|- >= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}\Big|x > +\frac{1}{\sqrt{2}}|y\rangle \tag{13}$$

$$\Big|- >= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}\Big|x > -\frac{1}{\sqrt{2}}|y\rangle \tag{14}$$

Error computation stage

To establish whether they should continue with error rectification or if they should halt the protocol in relation to a preset error tolerance reference, which is often about 11%, healthcare and key authority assess the mistake rate in their filtered keys. This allows them to decide whether or not they should continue with error rectification. User groups in the healthcare industry and key authorities sometimes permute their filtered keys in order to "flatten" the faults that occur across the bit string. It does this by comparing a portion of their sifting keys with the mistakes that have been flattened. This keeps the error calculate in check. The user clusters that are responsible for healthcare and key authorities use an interactive error correction procedure known as cascade through the public channel in order to find and rectify any incorrect bits that may be present in their filtered bit strings.

Error correction stage

The error validation step is validated and authenticated by key authorities and user groups in the healthcare industry by calculating the hash of their mistake-corrected keys using each organization's unique digest. The use of 64-bit key material was implemented for authentication purposes. Authentication was performed using a method known as linear feedback shift register (LFSR), which is a global hashing approach. This technique safeguards the data while at the same time authenticating the users.

4 Performance Evaluation

In this section, the suggested approach is verified and examined. This suggested methodology develops the MCPSs and executes it using a MATLAB/NS2 program. Several metrics, including encryption time, decryption time, processing time, and quantum key distribution time are taken into account while evaluating the suggested method. Table 1 lists the parameters of the suggested strategy.

The proposed methodology is validated by considering the encryption time which is presented in Fig. 3. The proposed method is contrasted with traditional methods like AES, ECC, and RSA, respectively. The proposed methodology is achieved 1 ms. After that, the conventional techniques are achieved as 2, 5 and 6 ms. Based on this procedure, the proposed approach attained optimal results in the measure of encryption time. The proposed methodology is validated by considering the decryption time which is presented in Fig. 4. The proposed approach is compared with the conventional techniques such as AES, ECC, and RSA, respectively. The proposed methodology is achieved 1 ms. After that, the conventional techniques are achieved as 2, 3 and 4 ms. Based on this procedure, the proposed approach attained optimal results in the measure of decryption time. The proposed methodology is validated by considering the processing time which is presented in Fig. 4. The proposed method is contrasted with traditional methods like AES, ECC, and RSA, respectively. The proposed methodology is achieved 2 ms. After that, the conventional techniques are

Table 1 Simulation parameters

S. No	Method	Description	Value
1	Proposed method	Data size	200 MB
2		Transactions per block	50
3		Ledger model	CouchDB
4		Number of channels	1
5		Consensus mechanism	Raft

achieved as 2, 4 and 6 ms. Based on this procedure, the proposed approach attained optimal results in the measure of processing time. The proposed methodology is validated by considering the quantum key distribution time which is presented in Fig. 5. The proposed method is contrasted with traditional methods like AES, ECC, and RSA, respectively. The proposed methodology is achieved 0.2 ms. After that, the conventional techniques are achieved as 0.5, 0.6 and 0.8 ms. Based on this procedure, the proposed approach attained optimal results in the measure of quantum key distribution time (Figs. 6, 7 and 8).

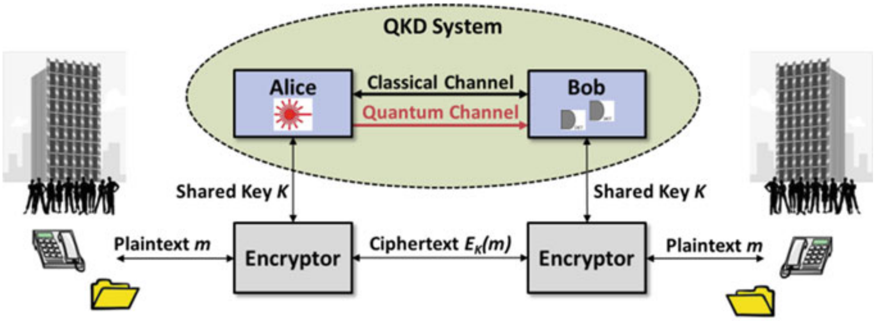


Fig. 4 Quantum key distribution

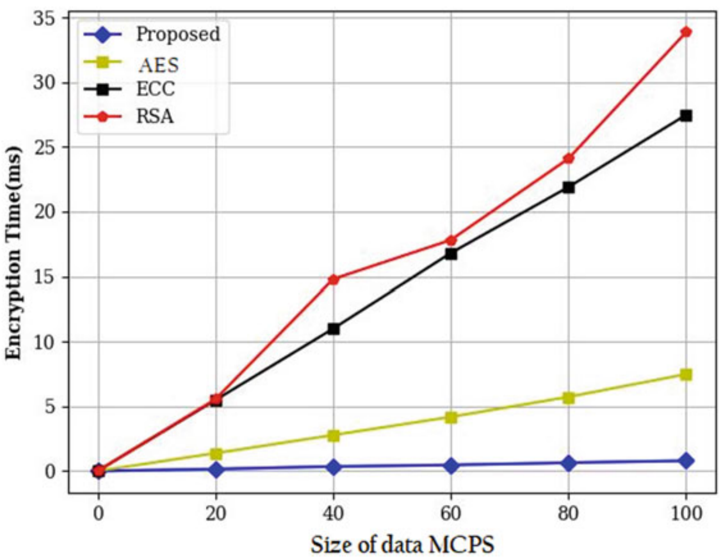


Fig. 5 Encryption time

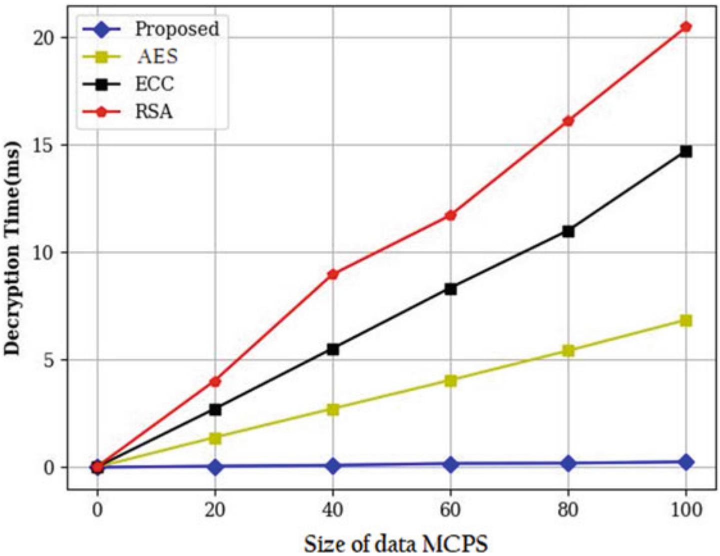


Fig. 6 Decryption time

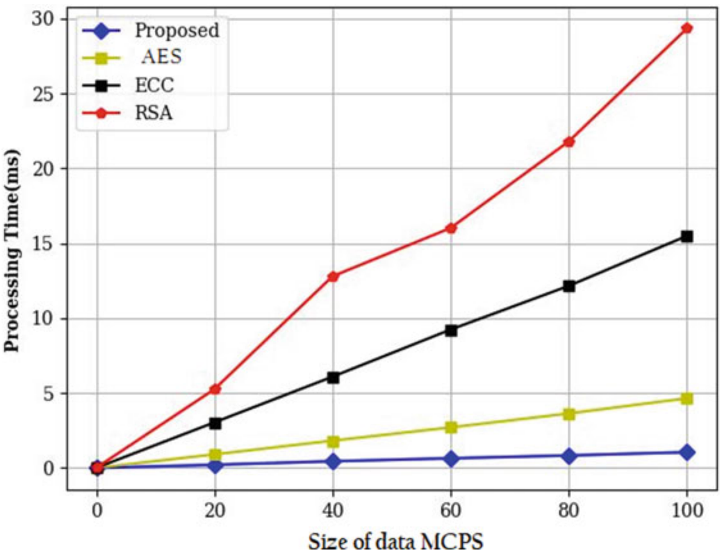


Fig. 7 Processing time

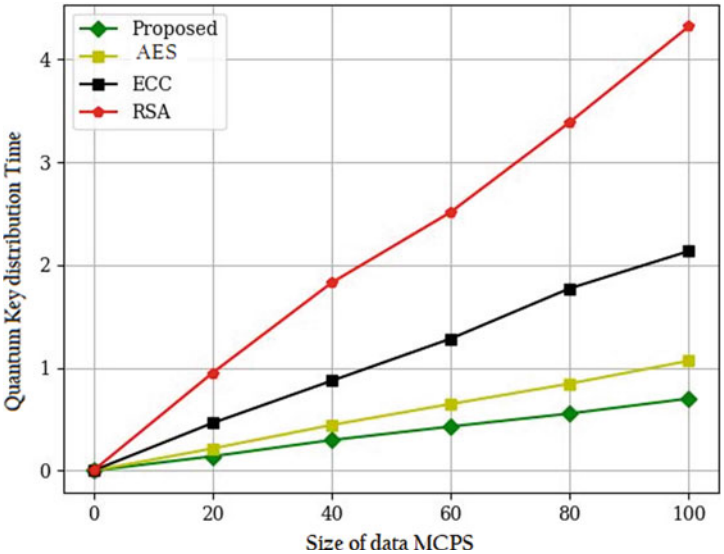


Fig. 8 Quantum key distribution time

Issues and challenges

The developer takes into account the consequences of several aspects when creating the security and privacy system for MCPS in order to strike a better balance between them. A more secure environment must be established, and some issues require special attention.

- **Sharing Data Security:** The patient's medical records, which include outpatient records, hospitalization records, body temperature lists, medical order lists, laboratory test lists, medical image examination data, consent lists for special examinations, consents for operations, lists of operations and anesthesia records, pathological data, nursing records, and other records, are important pieces of information for MCPS. The medical field will significantly develop if these medical records can be shared with patients, hospitals, and researchers. Thus, a current research hot topic is how to execute data exchange of heterogeneous and different big medical data.
- **Lightweight Protocol for Devices:** Low-cost devices and software programs based on medical sensors must adhere to specified operational standards since medical gadgets have limited storage space. Currently, in order to provide sophisticated security protection for medical sensors, we must design many levels of security protocols in line with the application scenarios, notably lightweight security protocols to secure the communication security of sensors.
- **Unsecure Network:** Many different types of enterprises utilize WLAN. Wi-Fi, and other wireless networks are employed in the MCPS industry for hardware and software services due to their ease of use. The vulnerability of Wi-Fi to intrusions

of many kinds, such as unauthorized router access, man-in-the-middle attacks, deception, denial-of-service assaults, violent attacks, etc., is well recognized. Additionally, MCP'S top concerns include network security and the majority of untrusted or unverified free wireless networks in public spaces.

The quantity of medical data is expanding at an exponential rate as a direct result of the fast growth of information technology and the Internet of things (IoT). As a result, the storage and processing of a huge number of medical data have become a new issue. The typical database that is used for the storing and processing of a large quantity of medical data has various flaws, such as a poor processing efficiency, a significant consumption of computer resources, and a sluggish transmission speed. New approaches to the storing and processing of medical data have become available as a result of technological developments such as big data, cloud computing, and blockchain. When dealing with a significant number of medical data storage and processing issues, these technologies offer the benefit of reducing the amount of time spent computing, enhancing the efficiency with which processing is accomplished, and providing a high level of security. Additionally, it will bring forth new issues and obstacles, such as concerns around the safety of storage and the use of resources.

5 Conclusion

An EAES has been designed in this research to improve the security and authentication of the MCPSs. In order to safeguard the sensitive healthcare data, this study suggests a secure authentication method. At first, "the patient's medical data was gathered and encrypted and stored in the cloud, where the proposed ECDH encryption was applied to ensure secure access to the stored data to be used for medical data analysis". In the ECFH, the key has been managed by using quantum key distribution. Additionally, blockchain has been considered for the mutual authentication of users in the MCPSs. The proposed methodology has been implemented and evaluated by using performance measures. Additionally, it has been compared with conventional techniques.

References

1. Rameshwar SK, Malathi P (2023) Splunk-based threat intelligence of cyber-physical system: a case study with smart healthcare. Int J Intell Syst Appl Eng 11(2):537–549
2. Xugui Z, Ahmed B, Aylor JH, Asare P, Alemzadeh H (2023) Hybrid knowledge and data d riven synthesis of runtime monitors for cyber-physical systems. IEEE Trans Dependable and Secure Comput
3. Kamal T (2023) Identifying and protecting cyber-physical system's influential devices for sustainable cybersecurity. IEEE Trans Sustain Comput

4. Vatsal V, Chakraborty A, Rajarajan G, Leo Fernandez A (2023) A detailed investigation of popular attacks on cyber physical systems. In: Cyber security applications for industry 4.0, Chapman and Hall/CRC, pp 1–42
5. Gao Y, Chen J, Miao H, Song B, Yiqin L, Pan W (2022) Self-learning spatial distribution-based intrusion detection for industrial cyber-physical systems. IEEE Trans Comput Soc Syst 9(6):1693–1702
6. Javed S (2022) Towards digitization and machine learning automation for cyber-physical system of systems. Luleå University of Technology, PhD diss.
7. Niu X, Chen Y, Jin H (2022) CRSM: computation reloading driven by spatial-temporal mobility in edge-assisted automated industrial cyber-physical systems. IEEE Trans Industr Inf 18(12):9283–9291
8. Verma R (2022) Smart city healthcare cyber physical system: characteristics, technologies and challenges. Wireless Pers Commun 122(2):1413–1433
9. Keshav K (2022) Exposing security and privacy issues on cyber-physical systems. In: Cyber-physical systems: foundations and techniques, pp 273–288
10. Devi MP, Choudhry MD, Boopathi Raja G, Sathya T (2022) A roadmap towards robust IoT-enabled cyber-physical systems in cyber industrial 4.0. In: Handbook of research of internet of things and cyber-physical systems, Apple Academic Press, pp 293–313
11. Kocabas O, Soyata T, Aktas MK (2016) Emerging security mechanisms for medical cyber physical systems. IEEE/ACM Trans Comput Biol Bioinf 13(3):401–416
12. Qiu H, Qiu M, Liu M, Memmi G (2020) Secure health data sharing for medical cyber-physical systems for the healthcare 4.0. IEEE J Biomed Health Inform 24(9):2499–2505
13. Ashish K, Chatterjee K (2023) A lightweight blockchain-based framework for medical cyber-physical system. The J Supercomput 1–29
14. Anand Singh R, Goyal SB, Bedi P, Verma C, Ionete EI, Raboaca MS (2023) 5G-enabled cyber-physical systems for smart transportation using blockchain technology. Mathematics 11(3):679
15. Zhiqiang L, Li Q, Ding D-W, Wang H (2023) Event-based fixed-time secure cooperative control for nonlinear cyber-physical systems under denial-of-service attacks. IEEE Trans Control of Netw Syst
16. Michail S, Ong MT, Sridharan RV, Nakamura J, Ohmura R, Khor JH (2019) Ultralightweight mutual authentication RFID protocol for blockchain enabled supply chains. IEEE Access 7:7273–7285
17. Chao L, He D, Kumar N, Huang X, Vijayakumar P, Raymond Choo K-K (2019) HomeChain: a blockchain-based secure mutual authentication system for smart homes. IEEE Internet of Things J 7(2):818–829
18. Deebak BD, Memon FH, Khowaja SA, Dev K, Wang W, Qureshi NMF, Su C (2022) Lightweight blockchain based remote mutual authentication for AI-empowered IoT sustainable computing systems. IEEE Internet of Things J
19. Kumar M, Iqbal A, Kumar P (2016) A new RGB image encryption algorithm based on DNA encoding and elliptic curve Diffie-Hellman cryptography. Signal Process 125:187–202
20. Perumal AM, Nadar ERS (2021) Architectural framework and simulation of quantum key optimization techniques in healthcare networks for data security. J Ambient Intell Humanized Comput 12:7173–7180
21. Alzubi JA, Alzubi OA, Beseiso M, Budati AK, Shankar K (2022) Optimal multiple key-based homomorphic encryption with deep neural networks to secure medical data transmission and diagnosis. Expert Syst 39(4):e12879
22. Kiran P, Parameshachari BD (2022) Resource optimized selective image encryption of medical images using multiple chaotic systems. Microprocess Microsyst 91:104546
23. Li C, Dong M, Li J, Gang X, Chen X-B, Liu W, Ota K (2022) Efficient medical big data management with keyword-searchable encryption in healthchain. IEEE Syst J 16(4):5521–5532

Heart Disease Diagnosis Detection Based on Chest Pain Using Machine Learning Stacking Classifier

Jalpa Shah, Chintan Shah, and Premal Patel

Abstract One of the most common illnesses today is heart disease, and therefore, it is essential for those who work in healthcare to work with their patients to find ways to protect their health and prolong their lives. In this study, the performance of various classifiers was examined in order to correctly identify the guts disease dataset and/or predict the cases of heart condition using minimal criteria. Healthcare professionals have accumulated a lot of knowledge, including some private data. Taking wise decisions is made easier by this data collection. In this scenario, a heart disease prediction system (HDPS) is built to predict the risk level of gastroenteritis applying K-nearest neighbor, logistic regression, decision tree, support vector machine, and random forest classifier methods.

Keywords Heart disease · Support vector machine · Accuracy · Decision tree

1 Introduction

The most vital and fundamental component in the human body is the heart. Because the heart has a number of issues, the examination of heart vaticination has to be correct. To develop a solution, online research in this field is required. Finding thorough algorithms for situations like vaticination is necessary because these circumstances are often prognosticated at the very end, which is the main reason why cardiac instances end in mortality due to a lack of correctness. One of the most comprehensive and effective knowledge bases for machine techniques is one that was created using Python and its libraries for specialized training and testing. In accordance with this training, test should be conducted on a variety of various types of requirements in line with the proper algorithms. The software learns by gathering data and expertise.

J. Shah (✉) · C. Shah · P. Patel
Department of Computer Engineering, College of Technology, Silver Oak University, Ahmedabad, Gujarat, India
e-mail: Shahjalpa.comp@gmail.com

P. Patel
e-mail: premalpatel.ce@socet.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 782, https://doi.org/10.1007/978-981-99-6568-7_55

Testing and training can both benefit from machine literacy. Complaint (CVD) [1, 3, 4], also known as disease-related to the guts, refers to a number of illnesses that affect the digestive tract in conjunction with the complaint. One of the circumstances where the body's plasma is pumped and circulated is cardiomyopathy. Other conditions include coronary artery complaints and cardiovascular health issues. Giving an opinion is a significant task that must be well fulfilled. Most of this is done under the instruction of a croaker. As a result, people with incorrect diagnoses end up undergoing expensive therapy. Therefore, we infer that an automated vaticination and opinion method would be quite helpful [2, 5].

Cardiovascular diseases reduce a person's longevity in the modern day. The World Health Organization (WHO) therefore established objectives for the avoidance of non-communicable illness (NCDs) in 2013, with a 25% relative decrease in cardiovascular illnesses as the main target. At least 50% of individuals with cardiovascular illnesses are anticipated to get access to the required drugs and counseling by 2025. A 17.9 million individuals worldwide passed away in 2016 from cardiovascular diseases alone, making up 31% of all fatalities. The detection of heart disease is a big barrier. It might be hard to tell if someone has a heart condition or not. Although there are technologies that can predict the risk that an individual would get heart disease, these tend to be expensive or in efficient [1–3] .

Only 67 heart disorders can be identified by medical professionals, based on the World Health Organization (WHO), so there is a great need for study in this field. In India, it is extremely challenging to locate skilled medical professionals and institutions in rural areas. Only 58 among the croakers for urban areas and 19 for the countryside have medical degrees, based to a 2016 WHO survey. Medical expertise faces a significant difficulty when it comes to diagnosing heart disorders in people. In this case, machine literacy might be a good solution. Heart issues can be predicted using a variety of techniques, including decision trees, neural networks, and K-NNs [5, 7].

There are several different types of cardiac disease, according to CP readings [9].

1. **Typical Angina** is a type of chest pain caused by reduced blood flow to the heart.
2. **Atypical Angina** pectoris does not have associated classical symptoms of chest pain.
3. **Non-Angina or Non-cardiac** chest pain brought on by heart illness, angina, tends to be compared to discomfort in the chest. Particularly behind your sternum, it feels like a severe squeezing, tightness, pressure, or heaviness in your chest. You can feel it in the centering, on the left, or on the right.
4. **Asymptomatic** if a patient tests as a carrier for a disease or infection but experiences no symptoms.

2 Methodology

According to the World Health Organization, 12 million deaths per year worldwide are caused by cardiac issues. In the US and other advanced nations, cardiovascular illnesses are the leading cause of death. In high-risk settings, early cardiac condition prognostication can help in making judgments on significant life changes, which in effect assists in preventing complications. This inquiry attempts to determine the major risk factors for heart complaints as well as project the overall level of danger by using logistic retrogression data processing. In statistics, a technique known as stacking classifier is used to extract a dependent variable that is categorical from a set of predictors. A double is usually used as a dependent variable for a stacking classifier. Logistic retrogression is used extensively in vaticination and success estimation [6, 9].

Figure 1 displays the dataset from a cardiac study that is now being carried out on citizens of the Massachusetts city of Framingham. The bracket challenge entails determining whether there is a 10-year risk of fetal coronary heart disease (CHD). The issue has become less clear in recent years, in part because of worries regarding heart failure, a lack of systematic blood pressure audio recordings before to the start of coronary artery disease and its management, and a lack of systematic visualization of the easily observable epicedial coronary highways. Smoking causes real quick damage to the heart and blood arteries, but for the majority of smokers who quit, the harm is quickly restored. Since many smokes nowadays also harm the heart, protecting your heart is the only method that has been shown to work. To stop smoking is to stop its benefits. Proteins has also been related to enhanced blood vessel contraction, according to researchers, who also believe that elevated blood sugar (glucose) produces stronger blood vessel contraction. Figure 5 illustrates the findings that could result in fresh therapies to ease problems following a heart attack or stroke. The risk of acquiring cardiovascular or heart diseases generally triples with each decade of age, making it the most significant risk factor. Coronary adipose stripes can start to develop in early adulthood. Approximately 82% of those who pass away from coronary heart disease are 65 years of age or older. Stroke risk currently doubles every ten years beyond the age of 55 [4–6] (Fig. 2).

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.

The block diagram of the suggested system for predicting heart disease is shown in Fig. 3. The training dataset is the first step in the block diagram, and the test data is the second step.

1. **Naïve Bayes Classification**
   For sentiment analysis tasks, the Nave Bayes classification method frequently serves as the starting point. The Naive Bayes technique's fundamental tenet is to use the combined probabilities of words and classes to determine the probabilities of classes assigned to texts. Let us look at math for a moment.

   Given the Bayes theorem: $P(A|B) = \underline{(B|A)\ P(A)}\ P(B)$.

	age	sex	cp	trestbps	chol	fbs	...	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	...	0	2.3	0	0	1	1
1	37	1	2	130	250	0	...	0	3.5	0	0	2	1
2	41	0	1	130	204	0	...	0	1.4	2	0	2	1
3	56	1	1	120	236	0	...	0	0.8	2	0	2	1
4	57	0	0	120	354	0	...	1	0.6	2	0	2	1
5	57	1	0	140	192	0	...	0	0.4	1	0	1	1
6	56	0	1	140	294	0	...	0	1.3	1	0	2	1
7	44	1	1	120	263	0	...	0	0.0	2	0	3	1
8	52	1	2	172	199	1	...	0	0.5	2	0	3	1
9	57	1	2	150	168	0	...	0	1.6	2	0	2	1
10	54	1	0	140	239	0	...	0	1.2	2	0	2	1
11	48	0	2	130	275	0	...	0	0.2	2	0	2	1
12	49	1	1	130	266	0	...	0	0.6	2	0	2	1
13	64	1	3	110	211	0	...	1	1.8	1	0	2	1
14	58	0	3	150	283	1	...	0	1.0	2	0	2	1
15	50	0	2	120	219	0	...	0	1.6	1	0	2	1
16	58	0	2	120	340	0	...	0	0.0	2	0	2	1
17	66	0	3	150	226	0	...	0	2.6	0	0	2	1
18	43	1	0	150	247	0	...	0	1.5	2	0	2	1
19	69	0	3	140	239	0	...	0	1.8	2	2	2	1

Fig. 1 Dataset distribution [4, 8]

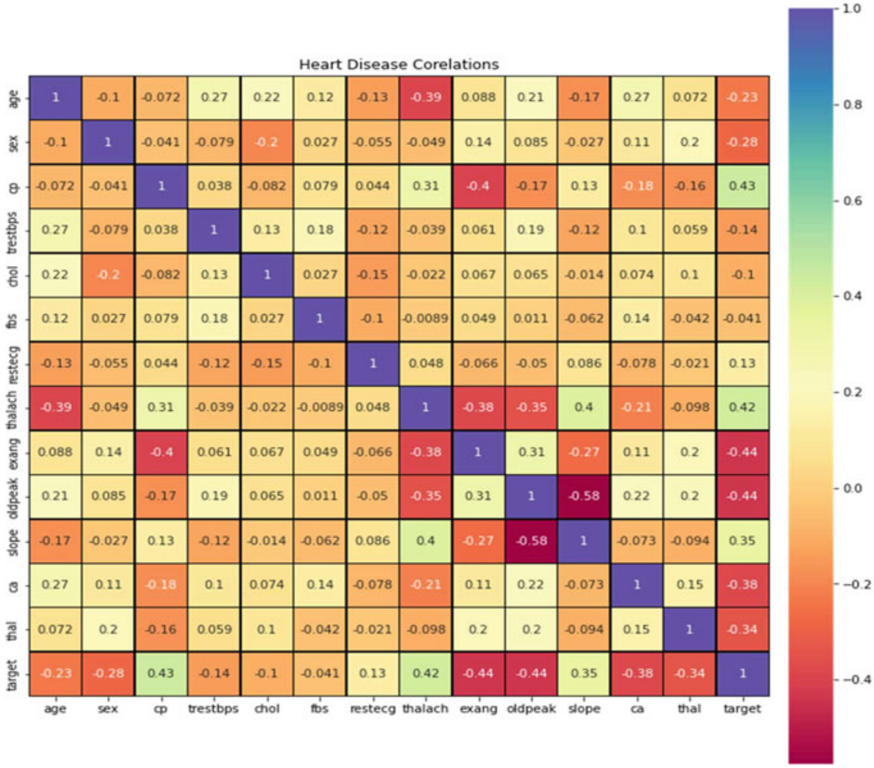


Fig. 2 Correlation matrix of the model

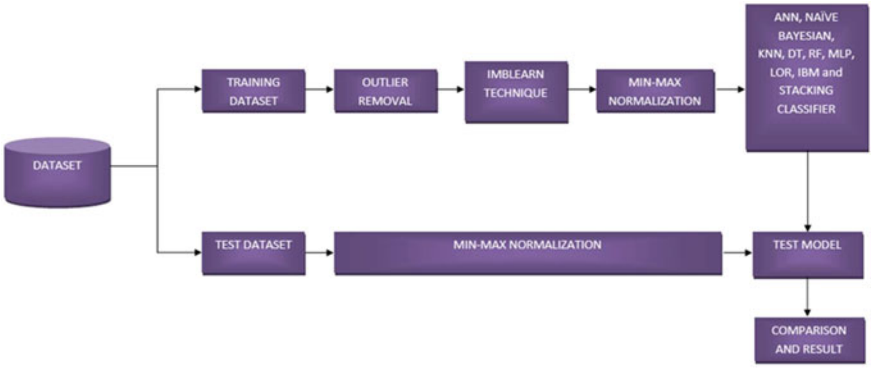


Fig. 3 Block diagram [11]

A comparable theorem will be applied to carry out the bracket: P (A) is the likelihood of the situation of component A, P (B) is a likelihood of situation of element B, P (A|B) is the estimated likelihood of component A given that element B occurs. Calculations are made for elements A and B and their probabilities under the scenario P (X). Therefore, given independent qualities, the aforementioned theory would directly sum the probability of each point passing [2–4].

2. **Dataset Information**
The dataset is called heart.csv. This dataset contains 303 cases, all of which are either healthy individuals or those who have heart issues. Among the 303, 165 (54.45) cases involve persons who have a heart condition, whereas 138 (45.54) involve people who do not. There are fourteen qualities. The dataset will not contain any null values or missing values.
A few of the characteristics include age, coitus, type of casket pain, resting levels of lipids, blood sugar levels position, electrocardiogram (EC outcome), greatest heartbeat achieved, exercise-induced angina, the ST depressive disorders, elevated ST, a member pitch, the number of major blood vessels, and heart impairment as of 3-normal, 6-fixed, and 7-reversible disfigurement [4, 8].
3. **KNN Methods**
K-NN is K-nearest neighbor. A data point whose classification is unavailable is selected for the K-NN algorithm, after which the number of neighbors, k, is established. On the basis of the smallest distance measured by Euclid between the points that were chosen and their neighbors, the next step is to select k neighbors. After that, the selected data point is placed in the category that includes the large majority of the K-neighbor's neighbors [2–4].
4. **Random Forest Methods**
The information that is trained in the random forest (RF) technique is used to build several decision trees. Every decision tree predicts a class to be a relationship, and in the scenario of the brackets, the class which is the affairs of the most decision trees is selected as the ultimate result. We must specify how many trees

we want to build before we can use this algorithm. A bootstrap aggregating or bagging method is random forest. This method is employed to reduce friction in the outcomes.

The algorithms used for all classification and regression are random forest. Predictions are made based on the information's classification into a tree. When used to enormous datasets, the random forest technique can still deliver the same results even when a sizable portion of record values are missing. The samples produced by the decision tree can be saved and applied to various datasets. A random forest has two stages: the first is the generation of the random forest, and the second is the prediction using the classifier created in the first step [2–4].

5. **Decision Tree**
The core node of the decision tree method represents the characteristics of the dataset, and the lateral branches will achieve a certain goal. Decision trees are employed due to their speed, dependability, clarity, and minimal data preparation requirements. The predicted class label is determined by the decision tree's root. Both the value of the core attribute and the record's attribute are assessed [2–4].
6. **Stacking Classifier**
A technique for putting together regression or classification models is arranging them, which includes two-layer estimators. The first layer contains all of the basic models that are employed to predict the outcomes on the test datasets. All of the forecasts from the baseline models are entered into the Meta-classifier or regressor in the subsequent level to generate fresh predictions.
We will specifically assess the following algorithms [2–4].

- Multi-layer perceptron
- Adaboost classifier
- Decision tree.
- Extra tree classifier

3 Results

Our primary objective in doing this study was to use several machine learning algorithms to predict heart disease. K-nearest neighbor (K-NN), ANN, DT, RF, IBM, MLP, LOR [2–4] and stacking classifier were used to make our predictions. Stacking classifier provides superior results with 99.96% accuracy. We have noted the following accuracy, precision, sensitivity, and specificity for each algorithm (Figs. 4, 5, 6 and 7).

4 Comparison Chart

See Figs. 8 and 9.

```
STACKING CLASSIFIER Accuracy is :91.00%

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       0.87      0.97      0.92       117
           1       0.92      1.00      0.96        33
           2       1.00      1.00      1.00        66
           3       0.00      0.00      0.00        17

    accuracy                           0.91       233
   macro avg       0.70      0.74      0.72       233
weighted avg       0.85      0.91      0.88       233
```

Fig. 4 Without SMOTE and CP based stacking classifier

```
STACKING CLASSIFIER Accuracy is :99.00%

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       0.98      1.00      0.99       117
           1       1.00      1.00      1.00        33
           2       0.98      0.97      0.98        66
           3       1.00      0.94      0.97        17

    accuracy                           0.99       233
   macro avg       0.99      0.98      0.98       233
weighted avg       0.99      0.99      0.99       233
```

Fig. 5 With SMOTE and CP based stacking classifier

5 Conclusion

The maximum number of heart conditions is only limited by the current script. Heart discomfort is difficult, and this complaint frequently results in many fatalities. Manually calculating the likelihood of developing a cardiac problem based on hazard factors that have already been identified is delicate. This work's primary focus on

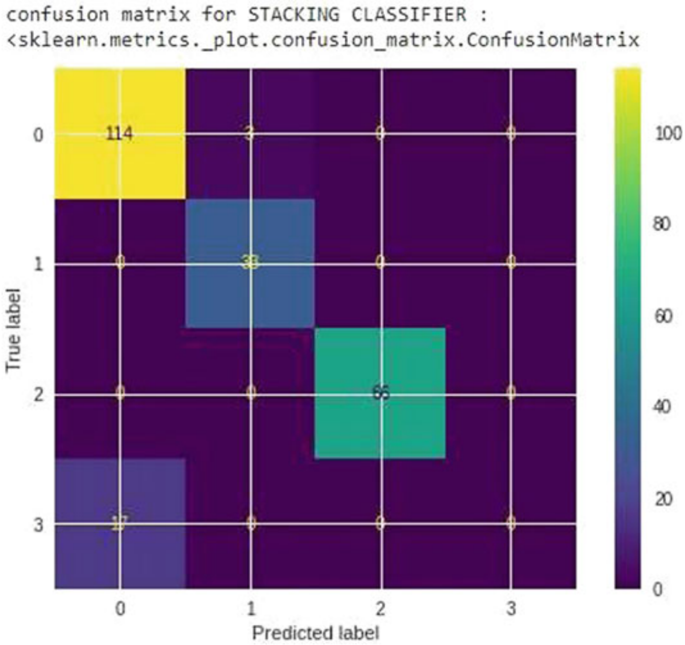


Fig. 6 Without SMOTE and CP based confusion matrix

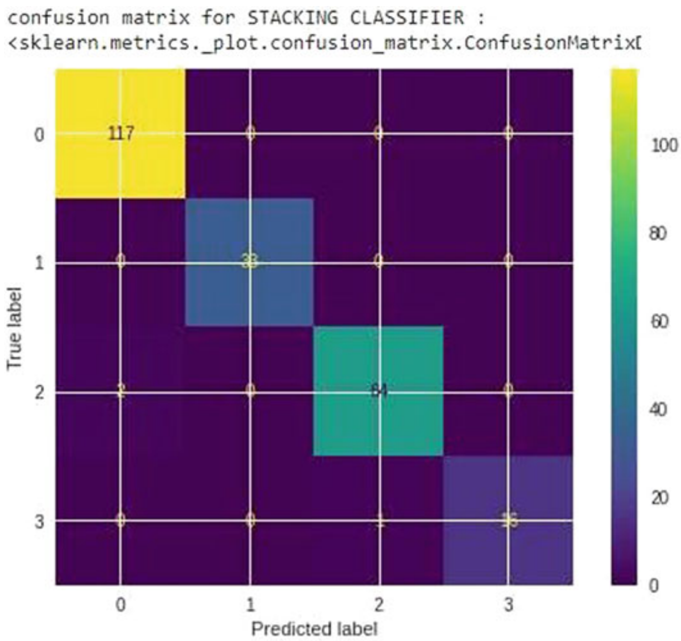


Fig. 7 With SMOTE and CP based confusion matrix

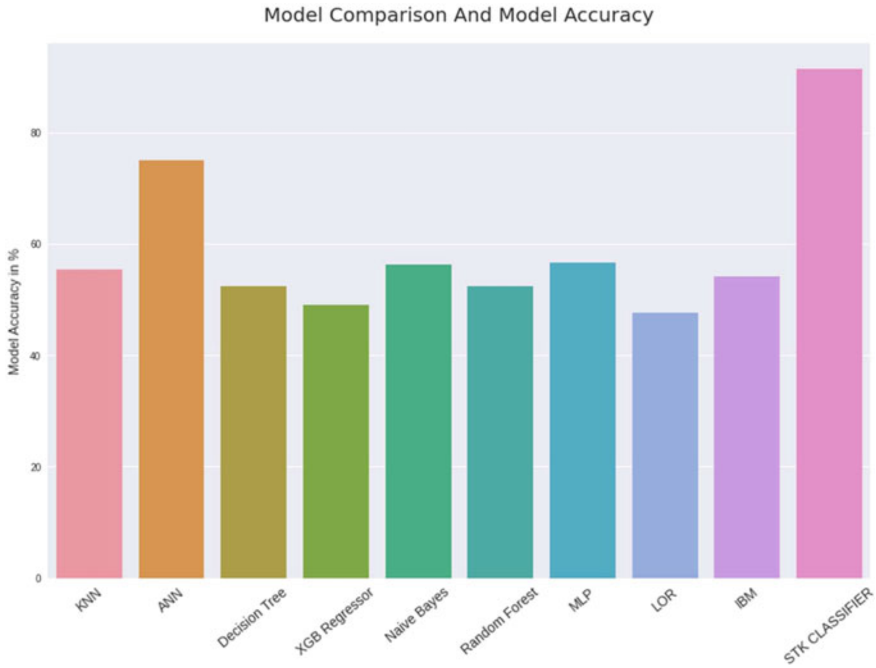


Fig. 8 Without SMOTE and CP based comparison chart

the execution of categorization algorithms and approaches for cardiac complaints is one of its biggest flaws. In order to employ machine learning techniques in the logistic retrogression techniques to determine if a patient has a cardiac problem or not, we have studied colorful data cleaning and mining approaches that prepare and create a dataset suitable for data mining. This software simplifies the intricacy of the croaker's waiting periods by enabling any medical expert to anticipate heart problems. It is nevertheless possible to strengthen the delicacy and enforce heart complaint vaticination. After employing interesting algorithms, it can be said that machine literacy is proving to be quite helpful in predicting heart complaints, which are currently one of society's most prevalent difficulties. As more and more research is done in the area of machine literacy, new approaches to use machine learning in the healthcare industry may soon be developed. Utilizing the relevant attributes, the algorithms utilized in this trial have demonstrated good performance. Eventually, the conclusion can be reached. By predicting heart complaints, machines can lessen the harm done to individuals' physical and mental health. Therefore, in this study, the stacking classifier model with SMOTE and CP value yielded the best accuracy, 99% forecast the accuracy, and the stacking classifier model without SMOTE and CP value yielded 91% accuracy to identify the risk of heart disease based on CP value [3, 5–7, 11].

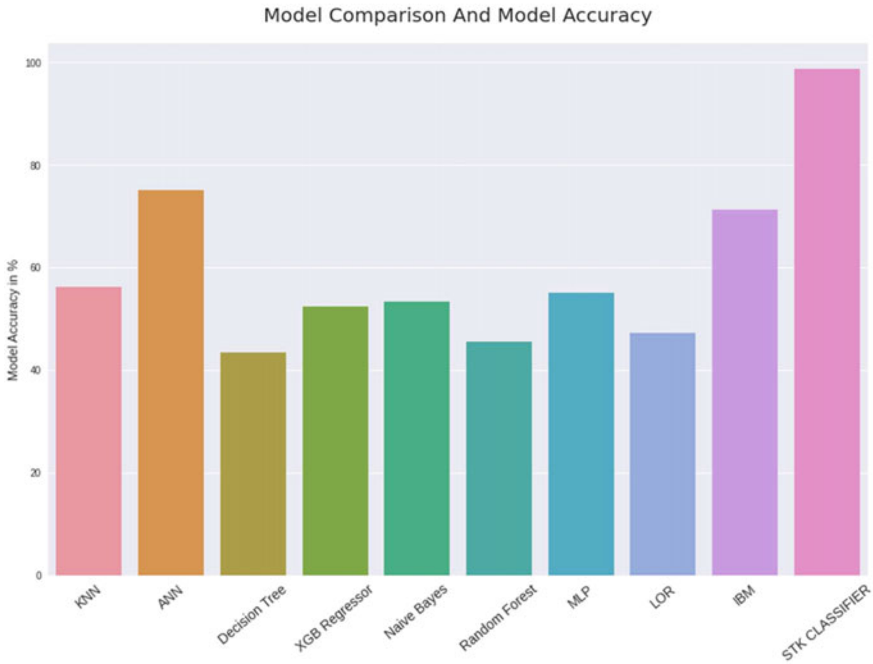


Fig. 9 With SMOTE and CP based comparison chart

References

1. Golande A, Pavan Kumar T (2019) Heart complaint vaticination using effective machine learning ways
2. The Lancet Global Health (2018) The changing patterns of cardiovascular conditions and their threat factors in the countries of India The global burden of complaint study 1990–2016
3. Sharma H, Rizvi MA (2022) Vaticination of heart complaints using machine literacy algorithms Acheck. World health ranking
4. Bharti S (2022) Analytical study of heart complaint vaticination compared with different algorithms. In: International conference on computing, communication, and robotization (ICCA 2022)
5. Gandhi M (2022) Vaticination in heart complaint using ways of data mining. In: International conference on futuristic trend in computational analysis and knowledge operation (ABLAZE-2022)
6. Babu S (2017) Heart complaint opinion using datamining fashion. In: International national conference one lectronics, communication and aerospacetechnology (ICECA2022)
7. Chen AH (2011) HDPS heart complaintv at icination system
8. Grover SA, Lowensteyn I, Esrey KL, et al. (1995) Docroakers directly assess the coronary threat in their cases. Primary results of the Coronary Health Assessment Study
9. Friedmann PD, Brett AS, Mayo- Smith MF (1996) Differences in generalists and cadiologist's comprehension so cardiovascular threat and the issues of preventative remedy in the cardiovascular complaint
10. Meland E, Laerum E, Lehmann EH (2021) Assessment of coronary heart complaint threat, IIaclin icalmulti-center study of general interpreters threat assessment

11. Mohan S, Thirumalai C, Srivastava G (2021) Effective heart complaint vaticinationu singmodern machine-learning ways