Lecture Notes in Electrical Engineering 903

Sudhan Majhi
Rocío Pérez de Prado
Chandrappa Dasanapura Nanjundaiah *Editors*

# Distributed Computing and Optimization Techniques

## Select Proceedings of ICDCOT 2021

Springer

# Lecture Notes in Electrical Engineering

## Volume 903

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

**China**

Jasmine Dou, Editor (jasmine.dou@springer.com)

**India, Japan, Rest of Asia**

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

**Southeast Asia, Australia, New Zealand**

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

**USA, Canada:**

Michael Luby, Senior Editor (michael.luby@springer.com)

**All other Countries:**

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at https://link.springer.com/bookseries/7818

Sudhan Majhi · Rocío Pérez de Prado ·
Chandrappa Dasanapura Nanjundaiah
Editors

# Distributed Computing and Optimization Techniques

## Select Proceedings of ICDCOT 2021

Springer

*Editors*
Sudhan Majhi
Electrical Engineering
Indian Institute of Technology Patna
Patna, Bihar, India

Rocío Pérez de Prado
Telecommunication Engineering
University of Jaén
Jaén, Spain

Chandrappa Dasanapura Nanjundaiah
Electronics and Communication
SJB Institute of Technology
Bengaluru, Karnataka, India

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-19-2280-0 ISBN 978-981-19-2281-7 (eBook)
https://doi.org/10.1007/978-981-19-2281-7



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

# Preface

This volume of the conference proceedings encompasses top-notch papers presented at the International Conference on Distributed Computing and Optimization Techniques (ICDCOT–2021), which took place on the 25th and 26th of June 2021, at the SJB Institute of Technology, Bangalore, Karnataka, India.

## Mode of Conference

In view of the COVID-19 pandemic, a complete lockdown was imposed by the Indian Government. Public gatherings were strictly restricted and treated as a high offence by the task force not only within the country but world over. Travel restrictions were completely banned worldwide. Therefore, in order to abide by the government rules and to not risk the lives of the public, the conference was held online through "Microsoft Teams". The platform conveniently provided the facilities of seamless and uninterrupted interactions amongst all the conference participants which involved researchers and speakers from different parts of the world coming together under one roof to conduct the event as per planned to avoid postponement. Conducting it offline was not a viable option as it would be possible only after the pandemic, causing a long delay, which can be easily prevented through a virtual meeting. We are delighted to announce that all the conference proceedings were successfully executed through the will and cooperation of all its organizers, hosts, participants and all other contributors.

International Conference on Distributed Computing and Optimization Techniques (ICDCOT–2021) is the premier conference, which aims to provide an opportunity to exchange valuable insight about the advanced trends in communications, signal processing, power engineering and VLSI designs. This peer-reviewed conference provides the forum for the participators, scholars and industry experts to discuss, debate and share their innovative ideas in advance trends in Electronics, Electrical and Communication Engineering. Field experts share their solutions for the complex

problems that occur in various fields/domains. This year ICDCOT–2021 Proceedings received an overwhelming global response from countries across the world. The presentations consisted of novel and original works undertaken in the areas of Electronics, Electrical and Communication Engineering as well as their applications by experts and budding researchers, across the world. This conference provided opportunities for researchers to exchange new ideas and application experiences to establish research relations and to find global partners for future collaboration. The presentations given at the proceedings encompassed a wide variety of research topics that majorly involved but were not limited to the following domains:

- 4G;5G, 802.16 & WiMAX
- Antenna design (lower band/higher band)
- Design of controllers using electrical circuits
- Image, signal and video processing
- Micro-electromechanical system (MEMS)
- Optical communication
- Renewable energy

The conference has received papers from countries such as Iraq, Egypt, Bangladesh, Belgium, as well as from various parts of India: Karnataka, Tamilnadu, Andhra Pradesh, Uttarakhand, Jalandhar, Rajasthan, Mathura and Maharashtra. We received a total of 332 papers for the conference. All the papers have been subjected to a thorough peer review by at least two referees, until each of the papers levelled the quality expected at the conference. A total of 73 papers had received the final acceptance for submission at the conference. Overall, the proceedings consisted of a total of eight sessions, spanned over two days. Each author was given a time of 15 minutes under which both paper presentation and the Q & A discussion had to be completed.

[Day 1]: 25/06/2021

[11:30 AM – 1:30 PM] = Speech given by our honourable speaker, Dr. Sudhan Majhi, followed by two sessions dedicated to paper presentations [Session 1 & 2], conducted parallely.

[2:30 PM – 4:30 PM] = Speech given by our honourable speaker, Dr. Gabriella Casalino, followed by two sessions dedicated to paper presentations [Session 3 & 4], conducted parallely.

[Day 2]: 26/06/2021

[10:15 AM – 1:00 PM] = Speech given by our honourable speaker, Dr. Mincong Tang, followed by a half an hour break, after which two sessions dedicated to paper presentations [Session 1 & 2], were conducted parallely.

[2:15 PM – 4:30 PM] = Two sessions of the proceedings were conducted, parallely. A total of 13 papers were presented over the span of two sessions, with nine papers presented in [Session 3] and ten papers presented in [Session 4]. This was followed by a speech given by our honourable speaker, Dr. Tu N. Nguyen.

Patna, India — Sudhan Majhi
Jaén, Spain — Rocío Pérez
Bengaluru, India — Chandrappa Dasanapura Nanjundaiah

**Acknowledgements** On behalf of the Departments of Electronics and Communication Engineering and Electrical and Electronics Engineering, SJB Institute of Technology, Bangalore, Karnataka, India, we would like to thank all our trustees, invited speakers, session chairs, conference coordinators, convener, the organizing committee, members of the discussion panels and presenters, for making the conference such a great success.

Some notable guests of ICDCOT–2021:
Speakers of the conference:
Dr. Sudhan Majhi, Professor, IIT-Patna, India
Dr. Gabriella Casalino, University of Bari, Bari, Italy
Dr. Mincong Tang, Beijing Jiaotong University, China
Dr. Tu N. Nguyen, Director of the Intelligent Systems Laboratory (ISL) in the Department of Computer Science at Kennesaw State University, Georgia, USA

# List of Editorial Board Members

# Contents

# About the Editors

**Sudhan Majhi** is working as Associate Professor in the Department of Electrical Engineering at the Indian Institute of Technology (IIT) Patna. He received a Ph.D. degree from Nanyang Technological University, Singapore, in 2008. His research interest areas are signal processing for wireless communication, blind signal classification, blind signal synchronization, blind parameter estimation, secrecy capacity of cognitive radios and cooperative communications, MIMO, OFDM, MIMMO-OFDM, SC-FDMA, NOMA, UWB systems, receiver design and implementation on the testbed, and sequence design for wireless communication. Dr. Majhi has published over 90 papers in national and international journals.

**Rocío Pérez de Prado** received his M.S. degree in telecommunication engineering from the University of Seville, Spain, in 2008, and a Ph.D. degree in telecommunication engineering with European Mention from University of Jaén, Spain, in 2011. She is currently Associate Professor (Professor Titular de Universidad) at the Telecommunication Engineering Department, University of Jaén, Spain. Her current research interests include artificial intelligence, machine learning, grid computing, cloud computing and optical fibre. She has authored over 45 publications and has been an editor of the Journal Citation Report (JCR)-indexed journal Applied Soft Computing and Energies since 2014 and 2020, respectively. Also, she belongs to the reviewer board of more than 25 JCR-indexed journals and participated as a technical program committee (TCP) member in 30 international conferences in the field of artificial intelligence and cloud computing. Besides, she has collaborated in nine research projects and contracts and is currently serving as an external evaluator of the National Committee of Science and Technology and Technology Innovation (CONCYTEC) in Pérú.

**Chandrappa Dasanapura Nanjundaiah** is Professor in the Department of Electronics and Communication Engineering, SJB Institute of Technology, Bangalore, India. Dr. Chandrappa has over 20 years of experience in the field of engineering education. He received his Ph.D. from Visvesvaraya Technological University, India. His research interests include digital image processing, embedded system, electronics circuit design and power electronics. He is supervising five research scholars for their Ph.D. thesis. He has over 35 research publications in refereed international journals and conferences.

## List of Contributors

**Abdullah Hanifa** University of South Africa, Pretoria, South Africa

**Abood Amany Mohammad** Computer Engineering Techniques Department, Al Esraa University College, Baghdad, Iraq

**Agarwal Unnati** Department of Information Technology, National Institute of Technology Raipur, Raipur, India

**Agieb Ramy Said** Faculty of Engineering, Department of Electrical Engineering, MTI University, Cairo, Egypt

**Ahammed Md. Sobuj** Department of Computer Science and Engineering, Khwaja Yunus Ali University, Sirajganj, Bangladesh

**Ahmed Mohammed Riyaz** REVA University, Bangalore, India

**Amaranageswarao Gadipudi** Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli, India

**Amarsingh Feroz C.** Department of ECE, Francis Xavier Engineering College, Tirunelveli, India

**Anila M.** Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

**Anita J. P.** Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

**Annepu Arudra** Rajiv Gandhi Institute of Technology, Bangalore, Karnataka, India

**Anuradha T.** Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, AP, India

**Aparanji V. M.** Siddaganga Institute of Technology, Tumkur, Karnataka, India

**Archana N. V.** Electrical and Electronics Department, NIE Institute of Technology, Mysore, Karnataka, India

**Aron Rajni** SVKM's Narsee Monjee Institute of Management Studies (NMIMS) University, Mumbai, Maharashtra, India

**Asirvatham M.** Department of ECE, SCAD College of Engineering and Technology, Tirunelveli, India

**Azad Md. Selim** Department of CSE and CSIT, Shanto Mariam University of Creative Technology, Dhaka, Bangladesh

**Azad Mir Mohammad** Department of Computer Science and Engineering, Khwaja Yunus Ali University, Sirajganj, Bangladesh

**Azam Farooque** Department of Computer Science and Engineering, Sangam University, Bhilwara, Rajasthan, India

**Babu D. Vijendra** Department of ECE, Aarupadai Veedu Institute of Technology, Chennai, India

**Balaji G. N.** Department of ECE, SRM TRP Engineering College, Trichy, India

**Bapu B. R. Tapas** ECE, S.A. Engineering College, Chennai, Tamil Nadu, India

**Bhardwaj Kapil** NIT Jamshedpur, Jamshedpur, India

**Bhargav Madhavarapu Srinivasa Sai** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

**Bhoomika C. M.** REVA University, Bangalore, India

**Bhuvaneswari P.** Department of ECE, Sri Venkateswara College of Engineering and Technology, Chittoor, Andhra Pradesh, India

**Blessy P. Hannah** Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli, India

**Bosquez Carlos** GISTEL (Telecommunications Systems Research Group), Salesian Polytechnic University, Guayaquil, Ecuador

**Butale Vaishnavi** K. K. Wagh Institute of Engineering Education and Research, Nashik, India

**Chandra Budhi Veera Bharath** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

**Chandran K. Prabhu** Department of ECE, Sreenivasa Institute of Technology and Management Studies, Chittoor, Andhra Pradesh, India

**Chandrappa D. N.** Department of Electronics and Communication Engineering, SJB Institute of Technology, BGS Health and Education City, Bangalore, India

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

**Chandrasekaran Gokul** Department of Electronics and Communication Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

**Chandrika C. P.** M S Ramaiah Institute of Technology, Bangalore, India

**Charan Shree** Department of Electronics and Communication Engineering, SJB Institute of Technology, Bengaluru, India

**Dani Virendra** Computer Science and Engineering Department, Shivajirao Kadam Institute of Technology and Management, Indore, India

**Darney P. Ebby** Department of EEE, SCAD College of Engineering and Technology, Tirunelveli, India

**Dasanapura Nanjundaiah Chandrappa** Department of ECE, East West College of Engineering Yelahanka, Bangalore, India

**Deevi Hari Krishna** Sri Mittapalli College of Engineering, Guntur, Andhra Pradesh, India

**Deivalakshmi S.** Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli, India

**Devika B.** Department of TCE, KSIT, Bengaluru, India

**Dhanunjaya Babu P.** MCA Department, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Dhole Avinash** Department of Computer Science and Engineering, Raipur Institute of Technology, Chhatouna, Chhattisgarh, India

**Divyashree B.** M.S. Ramaiah University of Applied Sciences, Bangalore, Karnataka, India

**Diwakar Manoj** CSE Department, Graphic Era University (Deemed), Dehradun, India;
Department of CSE, Graphic Era Deemed to Be University, Dehradun, India

**Faisal Zainab G.** Computer Engineering Techniques Department, Al Esraa University College, Baghdad, Iraq

**Friska J.** Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli, India

**Gaikwad Vinayak** MPSTM & E, NMIMS, University, Mumbai, India

**Ganesh Chattu Sai** MCA Department, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Geervani Bandi** Vardhaman College of Engineering, ECE, JNTU Hyderabad, Hyderabad, India

**Geethesh T. G.** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Goswami Pooja** Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

**Goud Kalal Hanush** Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India

**Grewal Naseeb** Indian Institute of Technology Roorkee, Roorkee, India

**Gurram Rajendra** Department of Information Technology, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Guthikinda Likhitha** Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India

**Harold Robinson Y.** School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India

**Harshak Krishnaa K.** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Hussein Maysam Sameer** Computer Engineering Techniques Department, Al Esraa University College, Baghdad, Iraq

**Islam Md. Tarequl** Department of Computer Science and Engineering, Khwaja Yunus Ali University, Sirajganj, Bangladesh

**Jadhav Tushar** Government Polytechnic, Malvan, India

**Jagan R.** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

**Jangam Naga Raju** Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India;
St. Peters Institute of Higher Education and Research, Avadi, Chennai, India

**Jayakumar D.** Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India;
Department of CSE, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India

**Jayanth Reddy A.** Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

**Jayanthi K.** Government College of Engineering, Salem, Tamil Nadu, India

**Jayaprasad Madalai** CMR University School of Engineering, Bangalore, Karnataka, India

**Jeyabharathi M.** Government College of Engineering, Salem, Tamil Nadu, India

**Joshi Priyanka** NIT Jamshedpur, Jamshedpur, India

**Jugale Abhinandan Ajit** REVA University, Bangalore, India

**Juluri Varsha** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Ettimadai, India

**Kadirappa Ravindranath** Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli, India

**Kalaiselvi K.** Department of CSE, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

**Kallimani Jagadish S.** Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore, India

**Kalluri Siva Krishna** Sri Mittapalli College of Engineering, Guntur, Andhra Pradesh, India

**Kamaraju M.** Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India

**Kanavalli Anita** Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India

**Kannan Aiswarya** SRM TRP Engineering College, Trichy, Tamilnadu, India

**Kannan P.** Department of ECE, Francis Xavier Engineering College, Tirunelveli, India

**Karanwal Shekhar** CSE Department, Graphic Era University (Deemed), Dehradun, India

**Karthick Saran S.** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Karthikeyan C.** Department of CSE, Koneru Laksmaiah Education Foundation, Guntur, Andhra Pradesh, India

**Karthikeyan P. R.** Department of Electrical and Electronics Engineering, Velalar College of Engineering and Technology, Erode, India

**Katarya Rahul** Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science, Delhi Technological University, New Delhi, India

**Kaur Mandeep** Lovely Professional University, Jalandhar, India;
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

**Kavya K. S.** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Ettimadai, India

**Keerthana P** Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

**Kishore Nanda** Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India

**Kokate Priyanka** Computer Science and Engineering Department, Shivajirao Kadam Institute of Technology and Management, Indore, India

**Kore Sheela** Department of Electronics and Communication Engineering, KLE Dr. M S Sheshgiri College of Engineering and Technology, Belgaum, Karnataka, India

**Koushik Ch. V. N.** Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, AP, India

**Krishna Talluri Lakshmi Siva Rama** Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India

**Krishnan R. Santhana** Department of ECE, SCAD College of Engineering and Technology, Tirunelveli, India

**Kumar Aditya** Department of Computer Science and Information Technology, Mahatma Gandhi Central University, Motihari, Bihar, India

**Kumar Battula Mohan** Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India

**Kumar Kethavath Prem** Department of Computer Science and Engineering, ACE Engineering College, Hyderabad, India

**Kumar B. M. Kiran** M.S. Ramaiah University of Applied Sciences, Bangalore, 560058 India

**Kumar Mahapatra Medha Sampath** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

**Kumar Neelam Sanjeev** Department of Biomedical Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

**Kumar Sunil** Department of Computer Science and Engineering, Sangam University, Bhilwara, Rajasthan, India

**Kumar T. Rajesh** Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India

**Kumar Vipin** Department of Computer Science and Information Technology, Mahatma Gandhi Central University, Motihari, Bihar, India

**Kumari Sapna** Department of Computer Science and Information Technology, Mahatma Gandhi Central University, Motihari, Bihar, India

**Kumbla Ashwij** Electronics and Instrumentation Engineering, M S Ramaiah Institute of Technology, Bangalore, India

**Kumutha D.** SJB Institute of Technology, Kengeri, Bangalore, India

**Kunal** Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India

**Kushwah Surbhi** Computer Science and Engineering Department, Shivajirao Kadam Institute of Technology and Management, Indore, India

**Kutre Tejaswini Jayawant** Department of Electronics and Communication Engineering, KLE Dr. M S Sheshgiri College of Engineering and Technology, Belgaum, Karnataka, India

**Lakshmi Narayanan K.** Department of ECE, Francis Xavier Engineering College, Tirunelveli, Tamilnadu, India

**Lakshmi Sarvani V.** Department of Computer Science and Engineering, SaveethaSchool of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India

**Latha S.** Department of ECE, SJBIT, Bengaluru, Karnataka, India

**Madhuri S. L.** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Ettimadai, India

**Mahadik Yogesh** Institute of Chemical Technology, Matunga, Mumbai, India

**Mahantesh K.** Department of Electronics and Communication Engineering, SJB Institute of Technology, Bengaluru, India

**Malar P. Stella Rose** ECE Department, J.P. College of Engineering, Tenkasi, Tamilnadu, India

**Manivannan S. S.** School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India;
Computer Science Engineering, Sri Ramachandra Faculty of Engineering and Technology, Porur, Chennai, Tamil Nadu, India

**Manu R.** REVA University, Bangalore, India

**Marouthu Anusha** Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

**Megharjun V. N.** University of Pennsylvania, Philadelphia, PA, USA

**Mittal Ankush** Indian Institute of Technology Roorkee, Roorkee, India

**Moger Venktesh M.** Department of ECE, Siddaganga Institute of Technology, Tumakuru, Karnataka, India

**Mohan Poornima** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

**Murthy B. T. Venkatesh** Department of ECE, Siddaganga Institute of Technology, Tumakuru, Karnataka, India

**Muthamizhan T.** Sri Sairam Institute of Technology, Chennai, India

**Naragonahalli ShambuGowda Anil** Department of ECE, East West College of Engineering Yelahanka, Bangalore, India

**Nagaraju V.** ECE, Saveetha School of Engineering, Chennai, Tamilnadu, India; CSE, Saveetha School of Engineering, Tamilnadu, Chennai, India

**Nair Swathy** M.S. Ramaiah University of Applied Sciences, Bangalore, India

**Narendra Karlaputi** MCA Department, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Nasir Mostofa Kamal** Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh

**Naveen Chigurupati** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

**Nayak Jyoti** Department of Computer Science and Engineering, Raipur Institute of Technology, Chhatouna, Chhattisgarh, India

**Neel Kamal R. V.** Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, AP, India

**Negi Anvit** Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India

**Neha** St. Peters Institute of Higher Education and Research, Chennai, India

**Nijhawan Rahul** University of Petroleum and Engineering, Dehradun, India

**Nikose Mangesh D.** Sandip University, Nashik, India

**Nitin T.** Department of Electronics and Communication Engineering, SJB Institute of Technology, Bengaluru, India

**Pabboju Suresh** Chaitanya Bharathi Institute of Technology, Hyderabad, India

**Padmavathi S.** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Pallavi** St. Peter's Institute of Higher Education and Research, Chennai, India

**Parameshachari B. D.** GSSS Institute of Engineering and Technology for Women, Mysuru, India

**Parasuraman S.** Karpaga Vinayaga Engineering and Technology, Chennai, India

**Pavithra A. C.** NIEIT, Mysore, India;
Electronics and Communication Department, ATME College of Engineering, Mysore, Karnataka, India

**Patil Sujata N.** Department of Electronics and Communication Engineering, KLE Dr. M S Sheshgiri College of Engineering and Technology, Belgaum, Karnataka, India

**Peter Soosai Anandaraj A.** Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

**Pradeepini G.** Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

**Prajwal K. T.** M.S. Ramaiah University of Applied Sciences, Bangalore, India

**Pranav** Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science, Delhi Technological University, New Delhi, India

**Pranesh B.** Department of Electronics and Communication Engineering, SJB Institute of Technology, Bengaluru, India

**Prasanna A. Sai** Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

**Precila K.** Department of ECE, Francis Xavier Engineering College, Tirunelveli, India

**Priyadarshi Neeraj** CTiF Global Capsule, Department of Business Development and Technology, Aarhus University, Herning, Denmark

**Puttamadappa C.** Department of Electronics and Communication Engineering, Dayananda Sagar University, Bangalore, India

**Rachana P.** New Horizon College of Engineering, Bangalore, India

**Radha V.** Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India

**Raghavi P. Yamini** Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

**Ragunathan Thirumalaisamy** Department of Computer Science and Engineering, SRM University, Amaravathi, India

**Rahman Haaris** Electronics and Instrumentation Engineering, M S Ramaiah Institute of Technology, Bangalore, India

**Rahman Md. Wahidur** Department of Computer Science and Engineering, Khwaja Yunus Ali University, Sirajganj, Bangladesh

**Rahul B. P.** Department of ECE, Siddaganga Institute of Technology, Tumakuru, Karnataka, India

**Raj N. Pruthvi** Department of ECE, Siddaganga Institute of Technology, Tumakuru, Karnataka, India

**Rajakumar G.** Department of ECE, Francis Xavier Engineering College, Tirunelveli, India

**Rajalakshmi T. S.** Department of Mechatronics Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamilnadu, India

**Rajarajeswari S.** Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India

**Rajesh C. B.** Department of Electronics and Communication Engineering, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Rajeshwari A.** Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli, India

**Rajeshwari S. B.** Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore, India

**Raju Ayalapogu Ratna** Mahatma Gandhi Institute of Technology, Hyderabad, India

**Ramakrishna Ch.** Vardhaman College of Engineering, ECE, JNTU Hyderabad, Hyderabad, India

**Ramesh G. P.** St. Peters Institute of Higher Education and Research, Avadi, Chennai, India

**Ramesh K.** Karpaga Vinayaga Engineering and Technology, Chennai, India

**Rao S. Nagaraja** M.S. Ramaiah University of Applied Sciences, Bangalore, Karnataka, India

**Rao Ramisetty Rajeswara** University College of Engineering, JNTUK, Vizianagaram, India

**Reddy Kalyan** Department of ECE, AITT, Tirupathi, India

**Ringangonkar Archish Amar** Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India

**Robinson Y. Harold** School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India

**Sahu Devbrat** Department of Computer Science and Engineering, Raipur Institute of Technology, Chhatouna, Chhattisgarh, India

**Sahu Tirath Prasad** Department of Information Technology, National Institute of Technology Raipur, Raipur, India

**Sai Yaswanth Bysani** Department of ECE, Siddaganga Institute of Technology, Tumakuru, Karnataka, India

**Sandeep Guptha P. V.** Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

**Sandeep P.** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Ettimadai, India

**Sandeep Valavala** Department of Electronics and Communication Engineering, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Sankar Sesha S.** Department of Electronics and Communication Engineering, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Santhana Krishnan R.** Department of ECE, SCAD College of Engineering and Technology, Tirunelveli, India

**Santhoshini T.** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Ettimadai, India

**Sasikala N.** New Horizon College of Engineering, Bengaluru, Karnataka, India

**Sathyanarayana N.** Vemana Institute of Technology, Bengaluru, India

**Saxena Devansh** Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India

**Senthilnathan R.** Department of Mechatronics Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamilnadu, India

**Sharma Sonali** Department of CSE, Graphic Era Deemed to Be University, Dehradun, India

**Shashidhara D. N.** SJB Institute of Technology, BGS Health and Education City, Bangalore, India

**Shubha Rao A.** Department of ECE, SJB Institute of Technology, Bangalore, India

**Shyamala Devi M.** Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

**Sikha O. K.** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Silas Stephen D.** Panimlar Engineering College, Chennai, India

**Sindhu A.** Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India

**Sivakumar Soubraylux** Computer Science Engineering, Sri Ramachandra Faculty of Engineering and Technology, Chennai, Tamil Nadu, India

**Sophia J. Jinu** Ràjalakshmi Engineering College, Chennai, India

**Srikanth V.** Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

**Sripathy B.** Department of Mathematics, School of Arts, Humanities and Education, SASTRA Deemed to be University, Thanjavur, TamilNadu, India

**Srivastava Mayank** NIT Jamshedpur, Jamshedpur, India

**Subramanian M.** Department of ECE, SCAD College of Engineering and Technology, Tirunelveli, India

**Subramanian P.** St Peter's Institute of Higher Education and Research, Chennai, India

**Sudha P. N.** Department of ECE, KSIT, Bengaluru, India

**Sunagar Pramod** Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India

**Sundararajan S.** SCAD College of Engineering and Technology, Tirunelveli, Tamilnadu, India

**Suneja Kriti** Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India

**Suresh G. R.** Department of CSE, SRM TRP Engineering College, Trichy, Tamil Nadu, India

**Surya Prasad S.** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Swathy S.** Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Ettimadai, India

**Talasila Viswanath** Electronics and Telecommunication Engineering, Center for Imaging Technologies, M S Ramaiah Institute of Technology, Bangalore, India

**Tarun Ch.** Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, AP, India

**Tejash D. P.** Department of Electronics and Communication Engineering, SJB Institute of Technology, Bengaluru, India

**Tejaswi Adusumalli Sai** MCA Department, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Tejeswini J.** Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

**Thakre Mohan** K. K. Wagh Institute of Engineering Education and Research, Nashik, India

**Thekiya Md. Shahid** Sandip University, Nashik, India

**Tiwari Harsha** Department of Computer Science and Engineering, Raipur Institute of Technology, Chhatouna, Chhattisgarh, India

**Tripathi Rajesh Kumar** Computer Engineering and Applications, GLA University, Mathura, India

**Valencia Wilson** GISTEL (Telecommunications Systems Research Group), Salesian Polytechnic University, Guayaquil, Ecuador

**Vasumathi Devara** Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Hyderabad, India

**Veerabhadra** M.S. Ramaiah University of Applied Sciences, Bangalore, Karnataka, India

**Velammal M. Navaneetha** Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli, India

**Velu C. M.** Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India;
Department of CSE, SRM TRP Engineering College, Trichy, Tamil Nadu, India

**Venkata Lakshmi T.** Department of ECE, JNTU, Kakinada, Andhra Pradesh, India

**Venkata Thanooj K.** Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

**Vigneshwar S.** Department of Electronics and Communication Engineering, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India

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

**Vishal Linga** Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India

**Viswesh K. S.** Department of Electronics and Communication Engineering, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India

# 1 T-1D Single-Ended SRAM Cell Design for Low Power Applications Using CMOS Technology


**T. Venkata Lakshmi and M. Kamaraju**

**Abstract** In recent decades, there is an increasing demand for high-density, and Very Large Scale Integrated Circuits (VLSI). SRAM cell stability is the central issue, because of the CMOS technology scaling. In recent periods, the demand for VLSI low-power circuits is very high. SRAM is the central part of the cache, and Memory Caching (MC) seems very effective as most programs access similar data repeatedly. The major challenge in SRAM cells is balancing power and delay. CMOS-based SRAM cells face issues such as less reliability, high cost, and significant variation in parameters. In addition, the gate in CMOS devices starts to lose its control over the channel. Due to this reason, CMOS technology is used to design a 1 T-1D SRAM cell in this research paper. This work aims to reduce the leakage of power without affecting SRAM cell logic state. The proposed cell structure is simple to design also makes it exceptionally accessible and cost-effective.

**Keywords** Delay · One transistor-one diode · Power consumption · Static random access memory · Very large scale integrated circuits

## 1 Introduction

Random-access memory is classified as SRAM and DRAM. Static RAM retains data bits in its memory until power is given [1]. SRAM directs the total system performance also uses a system-on-chip area [2]. Evolving applications, for example, wireless body sensing networks, implanted medical devices, seeks the necessity of low-power SRAMs [3]. It is challenging to design low-power SRAM in sub-micrometer technology as it has some effects like leakage power and reduced design margins [4]. Periodic refreshing is not needed in SRAM [5]. While comparing with

T. Venkata Lakshmi (✉)
Department of ECE, JNTU, Kakinada, Andhra Pradesh, India
e-mail: tvlthota@gecgudlavalleru.ac.in

M. Kamaraju
Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India
e-mail: profmkr@ieee.org



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_1

DRAM, static RAM is also costly, and it offers quick data access [6]. SRAM is used as a feature of RAM D/A converter on a video card for cache memory in personal computers [7]. SRAM stores '1' and '0' bits as volatile semiconductor memory [8]. Static RAM contains three conditions of activity in particular; hold—write—read. SRAM cell comprises access transistors and double cross-coupled inverters for reading and writing the data [9]. The familiar characteristics intended for SRAM cell stability measurement are Static Noise Margin, which is well-defined in all operational modes [10]. Various 7 T, 8 T, 9 T, and 10 T SRAM cells suggested improved read SNM, hold SNM, and read–write conflicts [11]. SRAM cells function at lower voltage with minor delay [3]. SRAM is preferred because of its low power requirements, speed, low access time, reliability, and simple construction. Power consumption is decreased by reducing the supply voltage, which also lowers the transistor's threshold voltage.

## 2 Related Works

Pal et al. [12] presented a vital 7 T low-power SRAM cell. The read, write SNM was delivered with this cell by dividing data stored from the read path and feedback path removal for the period of the write process when the only one-bit line was exerted in this design. The simulation result was based on the CMOS technology revealed the dominance of new cells. Average power, cell space, static power, and power delay product were improved in this cell compared with other cells averaged employing 69%, 26%, 41%, then 68% correspondingly. Wei et al. [14] proposed technology advances required to shrinking of CMOS transistors. MOS transistors channel length was decreased as stability and power are vital concerns. Less power intake and a high power stable SRAM model were suggested in this article. At the time of switching activity, this architecture was liable for swing voltage minimization. At high-speed switching activities, the reduction of voltage swing limits dynamic power use. The stability of SRM improved by the proper width ratio. The 8 T SRAM model afforded low power besides high stable design for high power and speed devices. Chaudhuri et al. [13] proposed the evaluation of various n-T SRAM cells. The standard power and read delay utilization are investigated for various 1-bit SRAM cells named 4 T, 5 T, 6 T, 7 T, 8 T, 9 T, 10 T, 11 T. The presentation of various SRAM cell topologies is mentioned in this work. The structure of SRAM cells relies upon the size and speed of the cell. To achieve high performance in the memory structure, massive transistors were accommodated on a single chip to boost the speed of the microchip and improve the SRAM cell execution. Also, the comparison of various structures has been finished.

## 3 Proposed Work

A new SRAM cell with only one transistor and a diode (1 T-1D) has been proposed. 1 T-SRAM uses single-transistor cell storage yet encompasses bit cell with control hardware that marks the memory practically proportional to SRAM. Because of its one-transistor bit cell, 1 T—SRAM is smaller than other n-T SRAM cells also closer in size and thickness to embedded DRAM. The proposed cell design uses a gated diode to accomplish significant enhancement in reading time, write access time, and performance; however, it consumes low power than the conventional DRAMs.

The proposed 1 T-1D SRAM cell is represented in Fig. 1. It consists of one transistor and a voltage-controlled diode capacitor. The source and drain of the NMOS structure are tied together and represent the structure of the diode. Each bit cell is connected to a word line to select a cell, and bit lines are used to read and write operations. Memory arrays are built as an array of bit cells, each store 1-bit data. For each combination of address bits, the memory asserts a single word line that activates the bit cells in that row, and the bit line is initially left floating to read a cell. Then the word line is turned ON, enabling the stored value to drive the bit line to 0 or 1. The bit cell is strongly driven to the desired value to write a bit cell. Then the word line is turned ON, interfacing the bit line to the stored bit. The SRAM cell provides extensive noise margin and high speed. The proposed design includes three modes of operation. They are standby mode (the circuit is idle), read mode (the data has been requested), and write mode (updating the contents). When the circuit is in ideal mode, this is also called standby or else hold mode. SRAM bit cells are enabled through word lines on the selected row, connecting each cell to a pair of bit lines. The bit cells carry the analog signals from the enabled bit cell to the sense amplifier, converting the analog signals to digital data while reading the data. The incoming data is driven onto the bit lines and stored into the enabled bit cells through write processes.

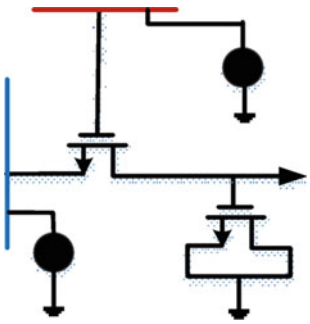


**Fig. 1** 1 T -1D SRAM cell

### 3.1 CMOS Based 1 T-1D SRAM Cell Design

The representation of the 1 T-1D SRAM cell in CMOS is shown in Fig. 2. SRAM cell based on CMOS design offers higher noise margin (NM) and stores binary information in a single bit comparing with different circuit configurations. 6 T SRAM cell design based on CMOS is popular because of its lower static power dissipation. However, the primary necessity in CMOS design is, SRAM cells have to offer high speed and superior NM. It seems complicated that even if high speed is needed and so the power leakage also gets increased.

The drawbacks noticed in CMOS design lead to the emergence of FinFET design. The cell design in CMOS operates at high speed with increased leakage current and high power dissipation. Hence, it is essential to reduce the leakage characteristics in order to boost cell stability. The diagrammatic illustration of 4 × 4 1 T-1D SRAM Cell is displayed in Fig. 3. The row address is the inputs to the row decoder. The word lines are attached to the row decoder to access the SRAM memory cell. The intersection of a row and column line represents a cell, where word lines are driven only from outside. Bit lines represent data flow in both directions. To access a cell, select a particular row and column for reading the data. Word lines lie in a horizontal direction, whereas bit lines are vertical, and the access transistor controls whether the cell should be connected to bit lines or not. Data transfer for reading as well as writing are done using the word line and bit line. Each cell is separately addressable, which can be arranged in matrix form. Most SRAM memories select a whole row of cells at any given moment and read out the data of the considerable number of cells in a row along column lines. The intersection of word line as well as bit line displays one transistor one diode SRAM memory cell. One of the vital components in SRAM is precharge circuits used to charge the bit lines to drain voltage. These circuits allow the bit line to charge high at all times except during the read and write process. A single precharge circuit is used for each column. The array of memory

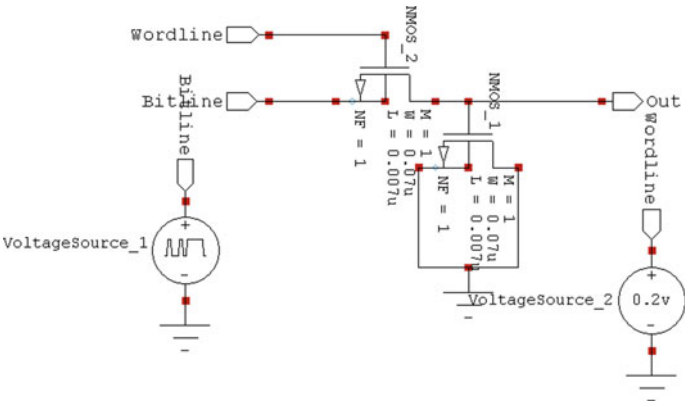



**Fig. 2** Schematic for 1 T-1D SRAM cell

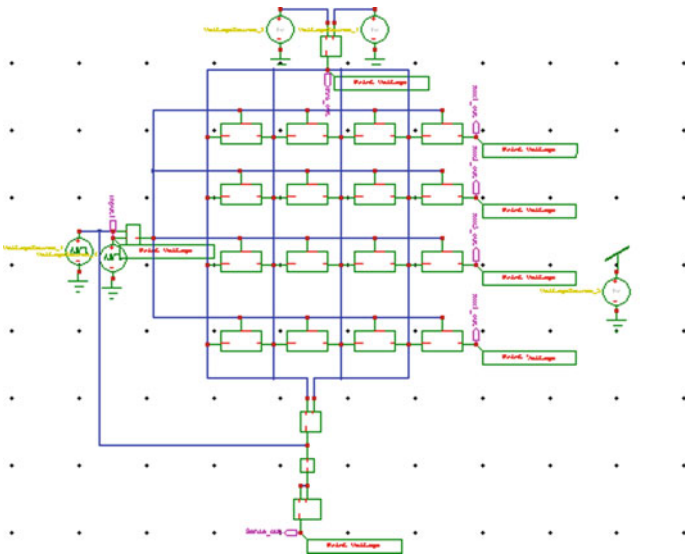


**Fig. 3.** 1 T-1D 16-bit memory array

cells consists of a diode structure in each cell for biasing the arrays unselected word lines and bit lines. If a programming pulse is applied to the selected word line and the bit line, reverse leakage of diode structures in the array may be reduced. The source and drain of a MOS device are tied together to form a diode structure. The diode is mentioned as a voltage-controlled diode capacitor that acts as variable capacitance.

The output of the row decoder is linked with a word line to select one of the word lines in a 4 × 4 memory array. The function of the row decoder is to select a particular word line depending on the inputs. The entire output of the memory array is linked with the column decoder. The column addresses are inputs to the column decoder. The source voltage of 5 V is supplied to all bit lines. If the bit lines are considered more prolonged, the diode's energy gets discharges with minimum variation in the bit line voltage. The column decoder takes the column address in binary logic and decodes it into the physical column location of the RAM array. Only one bit from one row and one column are accessed if the memory is one bit wide. For example, in an 8-bit memory, there are effectively eight identical planes of memory cells, each having the same number of rows and columns, and each will be selected when addressed. One bit is written to or read from each plane, giving the appearance of the data being 8-bits wide. Write driver circuit comprises a dual inverter circuit as well as dual NMOS transistors. These transistors drive logic 0 but not logic 1. Writing zero in BL represents the write operation. Inverters perform the conversion of logic '0' to logic '1' at other nodes. Data is transferred to bit lines if the write enable signal is activated. The sense amplifier is present at the edge of the array also bit lines are related to it. Sense amplifier senses the signal from the bit line. SA will detect which bit line is going towards high voltage also which bit line

is going towards ground potential, and afterward, a full voltage swing is acquired at the output. On each pair of bit lines, the purpose of SA is to turn weak differential signal into standard logic signal also fed to the additional data multiplexers and/or else input–output pin drivers. Finally, the digital output is obtained from the sense amplifier.

### *3.2 Read Operation*

During the read task, the drivers in the pre-charge circuit pre-charge the bit line to zero and Vdd to one. Every cell in the selected row pulls anyone of its two-bit lines to GND. Sense amplifiers rapidly identify the slight voltage difference developing between bit lines and generate the suitable digital output. Recognition of variation in voltage is highly delicate to electrical noise, and the SRAM utilizes a couple of bit lines for each bit and a differential sense amplifier to provide more excellent noise immunity.

### *3.3 Write Operation*

Drivers set the bit lines to the desired value during the write mode, and $V_{DD}$ equals zero value. Data has been placed on the bit line as well as turn ON the access transistor. At that point address decoder initiates any of the word lines to high esteem, choosing every cell in a specific row for write activity. Every cell is overpowered by the drivers that store values. If the storage node is initially storing one, then the data on the bit line is one.

## 4 Results and Discussion

Simulated outcomes of the 1 T-1D structure are displayed in this subdivision. Tanner 13 EDA tool is utilized for designing the circuit. SRAM cell performance is affected through the parameters such as power and delay. The performance of some SRAM cell categories is analyzed in this work. The read and write processes are performed in memory for one-bit storage of CMOS technology.

The input and output waveform of 1 T-1D SRAM cell with two inputs and one output is in CMOS shown in Fig. 4. Input bits are given at the bit line, and a voltage source of 5 V is connected with the word line. Based on the two inputs, the SRAM cell performs its functioning and displays the output. The input bits at the bit line may be composed of 0 & 1, and the voltage level giving at the word line is 5 V. For bit 1, there is an occurrence of charging, and hence the output will be logic 1. If the

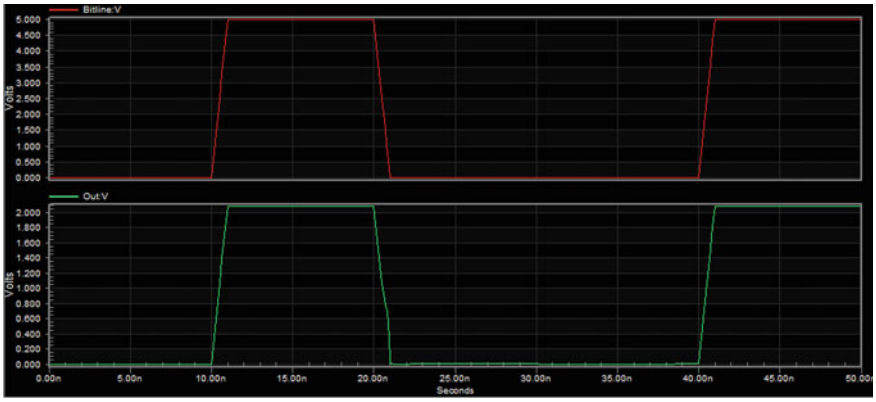



**Fig. 4** Representation of Input and Output waveform (CMOS Technology)

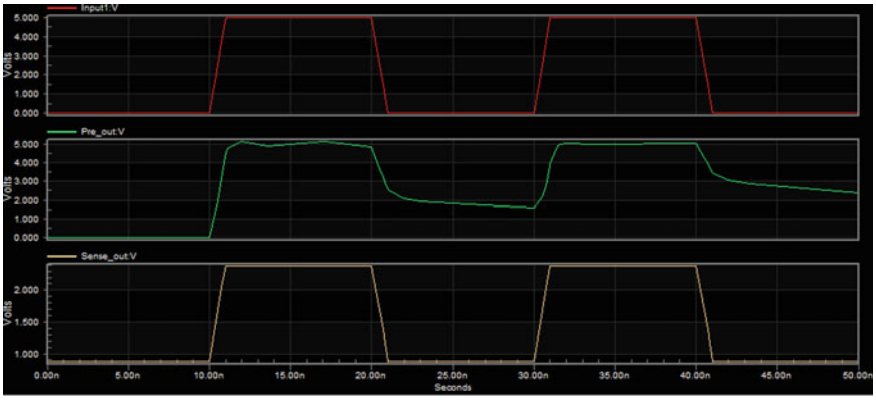



**Fig. 5** Read process (CMOS)

input bit is 0, there may be the possibility of discharging, and hence the result is logic 0.

The simulation results of read operation in CMOS technology are given in Fig. 5. Initially, the memory will have some value; the word line should also be high to perform read operation. The driver's pre-charge all bit lines to Vdd (1) and left them floating. The address decoder is activated by the word line, whereas the external word line driver is disabled. Word line is made active by selecting a row in addition turn ON access transistor. In read mode, charges stored in internal nodes will be interrupted by the bit line charges. The bit line is linked to SA acts as a comparator to generate the output.

The analysis of the write process in CMOS technology is shown in Fig. 6. During the write process, data must be placed on the bit line and turn ON the access transistor. If the value of data is zero, the output node finds a direct path to the ground. The

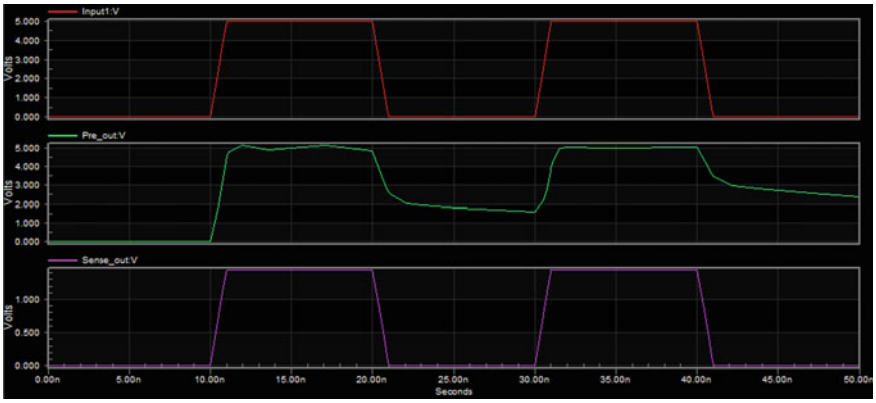



**Fig. 6** Write process (CMOS)

**Table 1** 1 T-1D Read, write power and delay (CMOS technology)

| $V_{dd}$ (V) | Read power (W) | Write power (W) | Read delay (sec) | Write delay (sec) |
|---|---|---|---|---|
| 0.8 | 8.50e-09 | 8.56e-09 | 2.055e-8 | 2.10e-8 |
| 0.9 | 9.45e-09 | 9.52e-09 | 2.056e-8 | 2.093e-8 |
| 1.0 | 1.12e-08 | 1.04e-08 | 2.0944e-8 | 2.097e-8 |
| 1.1 | 1.12e-08 | 1.13e-08 | 2.094e-8 | 2.098e-8 |

charge stored on the output node discharges through the access transistor. Initially, if the storage node stores zero, bit line data is one. Similarly, bit line data is zero, and the storage node stores also zero; there is no change in the cell state. The power, as well as delay, are the essential parameters in the 1 T-1D SRAM cell. Both read as well as write power and read, write delay concerning varying supply voltages in CMOS technology is shown in Table 1. The power range differs for each varying supply voltages. The read, write operations are fundamental operations in the SRAM cell.

Table 2 illustrates the power comparison of the proposed 1 T-1D CMOS SRAM cell with other structures (4 T–11 T) in 180 nm technology. Table 2 describes outcomes of 1 T-1D CMOS in 180 nm technology. The power of each SRAM cell is varied following the voltage variations. At 1 V, power for CMOS is 2.52E-8 W that describes low power consumption.

**Table 2** Power comparison of various existing SRAM cell with proposed work in 180 nm technology

| Supply voltage/ SRAM cell | 0.8 V | 0.9 V | 1.0 V | 1.1 V |
|---|---|---|---|---|
| 4 T [14] | 1.12e-07 | 1.42e-07 | 1.73e-07 | 2.03e-07 |
| 5 T [14] | 4.80e-07 | 6.15e-07 | 7.64e-07 | 1.13e-06 |
| 6 T [14] | 4.97e-07 | 6.37e-07 | 8.00e-07 | 1.18e-06 |
| 7 T [14] | 4.85e-07 | 6.17e-07 | 7.70e-07 | 1.14e-06 |
| 8 T [14] | 2.37e-07 | 3.06e-07 | 3.85e-07 | 5.70e-07 |
| 9 T [14] | 6.15e-07 | 8.12e-07 | 1.03e-06 | 1.54e-06 |
| 10 T [14] | 6.02e-08 | 8.31e-08 | 1.08e-07 | 1.73e-07 |
| 11 T [14] | 6.65e-08 | 8.95e-08 | 1.13e-07 | 1.62e-07 |
| **1 T-1D CMOS** (PROPOSED) | 2.016e-08 | 2.2713e-08 | 2.5269e-08 | 2.7833e-08 |

## 5 Conclusion

The 1 T-1D SRAM design permits the cost-effective embedding of enormous quantities of memory in SoC designs by employing the merits of one transistor bit cell. A diode structure can appear in a memory cell for various reasons, for instance, in certain types of solid-state memory devices that use a structural phase-change material as the programmable data storage mechanism. The read and write operations are permitted when the access transistors are turned on via word line and turned off at load condition. The functionality of the 1 T-1D cell structure is designed and justified by the simulation results. In the future, the low power design of 1 T-1D SRAM cell can be obtained in the latest Nano-meter technologies.

## References

1. Agrawal T, Kumar A, Saraswat SK (2016, November) Design of low-power SRAM on Artix-7 FPGA. In 2016 2nd international conference on communication control and intelligent systems (CCIS) IEEE, pp 203–209
2. Song T, Rim W, Park S, Kim Y, Yang G, Kim H, Baek S, Jung J, Kwon B, Cho S, Jung H (2016) A 10 nm FinFET 128 Mb SRAM with assist adjustment system for power, performance, and area optimization. IEEE J Solid-State Circuits 52(1):240–249
3. Moghaddam M, Timarchi S, Moaiyeri MH, Eshghi M (2016) An ultra-low-power 9T SRAM cell based on threshold voltage techniques. Circuits Syst Signal Process 35(5):1437–1455
4. Lundager K, Zeinali B, Tohidi M, Madsen J, Moradi F (2016) Low power design for future wearable and implantable devices. J Low Power Electron Appl 6(4):20
5. Ensan SS, Moaiyeri MH, Moghaddam M, Hessabi S (2019) A low-power single-ended SRAM in FinFET technology. AEU-Int J Electron Commun 99:361–368
6. Yang KH, Tsai HJ, Li CY, Jendra P, Chang MF, Chen TF (2016) eTag: Tag-comparison in memory to achieve direct data access based on eDRAM to improve the energy efficiency of the DRAM cache. IEEE Trans Circuits and Syst I: Regular Papers 64(4):858–868

7. Tripathi T, Chauhan DS, Singh SK, Singh SV (2017) Implementation of low-power 6T SRAM cell using MTCMOS technique. In: Advances in computer and computational sciences. Springer, Singapore, pp 475–482
8. Widjaja Y, Zeno Semiconductor Inc (2016) Semiconductor memory having volatile and multi-bit non-volatile functionality and method of operating. US Patent 9,257,179.
9. Singh S, Das A., Avago Technologies General IP (Singapore) Pte Ltd (2016) A memory cell having built-in read and write assist. US Patent 9,349,437
10. Zhang B, Arapostathis A, Nassif S, Orshansky M (2006, November). Analytical modeling of SRAM dynamic stability. In Proceedings of the 2006 IEEE/ACM international conference on Computer-aided design ACM, pp 315–322
11. Ahmad S, Gupta MK, Alam N, Hasan M (2016). Single-ended Schmitt-trigger-based robust low-power SRAM cell. IEEE Trans Very Large Scale Integration (VLSI) Syst 24(8):2634–2642
12. Moghaddam M, Timarchi S, Moaiyeri MH, Eshghi M (2016) An ultra-low-power 9T SRAM cell based on threshold voltage techniques. Circuits, Syst Signal Process 35(5):1437–1455
13. Pal S, Bose S, Ki WH, Islam A (2020) A highly stable reliable SRAM cell design for low power applications. Microelectron Reliab 105:113503
14. Chaudhuri D, Roy K, Nag A (2019) Comparison of different SRAM cell topologies using 180 nm technology. In: Advances in computer, communication and control. Springer, Singapore, pp 391–400
15. Wei M, Robin M, Portilla L, Ren Y, Shao S, Bai L, Cao Y, Pecunia V, Cui Z, Zhao J (2020) Air-stable N-type printed carbon nanotube thin film transistors for CMOS logic circuits. Carbon 163:145–153

# A Local Descriptor and Histogram of Oriented Gradients for Makeup Invariant Face Recognition Under Uncontrolled Environment


**Rajesh Kumar Tripathi**

**Abstract** Makeup is an important cosmetic which is used for beautification and for hiding their identity by the people or criminals. To recognize the face of a people or criminal after makeup is difficult due to the changes in the appearance of the face. This paper presents a descriptor which works on local region for extracting the important features for finding similarity against intra-class variation. The proposed pixel intensity difference is calculated from the local regions of the input image to generate the feature vector. Then, histogram of oriented gradient is also computed for shape feature and combined with the proposed descriptor feature. A machine learning classifier Multi-class SVM has been utilized and five-fold cross validation is used for calculating the rank-1 accuracy. The evaluation has been done carried on two standard datasets-VMU and YMU. The fusion approach performs better and outperforms to the most of the existing approaches of the standard datasets.

**Keywords** Descriptors · Face recognition · Makeup · Multi-class SVM

## 1 Introduction

Face is an important biometric trait for human identification or authentication under uncontrolled environment. Uncontrolled environment refers to the face image capturing under Pose, Illumination, Expression (PIE) and makeup variation. Most of the researchers have worked to handle face recognition under PIE variation. A few researchers have initiated to work for face recognition while makeup changes the appearance of the face. Makeup is a frequently used cosmetic which changes the appearances of the face photos. It changes the texture and shape of the face image. Most of the women uses makeup on their face daily basis and sometimes, criminals use the makeup to hide their identity. Therefore, authentication, identification, image tagging, criminal identification, human computer interaction [1] becomes more challenging using face images having makeup. There are two types of people who uses

R. K. Tripathi (✉)
Computer Engineering and Applications, GLA University, Mathura, India
e-mail: rajesh.tripathi@gla.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_2

light or heavy makeup. To handle the light or heavy makeup, a makeup invariant system is required to be develop.

Face appearance is altered due to the use of lipstick, powder, eye liner, mascara, foundation, contour, concealer, eye shadow etc. Face recognition under these appearance changing materials becomes very difficult. The results of the different approaches of face recognition under make-up [2–4] proves the degrading performance. Therefore, a robust approach is required for face recognition.

The local descriptors are used frequently for face recognition or texture classification. Many descriptors use the neighboring pixels of the reference pixel for encoding the relation. Sometimes, descriptors skips pixels of different radial widths which may cause the loss of important information. Therefore, a local region based descriptor has been presented for texture feature extraction and histogram of oriented gradients for shape feature.

This descriptor works on the local region to extract the intensity difference feature for face recognition. In this, all the pixels of local region are involved for feature extraction. The $k^{th}$ intensity difference is highly close to the $k^{th}$ intensity difference of intra-class similar images and discriminative features to make the distinction between inter-class similarity. Histogram of oriented gradients is more robust for extracting the shape feature. Fusion of shape feature with texture feature is more robust to handle makeup variation. Therefore, HOG is combined with the proposed descriptor.

The major contribution is robust to handle the appearance changes caused by makeup for face recognition, is discussed below:

- The presented descriptor finds the texture feature by calculating intensity difference of ordered pixel intensities of extracted local regions which creates robust feature vector for face recognition which handles appearance changes.
- The histogram of oriented gradients is helpful in shape feature extraction which combined with the proposed descriptor feature vector for face recognition under makeup.
- The fusion approach performed well and improved the result on standard datasets which includes synthetic and real makeup face images.

The remaining part of the research work is structured as follows: Sect. 2 explores the progress in handling of face recognition having makeup, Sect. 3 presents the proposed approach, Sect. 4 presents result and discussion, finally, Sect. 5 presents conclusion.

## 2 Related Work

The improvement in the area of face recognition having makeup variation has been presented in this section. Many researchers have presented their approaches to handle the texture and shape changes caused by makeup on the face image.

Guo et al. [6] used important features- smoothness, skin color tone, and texture changes of skin. Two dimensionality reduction methods PCA and LDA were used for feature extraction. Wang et al. [7] presented a face verification method for face recognition in which random space has been used to obtain different correlation space. Chen et al. [8] fused the local descriptors that are densely sampled LBP (DS-LBP), Histogram Ordinal Measures Ratio (HGORM), and Local Gradient Gabor Pattern (LGGP) and then used improved LDA for ensemble learning and generating scores of classifiers. Then, sum rule has been applied to fuse generated scores. This approach could not handle the face recognition for those face images which are having large pose and hair style. In next year, Poon et al. [9] generated the gradient faces and then used PCA for feature extraction and dimensionality reduction. Similarity was measured using Euclidean distance.

Zheng et al. [10] proposed learning approach based on a multi-level feature to handle appearance changes due to makeup. This approach used SVM for classification. It could not work well while face image having lighting effects and large pose variation. Sun et al. [11] constructed a triplet-network that has 3 face-inputs for the creation of negative-pair and positive-pair. The Alexnet model was used for testing and training.

In 2018, Sajid et al. [12] generated makeup faces from original input images and then a CNN model VGG-19 was used for face recognition. Recently, Wang et al. [13] proposed a network which is multi-branch based for face recognition under makeup.

## 3 Methods

The sequence of the presented fusion approach for makeup invariant face recognition. Figure 1 shows the block diagram which clearly shown the sequence of face recognition. The fusion approach has the following three steps-preprocessing, feature extraction using presented descriptor and HOG, and multiclass SVM to perform the classification and calculating the recognition accuracy. The face recognition process is explored in subsections which have been discussed below:

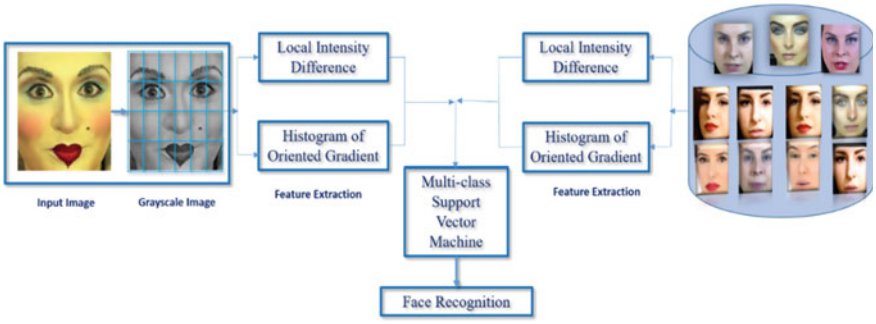



**Fig. 1** Shows the process of the proposed approach

### 3.1 Preprocessing of Face Image

The preprocessing step has been used for color to grayscale conversion and then dimension is reduced from $130 \times 150$ to $64 \times 64$. A descriptor which works on local regions are strongly able to handle face appearance changes due to makeup and PIE variation. The local regions of size $p \times p$ of input image are utilized for feature extraction.

### 3.2 Local Pixel Intensity Difference

The local descriptor computes intensity difference that has been applied over local region for feature extraction that is invariant to makeup with PIE variation. In this process, local regions of size $p \times p$ are taken out where pixel intensities are organized in increasing order. Considered a patch has 16 pixels- $\left[P^1, P^2, P^3, \ldots\ldots\ldots\ldots\ldots, P^{p\times p}\right]$. The pixel intensities are arranged in an ordered way $\left[I^1, I^2, I^3, \ldots\ldots\ldots\ldots\ldots, I^{p\times p}\right]$. An array containing the all the intensities in increasing order. In this process, alternate pixel intensity difference is calculated. To find the first order derivative (difference), first index intensity is subtracted from third, second indexed from fourth, fifth indexed from seventh, sixth indexed from eight and so on. This pattern provides difference values from 16 to 8 in one iteration. The iteration is done 5 times to calculate the difference value of a patch's pixel.

The equations given below helps to calculate local intensity difference value:

$$\delta^1 = \begin{bmatrix} I^1 \\ I^2 \\ I^3 \\ . \\ . \\ . \\ I^{p\times p} \end{bmatrix} \tag{1}$$

$$\delta^2_{1,2} = \Re(I_3) - \Re(I_1) \tag{2}$$

$$\delta^2_{2,2} = \Re(I_4) - \Re(I_2) \tag{3}$$

$$\delta^2_{3,2} = \Re(I_7) - \Re(I_5) \tag{4}$$

$$\delta^2_{4,2} = \Re(I_8) - \Re(I_6) \tag{5}$$

$$\delta^2_{5,2} = \Re(I_{11}) - \Re(I_9) \tag{6}$$

$$\delta^2_{6,2} = \Re(I_{12}) - \Re(I_{10}) \tag{7}$$

$$\delta^2_{7,2} = \Re(I_{15}) - \Re(I_{13}) \tag{8}$$

$$\delta^2_{8,2} = \Re(I_{16}) - \Re(I_{14}) \tag{9}$$

The first difference has been computed with the help of the above Eqs. 1–9. Further, four values are computed from eight difference values (using Eq. 10–13):

$$\delta^3_{1,3} = \Re(\delta^2_{2,2}) - \Re(\delta^2_{1,2}) \tag{10}$$

$$\delta^3_{2,3} = \Re(\delta^2_{4,2}) - \Re(\delta^2_{3,2}) \tag{11}$$

$$\delta^3_{3,3} = \Re(\delta^2_{6,2}) - \Re(\delta^2_{5,2}) \tag{12}$$

$$\delta^3_{4,3} = \Re(\delta^2_{8,2}) - \Re(\delta^2_{7,2}) \tag{13}$$

Next, difference values are computed using above four difference values:

$$\delta^4_{1,4} = \Re(\delta^3_{2,3}) - \Re(\delta^3_{1,3}) \tag{14}$$

$$\delta^4_{2,4} = \Re(\delta^3_{4,3}) - \Re(\delta^3_{3,3}) \tag{15}$$

Finally, Eq. 14–16 generates the final value to store as a feature vector:

$$\delta^5_{1,5} = \Re(\delta^4_{2,4}) - \Re(\delta^4_{1,3}) \tag{16}$$

The Eqs. 1–16 is applied over a $4 \times 4$ local region of the face. Equation 1 is representing to the sorted pixel intensities in $16 \times 1$ matrix, then Eqs. 2–9 are used to calculate the first order difference as $\delta^2$. Then, $\delta^2$ has dimension $8 \times 1$. Finally, $\delta^3$ generate $4 \times 1$, $\delta^4$ generates $2 \times 1$, and $\delta^5$ generates $1 \times 1$ matrix. The final value for $\delta^5$ is key value for feature vector.

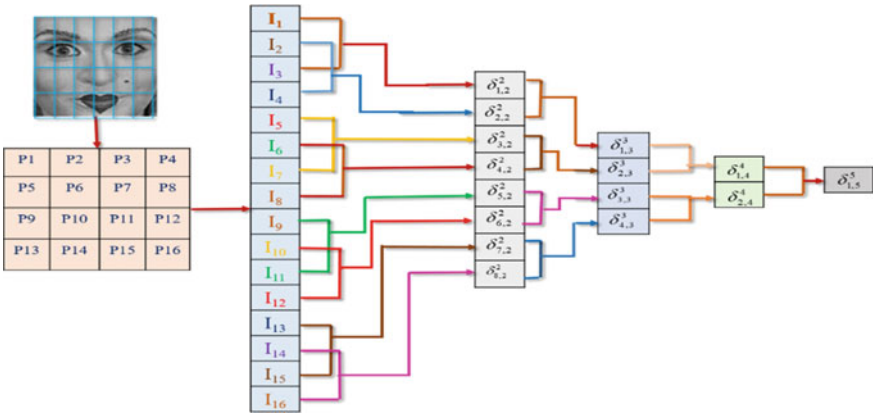



**Fig. 2** Proposed local intensity difference descriptor for feature extraction

The procedure related to feature extraction has been discussed in Fig. 2 to easily understand.

## *3.3 Histogram of Oriented Gradients*

The well-known descriptor HOG [14] has been used to compute the oriented gradients of the local regions. The HOG is highly helpful to compute the shape features of the face image which is combined with the above descriptor to generate a high discriminative feature vector. The $x$ and $y$ directional gradients are computed as $DG_x$ and $DG_y$ respectively over a $4 \times 4$ patch and then magnitude is computed with the help of the given Eq. 17:

$$G = \sqrt{(DG_x)^2 + (DG_y)^2} \tag{17}$$

Finally, computed local pixel intensity difference values and histogram of oriented gradients are combined into a feature vector.

## *3.4 Multi-class Support Vector Machine (MSVM)*

In order to calculate the accuracy of the proposed approach, Multi-class SVM. SVM works better in recognition for the small dataset and gives better result. The fivefold cross validation method is applied for training and testing purpose. This method is best to test all the fold's images. In this approach, MSVM divides the data into five different folds and complete the training and testing in five different iterations. The average result of fivefold iteration is computed as recognition accuracy.

**Table 1** Results of the state-of-the-art descriptors and proposed fusion approach using SVM, Euclidean and Chi-square measures on YMU dataset

| Works | Method | Chi-square | Euclidean distance | Multi-SVM |
|---|---|---|---|---|
| Dubey et al. [15] (2019) | Frequency decoded LBP | 65.37 | 64.52 | 68.52 |
| Chakraborty et al. [16] (2020) | R-Theta | 74.51 | 73.47 | 76.92 |
| **Proposed method** | **Local pixel intensity difference + HOG** | **82.39** | **81.27** | **86.25** |

## 4 Results and Discussion

The performance of the fusion approach is evaluated on two standard datasets Virtual Makeup (VMU) [2] and YouTube Makeup (YMU) [5]. The reason of choosing these datasets is the availability of the dataset publicly for the research purpose.

To prove the robustness of the fusion method, evaluation has been performed with existing methods such as HOG [3], LBP [3], and LGBP [3], fusion of DS-LBP, HGORM, LGGP [8], Sajid et al. [12], R-Theta [15], Poon et al. [9], and FDLBP [16]. The purpose of the fusion of HOG with proposed descriptor is to extract shape feature which helps in improvement. The proposed method outperforms to the existing methods which have used descriptors. The deep learning approaches performance are based on size of training set. Deep learning approaches performs well on large datasets. The proposed fusion method presented better result than FGGNet with six makeup face generation [12], and very close to FGGNet using Transfer learning and without augmentation [12]. Therefore, the proposed fusion of descriptors works better and improves the result. Tables 2 and 3 shows the performance analysis of the proposed fusion method and state-of-the-art methods.

### *4.1 Quantitative Analysis with Classifier and Similarity Measure*

Descriptors are more robust to makeup variation, PIE variation, therefore, most of the works have used the descriptors or fusion of descriptors. Evaluation has been done using Multi-SVM and some similarity measures and found that MSVM works better in comparison to Chi-square, Euclidean measures (Table 1).

**Table 2** Result of existing methods and the presented fusion approach on YMU dataset

| Works | Year | Method | Recognition accuracy (%) |
|---|---|---|---|
| [3] | 2013 | HOG | 71.90 |
| [3] | 2013 | LBP | 74.40 |
| [3] | 2013 | LGBP | 80.50 |
| [8] | 2016 | Fusion of HGORM, DS-LBP, and LGGP | 80.46 |
| [9] | 2017 | Gradient faces and PCA | 84.50 |
| [12] | 2018 | FGGNet using 6 face makeup styles | 75.61 |
| [12] | 2018 | FGGNet using Transfer learning and without augmentation | 86.37 |
| [15] | 2019 | Frequency decoded LBP | 68.52 |
| [16] | 2019 | R-Theta local neighborhood pattern | 76.92 |
| **Proposed method** | | **Local pixel intensity difference + HOG** | **86.25** |

**Table 3** Results of existing works and proposed fusion approach on VMU dataset

| Works | Year | Method | Recognition accuracy |
|---|---|---|---|
| [3] | 2013 | HOG | 73.60 |
| [3] | 2013 | LBP | 78.30 |
| [3] | 2013 | LGBP | 82.20 |
| [12] | 2018 | FGGNet using 6 face makeup styles | 78.19 |
| [12] | 2018 | FGGNet using Transfer learning and without augmentation | 88.48 |
| [16] | 2019 | Frequency decoded LBP | 68.52 |
| [15] | 2019 | R-Theta local neighborhood pattern | 76.92 |
| **Proposed method** | | **Local pixel intensity difference + HOG** | **88.75** |

## *4.2 Qualitative Analysis with Different Datasets*

**Experiments on YMU Dataset [5].** YMU has 604 face images of 151 white females where 2 face images for non-makeup and 2 for makeup images. The face images are having lipstick/gloss, foundation, eye liner etc. The face images have also PIE variation. This is the dataset with real makeup face images.

The proposed fusion of descriptors is able to handle the makeup variation with PIE variation. This local descriptor outperforms to the descriptors used in Guo et al. [3], fusion of descriptors in Chen et al. [8], R-Theta [15], Poon et al. [9], and FDLBP [16].

**Experiments on VMU Dataset [2].** VMU [2] has 204 faces of 51 subjects where every female has eye makeup, no makeup, full makeup, lipstick, and full makeup with blush, eye face images and foundation. In this, makeup face images are generated after applying synthetic makeup.

The proposed fusion approach works well with synthetic makeup face images having PIE variation. Histogram of oriented gradients are helpful to extract gradient feature which is more discriminative.

### *4.3 Comparative Analysis with State-of-the-Art*

**Experiments on YMU Dataset [5].** The proposed descriptor with oriented gradients is performed well and handle makeup variation. The fusion of descriptor outperforms to the used descriptors in Guo et al. [3], Poon et al. [9], fusion of descriptors in Chen et al. [8], and recent descriptors R-Theta [15], FDLBP [16]. Only, fusion of proposed descriptor with HOG is comparable to sajid et al. [12].

Table 2 proves the robustness and effectiveness of the fusion of descriptors in comparison to the existing works on YMU dataset.

**Experiments on VMU Dataset [2].** The evaluation of the proposed fusion of descriptors has been done with few state-of-the-art approaches who have used VMU dataset because of synthetic makeup. The fusion of descriptors outperforms to the existing works such as HOG [3], LBP [3], LGBP [3], Sajid et al. [12], R-Theta [15] and FDLBP [16].

Table 3 shows the performance which shows its robustness and effectiveness on VMU images in comparison to the existing works.

## 5 Conclusion

The proposed fusion approach improved the accuracy and proved its robustness for face recognition under face appearance changes due to makeup. The intensity difference descriptor extracts the texture feature and robust against makeup. The HOG extracts the shape feature of the face image and complement to texture feature for face recognition under makeup. The fusion of descriptors proved better in handling makeup variation and outperforms to the existing descriptors. The performance of the proposed method is comparable to existing deep learning approach.

In the future, improvement can be made with the help of makeup removal approach and deep learning approaches.

## References


1. Jain A, Ross A, Kumar KN (2011) Introduction to biometrics. Springer, Heidelberg
2. Dantcheva A, Chen C, Ross A (2012) Can facial cosmetics affect the matching accuracy of face recognition systems? In: Proceedings of conference on biometrics: theory, applications, and systems, BTAS, pp 391–398
3. Guo G, Wen L, Yan S (2014) Face authentication with makeup changes. IEEE Trans Circ Syst Video Technol 24(99):814–825
4. Marie-Lena E, Neslihan K, Jean-Luc D (2013) Facial cosmetics database and impact analysis on automatic face recognition. In: Proceedings of multimedia signal processing, MMSP
5. Chen C, Dantcheva A, Ross A (2013) Automatic Facial makeup detection with application in face recognition. In: Proceedings of 6th IAPR international conference on biometrics (ICB), Madrid, Spain
6. Guo G, Wen L, Yan S (2014) Face authentication with makeup changes. IEEE Trans Circ Syst Video Technol 24(5):814–825
7. Wang X, Kambhamettu C (2014) A new approach for face recognition under makeup changes. In: IEEE global conference on signal and information processing (GlobalSIP), pp 423–427
8. Chen C, Antitza D, Ross A (2016) An ensemble of patch-based subspaces for makeup robust face recognition. Inf Fusion 32:80–92
9. Poon B, Amin MA, Yan H (2017) PCA based human face recognition with improved method for distorted images due to facial makeup. In: Proceedings of the international multi conference of engineers and computer scientists, Hong Kong, vol 1
10. Zheng Z, Kambhamettu C (2017) Multi-level feature learning for face recognition under makeup changes. In: 12th international conference on automatic face and gesture recognition. IEEE, pp 918–923
11. Sun Y, Ren L, Wei Z, Liu B, Zhai Y, Liu S (2017) A weakly supervised method for makeup-invariant face verification. Pattern Recogn 66:153–159
12. Sajid M, Ali N, Dar SH, Ratyal NI, Butt AR, Zafar B, Shafique T, Baig MJA, Riaz I, Baig S (2018) Data augmentation-assisted makeup-invariant face recognition. Math Prob Eng 2850632:1–11
13. Wang W, Fu Y, Qian X, Jiang YG, Tian Q, Xue X (2020) $FM^{2u}$-Net: face morphological multi-branch network for makeup-invariant face verification. In: Computer vision and pattern recognition
14. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE computer society conference on computer vision and pattern recognition (CVPR 2005), vol 1, pp 886–893
15. Chakraborty S, Singh SK, Chakraborty P (2019) R-theta local neighborhood pattern for unconstrained facial image recognition and retrieval. Multimedia Tools Appl 78(11):14799–14822
16. Dubey S (2019) Face retrieval using frequency decoded local descriptor. Multimedia Tools Appl 78:16411–16431

# A Preventive Framework for Mine Representatives Utilizing Remote Sensor Networks with Optimized Routing


**D. Jayakumar, T. Rajesh Kumar, and C. M. Velu**

**Abstract** Security is the most basic piece of an industry. Inconsiderateness in the security part may cause hurting of first rate gear hampering of creation or may cause loss of human life also in absurd cases. Industrial safety is one of the fundamental parts of the Industry, particularly mining. Wellbeing is a vital factor in mining. To forestall material misfortune and harm to human wellbeing, a security framework and a dependable correspondence framework are needed in underground mines. To guarantee both security and profitability in mines, solid correspondence should be set up between laborers traveling through the mine and a fixed base station. Inside the mines, the link correspondence framework isn't as successful. In this research study, we will screen mine boundaries such as unusual gas, temperature and stickiness to keep destructive gas or high temperature from assaulting diggers and furthermore we examined successful directing calculation for power saving in ZigBee.

**Keywords** Gas sensor · Humidity sensor · Mine · Reliable · Temperature sensor · Wireless · ZigBee

## 1 Introduction

In the mining industry prosperity and security is a vital piece of all. To avoid such unwanted wonders all mining industry follows some central protection and miracles. Correspondence is the most major key factor today, to screen different limits relentlessly and to take fundamental actions in like way to avoid any sorts of dangers related to creation, security, managing of HR [1]. To avoid loss of material and hurting of human prosperity, security and security structure similarly as strong determined solid correspondence system is central in within the underground mines. To improve

D. Jayakumar (✉) · T. Rajesh Kumar
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India
e-mail: Djayakumar@kluniversity.in

C. M. Velu
Department of CSE, Saveetha School of Engineering, Saveetha University, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_3

security, prosperity and productivity in underground mines, a strong correspondence structure ought to be set up between workers, moving in the mine, and a fixed base station. The correspondence organization ought not to be upset at any second and at any condition [2, 3]. Inside underground mines, the wired correspondence network system isn't so amazing. The unfaltering quality and long presence of standard correspondences structures in severe mining conditions has reliably been an issue. Inside mines in view of abnormal condition the foundation cost similarly as help cost is high for wired correspondence associations. It is difficult to reinstall the wired correspondence system inside mines after a torrential slide or damage due to any clarification. If in view of some clarification any wire of the correspondence network hurts, it may cause brief obstruction of the consistent cycle or may cause a drawn out independent of the structure [4–6].

Because of rooftop slide, if using any and all means a few specialists caught inside mines, it is particularly needed to keep up the progression of the correspondence framework. It is particularly critical to know the real position and state of the caught laborers. To screen different boundaries during this condition it is a lot of important to keep up the correspondence framework not surprisingly [7]. Appropriately, advancement of mine checking framework to precisely identify temperature, to follow underground excavators and vehicles on continuous has critical importance to wellbeing creation and salvage of underground mine catastrophe [8]. Coal mineshaft wellbeing checking framework dependent on remote sensor organization can opportune and precisely reflect dynamic circumstance of staff in the underground locales to ground PC framework. In Existing System, in conventional technique the strange in any of the boundaries are sent in remote correspondence to hint the status to checking segment and in this strategy no alarm as voice is given to the laborers inside the mine. Consequently, we go for proposed framework [9, 10]. For the effectively remote information transmission, in this work the ZigBee determination is used. In this framework, the essential boundaries like temperature, moistness and dangerous methane gas will be observed and if any irregularity occur in any of the boundaries implies it will be hinted as voice inside the mine and sent to the checking area through ZigBee. For that we are having framework with Microcontroller, in that the sensors are interfaced with it. Here a voice IC named APR9600 is utilized for implying the strange status in voice design.

## 2 Related Works

Maimour [11] analyzed the impact of interference aware metrics in iterative multipath routing with AODV protocol. Subha and Malarkkan [12] has implemented a hybrid clustering energy aware routing protocol in WSN. Chehri et al. [13] has introduced a new energy efficient routing protocol in WSN to transmit the data with limited number of sensor nodes. The traffic was relayed by utilizing intermediate nodes to transmit the data to the final destination by generating a multi-hop network. Soomro and Jilani [15] predicts methane using artificial neural network to ensure the safety of miners.

Sunitha and Chandrikab [14] developed network condition aware node profiling and malicious node detection algorithm for determining the malicious sensor nodes. In addition to this, evolutionary computing assisted dual disjoint forwarding path algorithm for obtaining disjoint path with no shared elements for reducing energy consumption, and for ensuring QoS centric. Ramya et al. [16] proposed the modified AODV and DSR protocols for performing the routing. In this work, the throughput of the design is less due to the black hole attack.

## 3 Design of Embedded System Blocks

The block diagram of the design is shown in the Fig. 1(Mine section) & has a Microcontroller, LCD, Temperature sensor, Gas sensor, Humidity sensor, Voice play back circuit, ZigBee, and Power interact with each other as follows: When switched on, the power supply supplies 5 V to microcontroller for its operation. When the Digital signal board is on, a signal is transmitted by ZigBee.

The block diagram of the design is shown in the Fig. 2(Monitoring section) & has a ZigBee, Computer and Power supply blocks which interact with each other. Then the ZigBee at monitoring section will receive the signal and send to computer through serial communication circuit.

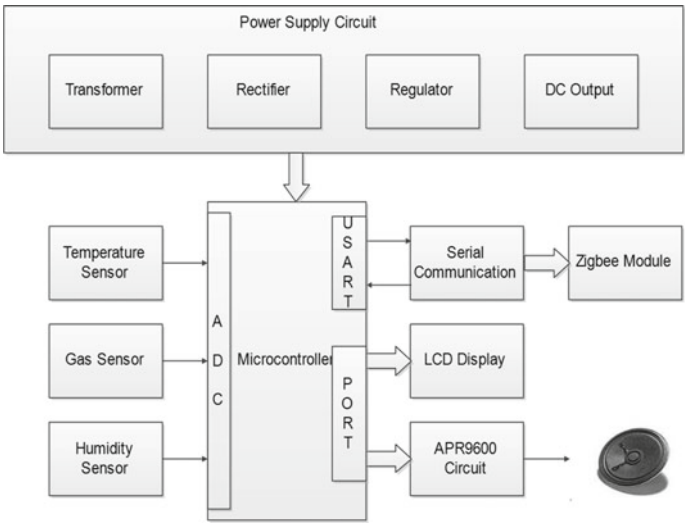



**Fig. 1** Mine section


**Fig. 2** Monitoring section

## 4 Mine Section

Digital sign board has a Microcontroller, LCD, Temperature sensor, Gas sensor, Humidity sensor, Voice play back circuit, ZigBee, and Power supply.

### *4.1 Microcontroller*

Microcontroller is the center of the arranged unit, which handles all of the signs. Any leftover interfacing blocks are interfaced to it. By enduring high heartbeat from the Digital sign board circuit it sends request to the exchange related with motor which cuts the affiliation or related with voice playback circuit. The ATmega8535 is a low-power CMOS 8-cycle microcontroller subject to the AVR improved RISC designing. By executing bearings in a lone clock cycle, the ATmega8535 achieves throughputs advancing toward 1 MIPS for every MHz allowing the system planner to improve power usage instead of taking care of speed. Predominant, Low-power AVR® 8-digit Microcontroller, Nonvolatile Program and Data Memories 8 K Bytes of In-System Self-Programmable Flash, Endurance: 10,000 Write/Erase Cycles, 512 Bytes EEPROM.

### *4.2 Power Supply*

The force supply incorporates Step Down transformer, and voltage controller. The transformer has two circle windings, the basic and the optional, around a common appealing core interest. The current spilling in the principal winding makes a period moving electro engaging field, which in like manner influences a yield voltage across the optional winding. The degree of the turns in the two windings picks the degree of

the data voltage and yield voltage. The higher voltage side has an all the flimsier (high check) wire with more turns while the lower voltage side has thicker (low measure) wire and less turns. The LM78XX series of three terminal positive regulators are available in the TO-220 package and with several fixed output voltages.

## *4.3 LCD*

Here, the LCD is related with the microcontroller. It is used to show messages. A liquid diamond show is an insignificant exertion, low power device prepared for showing text and pictures. LCD's are exceptionally essential in embedded systems. LCD can be found in different contraptions like watches, fax and copiers and little PCs. The LCD (HD44780U) used here. A variable or fixed resistor ought to be used on any LCD module as it appears in the above schematic. The HD44780U dab network fluid precious stone showcase regulator and driver LSI shows alphanumeric, Japanese kana characters, and images. It very well may be designed to drive a dab grid fluid gem show heavily influenced by a 4-or 8-digit microchip. Since every one of the capacities, for example, show RAM, character generator, and fluid gem driver, needed for driving a speck framework fluid gem show are inside given on one chip, an insignificant framework can be interfaced with this regulator/driver.

A solitary HD44780U can show dependent upon one 8-character line or two 8-character lines. The low force supply (2.7 V to 5.5 V) of the HD44780U is appropriate for any convenient battery-driven item requiring low force dispersal. The inward activity in the regulator chip is controlled by the signs sent from the MPU.

## *4.4 Temperature Sensor*

The LM35 game plan are precision facilitated circuit temperature sensors, whose yield voltage is straightly comparative with the Celsius (Centigrade) temperature. The LM35 in this manner appreciates a high ground over direct temperature sensors changed in Kelvin, as the customer isn't expected to remove a gigantic consistent voltage from its respect get profitable Centigrade scaling. The LM35 needn't bother with any external Calibration or figuring out how to give conventional correctness's of $\pm 1/4$ °C at room temperature and $\pm 3/4$ °C over a full $-55$ to $+150$ °C temperature range. Negligible exertion is ensured by overseeing and change at the water level. The LM35's low yield impedance, direct yield, and careful natural change make interfacing to readout or control equipment especially straightforward. It might be used with single power supplies, or with notwithstanding and short supplies. As it draws only 60 μA from its reserve, it has very low self-warming, under 0.1 °C in still air.

## 4.5 Gas Sensor

Gas sensor is utilized in gas spillage recognizing supplies in family and industry, are reasonable for identifying of LPG, iso-butane, propane, LNG, stay away from the commotion of liquor and cooking vapor and tobacco smoke. It has some features like, High sensitivity to LPG, iso-butane, propane, small sensitivity to LPG smoke, Fast response, Stable and long life, Standard detecting conditions.

## 4.6 Humidity Sensor

The Smartest dampness sensor is a two terminal capacitor, which extensions in regard as water particles are held into its dynamic polymer dielectric. It has a couple of features: unrivaled, long stretch security, close protections, strong turn of events and simplicity. Applications: Air conditioners, Climate control for green houses, Storage and stockrooms, meteorological applications, Food taking care of, Room comfort control, clinical applications.

## 4.7 Play Back Circuit

The APR9600 contraption offers authentic single-chip voice recording, non-flighty limit, and playback capacity for 40 to 60 s. The contraption maintains both self-assertive and progressive access of various messages. Test rates are customer selectable, allowing originators to adjust their arrangement for unique quality and limit time needs. Consolidated yield enhancer, beneficiary intensifier, and AGC circuits essentially develop structure plan. The contraption is ideal for use in adaptable voice recorders, toys, and various other client and present day applications. APLUS facilitated achieves these irrefutable levels of limit capacity by using its selective basic/staggered limit advancement executed in a general Flash non-flighty memory measure, where each memory cell can store 256 voltage levels. This advancement enables the APR9600 device to mirror voice signals in their ordinary design. It discards the prerequisite for encoding and pressing factor, which routinely present bowing.

# 5 Monitoring Section

## *5.1 ZigBee*

FSK Transceiver module, which is organized utilizing the ChipconIC (CC2500). It is an ensured single-chip handset; it depends upon 3 wire modernized progressive interface for cautious neighborhood oscillator age. So the rehash could be setting. It is an unparalleled and immaterial effort module. It gives 30 m range with locally open radio wire. In an average framework, this trans-finder will be utilized long side a microcontroller. It gives far reaching equipment backing to designate consideration of, information buffering, burst transmissions, clear channel assessment, interface quality sign and wake on radio. It very well may be utilized in 2400–2483.5 MHz ISM/SRD band structures. (for example RKE two-way Remote Keyless Entry, distant alert and security frameworks, AMR-adjusted Meter Reading, Consumer Electronics, Industrial seeing and control, Wireless Game Controllers, Wireless Audio/Keyboard/Mouse). It could no ifs, ands or buts to plan thing requiring inaccessible association. Working Range is 30 m without requiring any outer radio wire.

## *5.2 Computer*

PC is utilized to store information from the recipient. Programming, created here is to make an intuitive, dependable checking and the executives of detected information. The distinctive natural boundaries got by the ground control PC are shown in those habits in the LCD screen. The boundaries incorporate the temperature, mugginess, and gas. The PC stores the boundaries in the hard plate and ground staff can pick any of the boundaries for recording and replaying.

# 6 Safety System for Mine Workers

The made system can be disengaged into two regions. First is a hardware circuit that will be associated with the body of the earthmovers. This may be in a perfect world fitted with the prosperity cap of the experts also which should be obligatory in the premises of any underground mines. An additional save system can be fitted with the wrist of the underground backhoes at whatever point required. The circuit has a sensor module involving certain MEMS based sensors that activities continuous underground limits like temperature, tenacity combination of different gases inside mines, etc. Gas obsession is expected for the harmful gases. A segment of the gases is destructive and some are inflammable. A microcontroller is used with the sensors to get the sensor yields and to take the essential decision. The microcontroller can

store data's depending on the situation by the customer for keeping up of records. At the point when temperature is more than the security level set up at microcontroller, microcontroller deciphers blast alerts through the speaker and LCD related with. Again, when the intentional tenacity regard is more than the prosperity level coordinated at microcontroller, it deciphers different sort message. In like manner when gas center crosses the security level, microcontroller deciphers message. In each and every such case, this will send an alarm through a squeezing message and alert sound to the ground control terminal through ZigBee.

Second section is hardware and programming that will be placed in ground terminal. Considering the alert got ground staff takes decision and set up voice correspondence with the underground workers. The security division people are sent caution. Specific control and prosperity measures are taken in like way subject to the relentless seeing of condition and voice correspondence with underground people.

## 7 Routing Algorithm Used in WSN

### *7.1 Zone Based Routing*

The maximum min zPmin calculation needs right force level data for every one of the hubs in the organization. For huge scope sensor networks this is definitely not a potential presumption. Rather we propose to bunch along groups of sensors and gauge the in general steering force of the group for the point of the maximum min zPmin calculation. All the more explicitly we propose to arrange the organization primarily in geological zones, and progressively to control steering across the zones. The thought is to group along every one of the hubs that square measure in geographic vicinity as a zone, treat the zone as a substance in the organization, and permit each zone to settle on a choice an approach to highway a message across.

The hosts in a zone independently to process the ideal life, the message rates are known. The maximum min calculation won't have this information. This topographical dividing can be implemented essentially utilizing GPS information from each host. Direct neighborhood steering and take part in assessing the zone power level. Each message is steered across the zones abuse data concerning the zone power gauges. In our vision, a worldwide regulator for message steering deals with the zones. This might be the hub with the best force, albeit different plans, for example, circular robin may moreover utilize. In the event that the organization can be isolated into a relatively minuscule assortment of zones, the scale for the overall steering calculation is decreased. The worldwide information expected to send each message across is summed up by the office level gauge of each zone. We accept that in gadget networks this cost won't need continuous updates because of perceptible changes will happen exclusively once extensive stretches of time. The remainder of this part examines (1) how the hosts in a very zone team up to assess the force of the zone;

(2) how a message is steered inside a zone; and (3) anyway a message is directed across zones.

## *7.2 Zone Power Estimation*

The force gauge for each zone is constrained by a hub in the zone. This assessment estimates the quantity of messages which will move through the zone. Since the messages come from one adjoining zone and get coordinated to a unique adjoining zone, we propose a procedure during which the force assessment is done comparative with the heading of message transmission. The convention utilized by the regulator hub comprises of surveying every hub for its force level followed by running the maximum min zPmin calculation. The returned cost is then communicated to every one of the zones inside the framework. The recurrence of this strategy is contrarily relative to the measurable force level. At the point when the force level is high, the force assessment update will be done rarely because of a little Number of messages steered through the zone jars not adjustment the general force a lot. At the point when the force level is low, message transmission through the zone is probably going to change the force dispersion fundamentally. Without loss of consensus, we accept that zones territory unit sq. with the goal that they need four neighbors highlighted the North, South, East, and West3. We expect extra that it is potential to convey between the hubs that are closed to the boundary between 2 zones, so that in outcome the line hubs territory unit part of each zones. At the end of the day, adjoining zones that can speak with each option have a neighborhood of cover. The force gauge of a zone will be approximated as follows. We will utilize the maximum min zPmin algorithmic principle to gauge the office level, discover the maximum min zPmin way, reenact sending $\Delta$ messages through the way, and rehash until the organization is immersed. $\Delta$ is picked to be proportionate to the force level of the zone. All the more precisely, consider Figure three remaining. To gauge the force of zone B regarding sending messages inside the course from A to C, let the left piece of the cover among A and C be the source region and the correct piece of the cover among B and C the sink territory. The force of zone B toward the path from A to C is the maximal number of messages that can move from the source hubs to the sink hubs before a hub in B gets soaked. This can be registered with the maximum min zPmin calculation. We start with the force diagram of zone B and increase it. We make a nonexistent source hub S and associate it to all the source hubs. We make a fanciful sink hub T and interface all the sink hubs to it. Leave the loads of the recently added edges alone 0. The maximum min zPmin calculation run on this chart decides the force gauge for zone B toward A to C. Compared to AODV [11] protocol, DSR and modified AODV [16] protocol, Zone based routing achieved better performance in terms of throughput, which is shown in Table 1, and Fig. 3.

**Table 1** Comparative analysis

| Number of sensors | AODV [11] | DSR [16] | Modified AODV [16] | Zone based routing |
|---|---|---|---|---|
| 20 | 38 | 32 | 40 | **60** |
| 40 | 56 | 50 | 58 | **70** |
| 60 | 47 | 49 | 50 | **71** |
| 80 | 39 | 32 | 42 | **78** |
| 100 | 40 | 35 | 45 | **67** |

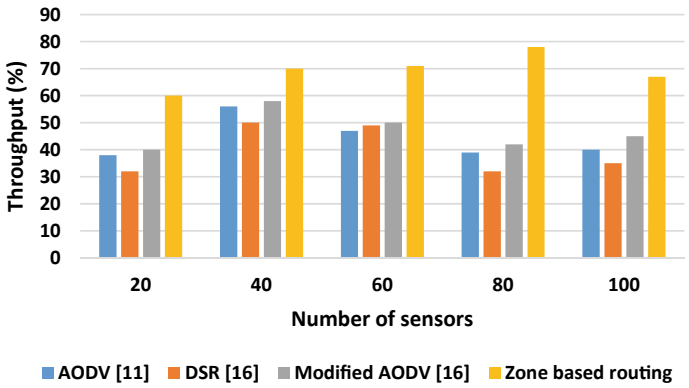



**Fig. 3** Graphical analysis

## 8 Conclusion

Traditional mine security system can be suitably replaced by the perception and prosperity structure proposed in the paper. This paper gives a system related to prosperity and security of underground mines. The system is strong, steadfast, constant, reasonable and straightforward. A greater district and greater significance inside unsafe underground mines are right now can be covered and potential accidents can be controlled effectively. The structure combined the low power, ease ZigBee based high repeat distant data transmission development. The sensor and ZigBee module can be obviously presented over the head defender of digger. Authentic checking and conversation is possible between the workers and the ground staff which can help with taking fitting actions even more rapidly and intelligently. The structure in like manner can be conveniently loosened up with ZigBee distant picture transmission office in future; it will improve flexibility of underground environment and expand accurate circumstance of tractors.

## References


1. Mineral Industry Safety and Health Centre (2012) Compliancegate database. http://www.mirmgate.com.au/index.php?gate=compliancegate. Accessed 8 Jan 2012
2. NSW Department of Primary Industries (2007) Preparing a health and safety management system. http://www.dpi.nsw.gov.au
3. National Research Centre for OHS Regulation (2012) About occupational health and safety regulation in Australia. http://ohs.anu.edu.au/ohs/index.php
4. Robens L (1972) Report of the Committee on Safety and Health at Work. Majesty's Stationery Office, London
5. Step Change in Safety (2011) Leading Performance Indicators. http://www.stepchangeinsafety.net, Accessed 8 Jan 2012
6. Jayakumar D, Omana J, Sivakumar M, Senthil B (2015) A safe guard system for mine workers using wireless sensor networks. Int J Appl Eng Res 10(8):21429–21441
7. Moore D, Leonard J, Rus D, Teller S (2004) Robusted distributed network localization with noisy range measurements. In: Proceedings of ACM SenSys 2004
8. Shannon C, Moore D (2004) The spread of the Witty worm. IEEE Secur. Priv. 2(4):46–50
9. Vural S, Ekici E (2007) Probability distribution of multi-hop- distance in one-dimensional sensor networks. Comput Netw Int J Comput Telecomm Netw 51:3727–3749
10. Trancă DC, Rosner D, Curatu R, Surpăteanu A, Mocanu M, Pardău Ş, Pălăcean AV (2017) Industrial WSN node extension and measurement systems for air, water and environmental monitoring: IoT enabled environment monitoring using NI WSN nodes. In: 2017 16th RoEduNet conference: networking in education and research (RoEduNet). IEEE, pp 1–6
11. Maimour M (2020) Impact of interference aware metrics on iterative multipath routing for industrial WSN. Internet Technol Lett 3(4):e159
12. Subha CP, Malarkkan S (2017) H-CERP: energy efficient scheme for industrial WSN applications. Wirel Pers Commun 94(4):1937–1950
13. Chehri A, Saadane R, Hakem N, Chaibi H (2020) Enhancing energy efficiency of wireless sensor network for mining industry applications. Procedia Comput Sci 176:261–270
14. Sunitha R, Chandrikab J (2020) Evolutionary computing assisted wireless sensor network mining for QoS-centric and energy-efficient routing protocol. Int J Eng 33(5):791–797
15. Soomro AH, Jilani MT (2020) Application of IoT and artificial neural networks (ANN) for monitoring of underground coal mines. In: 2020 international conference on information science and communication technology (ICISCT). IEEE, pp 1–8
16. Ramya TA, Mathana JM, Nirmala R, Gomathi R (2021) Exploration on enhanced quality of services for MANET through modified lumer and Fai-eta algorithm with modified AODV and DSR protocol. Mater Today: Proc

# A Quantitative Study of Image Fusion Using Hybrid Approach


**Budhi Veera Bharath Chandra, Mahapatra Medha Sampath Kumar, Chigurupati Naveen, Madhavarapu Srinivasa Sai Bhargav, R. Jagan, and Poornima Mohan**

**Abstract** Image fusion is a process that combines two complementary source images and gives a single image as output. It consolidates the beneficial information from different images into a single image. The image fusion methods that are studied and analyzed in this paper are DWT, SWT, DCT, and Hadamard Transform (HT). A new image fusion method called Stationary Wavelet Transform along with Hadamard Transform is proposed. Medical images such as Positron Emission Tomography (PET) images, Magnetic Resonance Image (MRI) images, and Computed Tomography (CT) images and multi-modal images, were used and analyzed using all the above-mentioned image fusion techniques including the proposed method. Obtained performance metrics such as PSNR, RMSE, Correlation, MAE and SSIM for all the techniques. The quantitative analysis has shown that the proposed method gave better results (Highest PSNR, least RMSE, Highest Correlation, least MAE and highest SSIM) and also the output images have good visual perception than input images.

**Keywords** Hadamard transform · Image fusion · SWT

## 1 Introduction

Image fusion is a prominent field in analyzing images. This helps to increase the image quality and reduce randomness and redundancy. This process helps to create more comprehensible images. Nowadays the advancements in image fusion are used in medical image processing like MRI, CT, PET scans. This image fusion can be carried out in two ways one is in the transform domain and the other is in the spatial domain. In the spatial domain, the fusion process deals with the image pixels whereas, in the transform domain, the fusion process uses the transformed coefficients such as DCT, DWT, SWT, and Hadamard transform.

B. V. B. Chandra · M. M. S. Kumar · C. Naveen · M. S. S. Bhargav · R. Jagan · P. Mohan (✉)
Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: poornimamohan12@gmail.com

B. V. B. Chandra
e-mail: bvbharathchandra8985@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_4

Average-DWT-Maximum and PCA-DWT-DSWT [6] are two hybrid methods for image fusion which gives high PSNR compared to the traditional methods. In our work, we are using a hybrid method SWT-Hadamard along with PCA for image fusion which gives better results compared to many widely used fusion methods like DWT, SWT, DCT, and Hadamard transform, which are studied for many images from different domains in our work. We claim that our method is better compared to many widely used methods because our method gives better values for performance metrics like PSNR, RMSE, Correlation, MAE and SSIM. We also have made a comparison in terms of performance measures with some recent works, the result of which is given in Table 6 of results section. The methodology and result sections explain more about the method and obtained results. Medical images such as MRI, CT, PET were taken from [16, 17].

## 2 Related Works

The research work in [9], has taken MRI and SPET images for performing the experiments in transform domain, and fusion methods like DCT, DWT, SWT, and Hadamard transform are applied. To measure the performance, spatial frequency, entropy, correlation, visibility, and standard deviation are found out. Mutual Information and Quality factors are used for comparison.

DWT is most commonly used as a Fusion technique for MRI and PET scans [6]. In [14] the input image is filtered and DWT is applied to get the fused image which contains more information than the both input images. [2] presents a method of Medical image fusion using the transform technique where DWT and SWT are used to perform the fusion in the transform domain for MRI and PET scans. Entropy was used to measure the image quality. The conclusion was that SWT is more efficient than DWT. This is where we got to know that SWT is better than the other transforms like DWT and DCT.

The authors in [6] says that they used some normal methods like PCA, DWT, DSWT for the medical image fusion and they didn't get the results they expected, so they have used the hybrid method-that is amalgamating two or more normal methods. They have achieved it by Average-DWT-Maximum and PCA-DWT-DSWT; these two methods have got high PSNR compared to the traditional methods. On the other hand, based on some studies in [5], the authors say the that: on comparison of the outputs of DWT and CNN for the input image fusion of PET and MRI, CNN shows a larger Signal to Noise Ratio and entropy, thus giving a better result when compared with others.

The researchers proposed different fusion methods for medical applications in [12]. Pixel, region, and decision levels fusion methods are used for image fusion in [11, 15]. The spatial methods are used to perform image fusion in [7, 12], wherein the principal component analysis (PCA), average, maximum, contrast pyramid, and minimum methods are included [7, 12]. In all these spatial domain methods, the major problem was to overcome the problem of decreasing signal to noise ratio. We can overcome this problem by making use of transform based techniques, such as

the discrete cosine transform (DCT) and stationary wavelet transform (SWT), [13, 15]

## 3 Methodology

Proposed Image fusion method using SWT with Hadamard Transform consists of the following steps:

**Step 1:** Take two source images im1 & im2 which are to be fused.

**Step 2:** Perform Stationary Wavelet Transform on both images.

SWT result consists of lower frequency (LL) and higher frequency (LH, HL & HH) bands.

**Step 3:** Divide the LL band obtained into 8*8 blocks and perform Hadamard Transform on each of the blocks.

**Step 4:** For the Hadamard Transform outputs of every 8*8 block obtained in step 3 for both the input images, we calculate the gradient, visibility, spatial frequency, correlation, and entropy, definitions of which are given below (Eqs. (1) to (7)) and do a comparison as given in Eq. (8) to obtain the best block. The Eq. (8) defines the fusion rule.

### *3.1 Gradient (G)*

Better quality block will have high gradient [18] using Eq. 1.

$$G = \frac{1}{(P-1)*(Q-1)} \sum_{r=1}^{P-1} \sum_{s=1}^{Q-1} \sqrt{\frac{1}{2}\left\{\left(\frac{dA(r,s)}{dr}\right)^2 + \left(\frac{dA(r,s)}{ds}\right)^2\right\}} \tag{1}$$

### *3.2 Visibility (VIS)*

It can be defined as a measure to verify the degree of transparency of a block using Eq. 2.

$\beta$ = mean, $\sigma$ = visual contrast and it is lies between 0.6–0.7 [10].

$$VIS = \sum_{r=1}^{P} \sum_{s=1}^{Q} \left| A(r,s) - \frac{\beta}{\beta^{\sigma+1}} \right| \tag{2}$$

### 3.3 Spatial Frequency ($S_f$)

It determines the sharpness of a block using Eq. 3 to Eq. 5.

$$S_f = \sqrt{(R_f)^2 + (C_f)^2} \tag{3}$$

where

$$R_f = \sqrt{\frac{1}{P * Q} \sum_{r=1}^{P} \sum_{s=1}^{Q} [A(r,s) - A(r,s-1)]^2} \tag{4}$$

and

$$C_f = \sqrt{\frac{1}{P * Q} \sum_{r=1}^{P} \sum_{s=1}^{Q} [A(r,s) - A(r-1,s)]^2} \tag{5}$$

### 3.4 Correlation Coefficient (C)

It is a measure of correlation between two blocks (X and Y) in [3] by using Eq. 6.

$$C = \frac{\sum_{r=1}^{P} \sum_{s=1}^{Q} \left(X(r,s) - \overline{X}\right)\left(Y(r,s) - \overline{Y}\right)}{\sqrt{\sum_{r=1}^{P} \sum_{s=1}^{Q} \left(X(r,s) - \overline{X}\right)^2 \sum_{r=1}^{P} \sum_{s=1}^{Q} \left(Y(r,s) - \overline{Y}\right)^2}} \tag{6}$$

### 3.5 Entropy (E)

It is used to measure the information contained in the block using Eq. 7 and Eq. 8.

$$E = \sum_{r=1}^{p} \sum_{s=1}^{Q} p(A(r,s)) loglog[p(A(r,s))] \tag{7}$$

### 3.6 Fusion Rule

$$H(r,s) = \begin{cases} h1 & if\ G1 > G2,\ VIS1 > VIS2,\ S_f1 > S_f2,\ C1 > C2, E1 > E2. \\ h2 & if\ G1 < G2,\ VIS1 < VIS2,\ S_f1 < S_f2,\ C1 < C2, E1 < E2. \\ \frac{h1+h2}{2} & Otherwise. \end{cases} \tag{8}$$

where h1 and h2 are the Hadamard Transform Coefficients obtained when the transform is applied to LL band data of SWT of the blocks. H(r,s) is the final Hadamard Transform coefficient after the Fusion process applied to Hadamard coefficients.

**Step 5:** Apply Inverse Hadamard Transform to get New LL band (NLL).

**Step 6**: Apply PCA in [20] to the higher bands.

**Step 7**: Perform ISWT to both newly formed higher bands and LL bands to get the final fused image. The whole process which is briefly described in the above steps is clearly shown in the block diagrams in Figs. 1 and 2 and can easily be understood.

The widely used image fusion techniques such as DWT, SWT, DCT, and Hadamard transform are studied and analyzed for images from different domains and compared the results. The study showed that among all the four techniques,

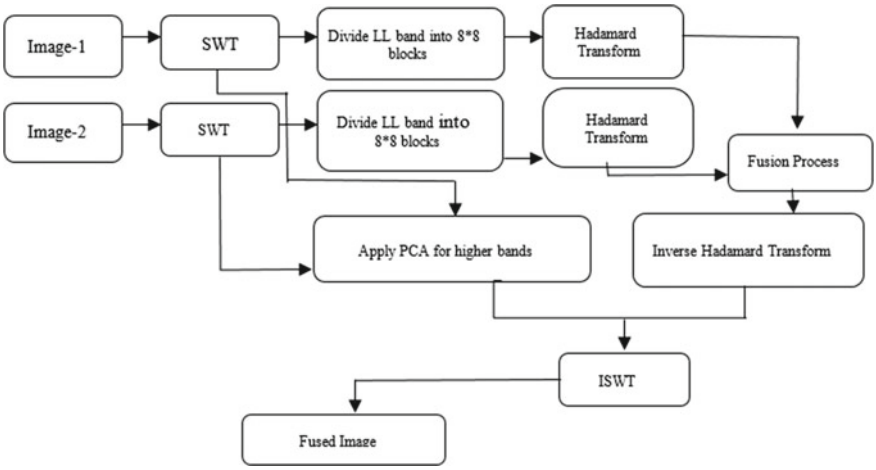



**Fig. 1** The layout of the Process flow

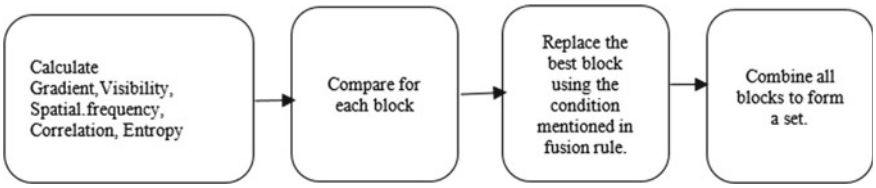



**Fig. 2** The layout of the Fusion process

SWT and Hadamard transform has given good results. SWT gives out maximum image features with more compression time whereas Hadamard Transform gives fewer image features with less compression time. To overcome the limitations of SWT and Hadamard Transform when applied alone, a new method by combining the above two transforms, a hybrid method, which is the Stationary Wavelet Transform along with Hadamard Transform, is used to obtain more image features useful for improving the quality of output. In addition to this, PCA is used which helps to reduce the larger data sets into smaller ones but still contains higher information that is good image features that are useful for fusion. Thus proposed method is used for image fusion. The block diagram representation of the method given in Fig. 1 and the fusion process in Fig. 1 is discussed in Fig. 2.

## 4 Results

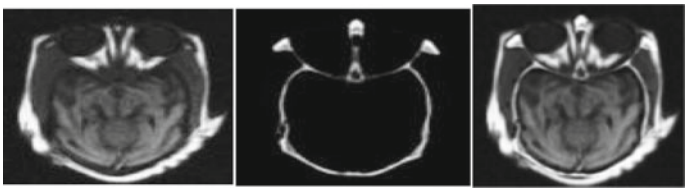
Medical images [14], Multimodal and other images of spatial resolution 256 × 256 pixels are given as inputs to our fusion method. The input images taken for fusion and fused images are shown in Figs. 3, 4, 5, 6 and 7. The performance evaluation is done using PSNR, RMSE, CORR, MAE, and SSIM which are defined as given below. Let us consider the ground image GI(r, s) and the fused image FI(r, s) of size P x Q.

(a) (b) (c)

**Fig. 3** Saras **a** and **b** Input images, **c** Output image

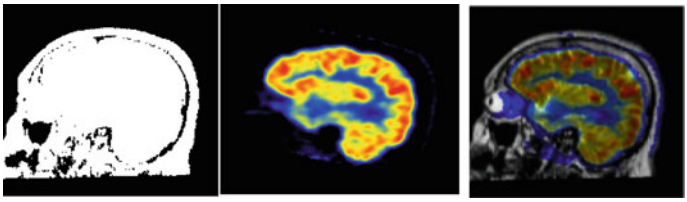
(a) (b) (c)

**Fig. 4** Pepsi **a** and **b** Input images, **c** Output image



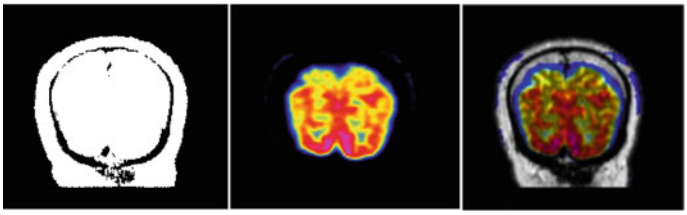
(a) (b) (c)

**Fig. 5** MRI and CT **a** and **b** Input images, **c** Output image



(a) (b) (c)

**Fig. 6** MRI-1 **a** and **b** Input images, **c** Output image



(a) (b) (c)

**Fig. 7** MRI-2 **a** and **b** Input images, **c** Output image

## *4.1 Peak Signal to Noise Ratio*

The Peak Signal to Noise Ratio [7, 19] is used to measure signal power to noise power ratio. Its maximum value is 100 and its minimum value is 0. Its expressed as Eq. 9.

$$PSNR(dB) = 10 * log_{10}\left[\frac{255^2}{\frac{1}{P*Q}\sum_{r=0}^{P-1}\sum_{s=0}^{Q-1}(\boldsymbol{GI}(r,s) - \boldsymbol{FI}(r,s))^2}\right] \quad (9)$$

### 4.2 Root Mean Square Error

It is used to estimate the closeness between the ground image and the fused image. Lower the value of RMSE indicates high closeness between ground and fused image [8] which is expressed in Eq. 10.

$$RMSE(dB) = \sqrt{\frac{1}{P * Q} \sum_{r=0}^{P-1} \sum_{s=0}^{Q-1} |GI(r,s) - FI(r,s)|} \tag{10}$$

### 4.3 Correlation (CORR)

Computes correlation between FI and GI. Value range (0,1). A value near 1 indicates more information and if near to 0 indicates the information is low which is expressed in Eq. 11.

$$C_{x,y} = \frac{\sum_{r=1}^{P} \sum_{s=1}^{Q} \left(\boldsymbol{GI}(\boldsymbol{r},\boldsymbol{s}) - \overline{\boldsymbol{G}}\right)\left(\boldsymbol{FI}(\boldsymbol{r},\boldsymbol{s}) - \overline{\boldsymbol{FI}}\right)}{\sqrt{\sum_{r=1}^{P} \sum_{s=1}^{Q} \left(\boldsymbol{GI}(\boldsymbol{r},\boldsymbol{s}) - \overline{\boldsymbol{G}}\right)^2 \left(\boldsymbol{FI}(\boldsymbol{r},\boldsymbol{s}) - \overline{\boldsymbol{FI}}\right)^2}} \tag{11}$$

### 4.4 Mean Absolute Error (MAE)

MAE is used to calculate the absolute error value between original and fused images using Eq. 12.

$$MAE = \frac{1}{P * Q} \sum_{r=1}^{P} \sum_{s=1}^{Q} abs(GI(r,s) - FI(r,s)) \tag{12}$$

The smaller value of MAE indicates finer similarities between ground and fused images.

### 4.5 Structural Similarity Index (SSIM)

It is used to measure the similarities between original and fused images. It is defined as Eq. 13.

$$SSIM = \frac{((2\mu_{FI}\mu_{GI} + c1) \times (2\sigma_{GF} + c2))}{(\mu_{FI}^2 + \mu_{GI}^2 + c1) \times (\sigma_{FI}^2 + \sigma_{GI}^2 + c2)} \tag{13}$$

where $\mu_{FI}$ and $\mu_{GI}$ are the average intensities of fused and ground images, $\sigma_{FI}$ and $\sigma_{GI}$ are the variances of fused and ground images, c1 and c2 are constants. The index value of SSIM is $\pm$ 1. If the index value is 1, it indicates that the source and fused images are the same.

The results obtained after fusion for different input images using our method are shown below.

Visually itself it can be understood that our method gives better quality output compared to the inputs given for fusion.

A comparative analysis of different fusion methods in terms of performance metrics are shown in the tables below:

The quantitative analysis of the performance metrics such as PSNR, RMSE, Correlation, MAE, and SSIM for the different types of images are shown in the tables from Tables 1, 2, 3, 4 and 5. For every input image, the performance measures are best

**Table 1** Comparison of performance measures for Saras Image

| Image | Method | PSNR | RMSE | CORR | MAE | SSIM |
|---|---|---|---|---|---|---|
| SARAS | DWT | 24.4862 | 15.2135 | 0.9978 | 4.6262 | 0.9603 |
| | SWT | 28.1231 | 10.0089 | 0.999 | 3.8974 | 0.9830 |
| | DCT | 25.1159 | 14.1496 | 0.9981 | 7.7225 | 0.9666 |
| | HT | 25.1231 | 10.0089 | 0.999 | 3.8974 | 0.983 |
| | PROPOSED | 31.9311 | 6.4563 | 0.9996 | 1.973 | 0.9933 |

**Table 2** Comparison of performance metrics for Pepsi Image

| Image | Method | PSNR | RMSE | CORR | MAE | SSIM |
|---|---|---|---|---|---|---|
| PEPSI | DWT | 30.0281 | 8.01 | 0.9973 | 3.6531 | 0.9852 |
| | SWT | 31.7907 | 6.5615 | 0.9981 | 3.4457 | 0.9903 |
| | DCT | 30.3638 | 7.733 | 0.9974 | 4.543 | 0.985 |
| | HT | 33.4037 | 5.4495 | 0.9987 | 3.038 | 0.9987 |
| | Proposed | 36.6782 | 3.7379 | 0.9994 | 1.8939 | 0.9965 |

**Table 3** Comparison of performance metrics for MRI and CT Image

| Image | Method | PSNR | RMSE | CORR | MAE | SSIM |
|---|---|---|---|---|---|---|
| MRI & CT | DWT | 10.0717 | 79.9756 | 0.2217 | 3.6531 | 0.9852 |
| | SWT | 25.8895 | 12.9439 | 0.9888 | 7.1596 | 0.9783 |
| | DCT | 11.793 | 65.5983 | 0.5865 | 51.142 | 0.2402 |
| | HT | 15.7422 | 41.6321 | 0.8261 | 28.308 | 0.6478 |
| | Proposed | 26.7243 | 11.7578 | 0.9911 | 6.7072 | 0.9983 |

**Table 4** Comparison of performance metrics for MRI-1 Image

| Image | Method | PSNR | RMSE | CORR | MAE | SSIM |
|---|---|---|---|---|---|---|
| MRI-1 | DWT | 53.5287 | 0.5372 | 0.6313 | 0.3083 | 0.6952 |
| | SWT | 65.4023 | 0.1369 | 0.9833 | 0.0202 | 0.9833 |
| | DCT | 57.5743 | 0.3372 | 0.8897 | 0.1156 | 0.8924 |
| | HT | 70.3248 | 0.0777 | 0.9945 | 0.0073 | 0.9946 |
| | Proposed | 70.9123 | 0.0726 | 0.9953 | 0.0061 | 0.9952 |

**Table 5** Comparison of performance metrics for MRI-2 Image

| Image | Method | PSNR | RMSE | CORR | MAE | SSIM |
|---|---|---|---|---|---|---|
| MRI-2 | DWT | 54.9415 | 0.4565 | 0.683 | 0.2228 | 0.9889 |
| | SWT | 66.6348 | 0.1188 | 0.9828 | 0.0153 | 0.9828 |
| | DCT | 60.2159 | 0.2487 | 0.9239 | 0.0638 | 0.9954 |
| | HT | 70.0475 | 0.0802 | 0.9920 | 0.0111 | 0.0997 |
| | Proposed | 72.4964 | 0.0605 | 0.9955 | 0.0043 | 0.9999 |

(Highest PSNR, least RMSE, Highest Correlation, least MAE and highest SSIM) for our method.

We did a comparison of PSNR values for our method with the results in [1, 4], RMSE values with the results in [1], MAE and SSIM values with the results in [4]. Table 6 shows a comparison of various performance measures.

The performance metrics of the proposed method are better when compared to the values mentioned in the research papers.

**Table 6** Comparison of performance metrics of our method with that in Research paper [4]

| Image | Method | PSNR | MAE | SSIM |
|---|---|---|---|---|
| Saras | SFDWT [4] | 29.1705 | 3.4824 | 0.9864 |
| | **Proposed** | **31.9311** | **1.973** | **0.9933** |
| Pepsi | SFDWT [4] | 35.3338 | 2.4741 | 0.9951 |
| | **Proposed** | **36.6782** | **1.8939** | **0.9965** |
| MRI&CT | SFDWT [4] | 24.8183 | 8.1639 | 0.9720 |
| | **Proposed** | **26.7243** | **6.7072** | **0.9983** |
| MRI & PET | SFDWT [4] | 63.1320 | 0.0765 | 0.9995 |
| | **Proposed** | **72.4964** | **0.0043** | **0.9999** |

## 5 Discussion and Conclusion

It is observed that the proposed method, SWT with Hadamard Transform has shown better results when compared to fusion methods using DWT, DCT, SWT, and Hadamard Transform and are giving similar results, in terms of various performance metrics, as obtained in some other work. The performance metrics of ideal fused image include high PSNR, low RMSE, low MAE, Correlation and SSIM values approximately equal to 1. The values we obtained for our method for these performance metrics are near optimum values. Thus, using the proposed method, the performance metrics such as PSNR, RMSE, CORR, MAE, and SSIM are improved.

On this note, it is concluded that by using image fusion method that is SWT with Hadamard Transform we have obtained fused images which have better image quality.

## References

1. Amala RV, Lalithakumari S (2020) A hybrid fusion model for brain tumor images of MRI and CT. In: 2020 international conference on communication and signal processing (ICCSP). IEEE, pp 1312–1316
2. Ashwanth B, Swamy KV (2020) Medical image fusion using transform techniques. In: 2020 5th international conference on devices, circuits and systems (ICDCS). IEEE, pp 303–306
3. Tyagi T, Gupta P, Singh P (2020) A hybrid multi-focus image fusion technique using SWT and PCA. In: 2020 10th international conference on cloud computing, data science & engineering (confluence). IEEE, pp 491–497
4. Niranjan B, Kullayamma I (2019) Image fusion using spatial frequency discrete wavelet transform and TYPE-2 fuzzy logic. Int J Eng Res Technol (IJERT) 08(11):1–6
5. Muthiah MA, Shanmugam EL, Reddy BVK (2019) Fusion of MRI and PET images using deep learning neural networks. In: 2019 2nd international conference on power and embedded drive control (ICPEDC). IEEE, pp 283–287
6. Prasad P, Subramani S, Bhavana V, Krishnappa HK (2019) Medical image fusion techniques using discrete wavelet transform. In: 2019 3rd international conference on computing methodologies and communication (ICCMC). IEEE, pp 614–618
7. Mamata MT, et al (2018) Image fusion based on DWT type-2 fuzzy logic system. IJSER 9(5):716. ISSN 2229–5518
8. Raju PR, Rani DS, Challaram G (2018) Comparison of medical image fusion methods using image quality metrics. In: 2018 international conference on communication, computing and internet of things (IC3IoT). IEEE, pp 449–454
9. Haribabu M, Hima Bindu C (2017) Feature level based multimodal medical image fusion with Hadamard transform. Int J Control Theory Appl 9(42):453–460
10. Haribabu M, Bindu CH (2017) Visibility based multi modal medical image fusion with DWT. In: 2017 IEEE international conference on power, control, signals and instrumentation engineering (ICPCSI). IEEE, pp 1561–1566
11. Sharma A, Gulati T (2017) Novel fusion rules for discrete wavelet transform based image fusion. IJST 10:19
12. Sanjay AR et al (2017) CT and MRI Image fusion based on discrete wavelet transform and type-2 fuzzy logic. IJIES 10(3):355–362
13. Kumar A, et al (2017) Image fusion using DTCWT with High Boost Filtering. IJEECE 6(1):16–23. ISSN No (Online) 2277–2626.

14. Vadhi R, Kilari VS, Kumar SS (2016) An image fusion technique based on hadamard transform and HVS. Eng Technol Appl Sci Res 6(4):1075–1079
15. Mishra Dhirendra, Palkar Bhakti (2015) Image fusion techniques: a review. Int J Comput Appl 130(9):7–13
16. http://www.med.harvard.edu/AANLIB/home.html
17. Rajini KC, Roopa S (2017) A review on recent improved image fusion techniques. In: 2017 international conference on wireless communications, signal processing and networking (WiSPNET). IEEE, pp 149–153
18. Deepa B, Sumithra MG, Bharathi TD, Rajesh S (2017) MRI medical image fusion using gradient based discrete wavelet transform. In: 2017 IEEE international conference on computational intelligence and computing research (ICCIC). IEEE, pp 1–4
19. Polinati S, Dhuli R (2019) A review on Multi-Model medical image fusion. In: 2019 international conference on communication and signal processing (ICCSP). IEEE, pp 0554–0558
20. Asish Reddy KS, Kumar KK, Kumar KN, Bhavana V, Krishnappa HK (2019) Multimodal medical image fusion enhancement technique for clinical diagnosis. In: 2019 3rd international conference on computing methodologies and communication (ICCMC), 2019, pp 586–589. https://doi.org/10.1109/ICCMC.2019.8819840

# A Review on Smart Road Traffic Management System Using LoRa WAN


**Naga Raju Jangam, G. P. Ramesh, and P. Rachana**

**Abstract** Smart Road Management System using LoRa WAN technology is helpful for the traffic police to clear a path for the Emergency vehicles at intersections to save human life and heavy damages due to any kind of fire accidents. This technology may be very useful to detect the road accidents and may provide density based traffic control system due to its versatile nature of being dynamic nature to connect with different end-nodes. It can be used as a tool to provide hierarchical road infrastructures for public transportation. This survey tends to help the users and traffic management systems by providing reliable information of the real-time traffic. This work also helps to identify the best methods/solutions to identify the rash drivers and to minimize the road accidents. Also, help to propose a method for easy exit path for the emergency vehicles and VIP passage within 1–2 km rages before the vehicle arrival time. To make efficient automatic traffic Chelan fine system for traffic rule breakers and also to design a density based traffic light system.

**Keywords** LoRa WAN · Road traffic management · Road site units · Smart road management system · Traffic light system

## 1 Introduction

The traffic management system is one of the oldest areas of discussions for a long time due to changing behavior of commuters with respect to facilities, innovative technologies, increasing population, and adamant behavior of commuters in different scenarios. The human- controlled (i.e. with a traffic police officer) traffic includes lot of intelligence and human brain to respond quickly for certain situations to take decisions and to maintain the safe and peaceful traffic flow. However, with the exponential rise in the population and vehicles on the roads demand wide range

N. R. Jangam (✉) · G. P. Ramesh
St. Peters Institute of Higher Education and Research, Avadi, Chennai 600054, India
e-mail: nagaraju.jangam@mlrinstitutions.ac.in

P. Rachana
New Horizon College of Engineering, Bangalore, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_5

of roads and monitoring eyes with more controlling facilities to maintain a systematic traffic flow on the roads. Advancement of technologies helped to various traffic management systems/departments across the globe, using wide range of electronic devices and computer aided tools to monitor, control, and guide the commuters on the roads in a proper direction. Sometimes these devices/tools also helped to identify the mischievous behavior of different individuals on the roads to punish them and to avoid future any kind of such incidents on the roads. However, these devices are functioning based on the instructions set stored in the central processing units (CPU) of the traffic control room; and cannot take certain decisions based on emotions, emergencies and on the urgency to a certain situation. Of course, there are some devices installed on the roads, which works based on the traffic density on certain roads and are functioning fine with the road management but human intelligence is missing all the way. On the other hand, these devices are directly/ indirectly controlled by satellite communication systems in most of the developed countries. They use extensive technological advancements of satellite technologies to operate, guide, monitor and to take quick decisions remotely from far-end distances. However, many of the countries do not have proper satellite communication facilities even to carry out various other important activities in their countries and depends on the neighboring countries or hardly could afford to buy a satellite for different purposes. Only twelve (12) Countries out of 300 have the capacity to build and launch their own satellites and rest of the nations need to depend on these twelve countries. Sometimes the satellite based systems does not work effectively on highways with low infrastructure and facilities, which in turn face the issues of low signals or signal dropouts. Especially, at the time of natural disasters such as heavy rains, winds, thunderstorms, etc. these kinds of problems are mostly heard and experienced at everywhere in the world. Therefore, it is very much essential to address a solution which can avoid expensive satellite technology for the road traffic management services using some of the smart devices which are affordable and can be effective to monitor and communicate the traffic information to the commuters on the roads. In the recent times, wide area networks (WAN) are popularly known for long range communication systems used for different types of traffic management systems. Low power & long range (LoRa) WANs are widely used by the traffic departments for effective usage due to the advantage of using this network with lower physical layers. Lora WAN works as a network layer protocol, which is based on the cloud-based medium access control (MAC) protocols. These protocols helps to establish a communication between the Lora WAN gateways and the end-nodes like a routing protocol. This technology gives a cutting-edge traffic management systems and methods to implement in very busy traffic roads in an efficient manner. This technology is also they responsible for the management of all communication frequencies, data rates, and power supplies. Apart from these it is highly reliable for moderate loads and with some of the issues related with performance while sending the acknowledgements.

In this paper, a detailed review on LoRa WAN is carried out to use it for the applications of smart road traffic management systems. In the Sect. 2, a background study on LoRa WAN will be discussed and the purpose of the LoRa WAN based traffic management systems in India will be discussed in detail. In the Sect. 3, different

types of innovative ideas and implementations using Lora WAN are discussed with appropriate research outcomes and deliverables. In the Sect. 4, a summarized SWOT analysis is carried out towards the LoRa WAN technology for a traffic management system. Finally, the Sect. 5 provides the references of authors' overall studies on the Road Traffic Management system design approaches with various development techniques and also LoRa WAN Module Interface.

## 2 Analysis of LoRa WAN Technology

Introduction of WAN helped most of the engineers to solve the problems related with remote access. To connect different local area networks installed at far-ends can be obtained by using WAN, covering larger geographical areas. The major advantage of WAN is with respect to the privacy, security, and consistency. In general, there are two popularly known circuit-switched WAN technologies are available in the market for the public usage. They are public switched telephone network (PSTN), and integrated services digital networks (ISDN). Later, low power based long range (LoRa) WAN has been introduced, which uses the spread spectrum modulation techniques that comes out of chirp spread spectrum (CSS) technology. This technology was introduced by Cycleo of France and later acquired by SEMTECH, which is also a founding member and alliance.

### *2.1 Technological Factors Related with LoRa WAN*

Using LoRa WAN the communication between two nodes can be established for a distance up to 15 km. Major advantage of this technology lies with the low power consumptions for covering a long range as compared to any other wireless system. It maintains different types of internet of things (IoT) standards and protocols are used with battery powered systems. The range of bytes for data transmission varies between 51–222 bytes in LoRa. The cellular networks are managed by a controller from a defined base station, but the LoRa WAN devices need not link up with any specific gateway and can operate independently as well. Gateways in this technology can forward the data to the network servers like any other link- layers and are allowed to add up couple of the bytes to payload. The network server network server ensures an appropriate gateway is identified for replying any particular nodes; and these gateways behave transparently with all end nodes. LoRa WAN is one of the fastest growing technologies and more than 50 countries deployed this network and presently many small countries are trying to implement for various IoT based network infrastructure development programs. Some of the countries using this technology for various applications such as waste management, GPS-less tracking, etc. due to high security and low installation expenses.

**Table 1** Shows different features of LoRa WAN as compared with LTE

| Features | LoRa WAN | Narrow-Band | LTE Cat-1 | LTE Cat-M | NB-LTE |
|---|---|---|---|---|---|
| Modulation | SS-Chirp | UNB/GSK/BPSK | OFDMA | OFDMA | OFDMA |
| Rx bandwidth | 500-125 k Hz | 100 Hz | 20 MHz | 20–1.4 M Hz | 200 kHz |
| Data range | 290-50 K bps | 100bps or 12/8 bytes | 10 Mbps | 200 K–1 M bps | 20 Kbps |
| Max. O/P power | 20 dBm | 20 dBm | 23–46 dBm | 23/30 dBm | 20 dBm |
| Battery life time | 105 months (9 Years) | 90 months (7.5 Years) | | 18 months (1.5 Years) | |
| Link budget | 154 dB | 151 dB | 130 dB | 146 dB | 150 Db |
| Security | Yes | No | Yes | Yes | Yes |

## *2.2 Advantages of LoRa WAN*

The main advantage of LoRa WAN is using the low-level physical layer with ISM bands of 868 and 915 MHz are available in the world for free to use. Low power consumption by these devices makes it more suitable for battery-powered applications/devices. The battery life- time using LoRa based devices can last for more than 9 years and which is far beyond any other type of wireless networks. The distance covered by a LoRa WAN is between 5–15 km based on urban and sub-urban areas respectively. The advantage of low power consumption using LoRa WAN is possible due to the usage of adaptive data rate (ADR), which helps to vary the output data rates based on the payload coupled with the CSS. On the other hand, frequency shift keying (FSK) modulation techniques are being implemented in this technology allows using low cost and high power amplifiers. This also helps to achieve the low cost connectivity aspect of any network provider. A single gateway of LoRa WAN accommodates up to 1000 end-nodes makes this technology to handle high network capacities. It also an interoperable technology due to the standardization process taken place at the time of design and implementation of LoRa architecture. Some of the key features of LoRa WAN as compared with the long-term evolution (LTE) technology are shown in Table 1.

## *2.3 LoRa WAN Based Traffic Management Systems in India*

India is the second largest populated country in the world with diversified cultures and religious practices. Due to the lack of proper Traffic rules awareness, it is seen that the people break the traffic signals/rules, due to which blockages occur on the roads frequently. Sometimes it is also observed to see the emergency service vehicles

are not in a position to come out of the traffic jams and traffic inspectors find it a very challenging task to deal with such scenarios.

## 3 Implementation Techniques of LoRa WAN for Traffic Management

Nellore and Hancke [1] recommended a design to be done by using the WSN based urban traffic management system. Such systems will optimally manage the traffic congestion by prioritizing the vehicle flow, provides exit way for emergency vehicles, reduce the waiting time of vehicles and fuel consumption, and finally good safety measures. Later, the work proposed by Santa et al. [2] mainly focused on a real monitoring platform for vehicles using LPWAN has been designed and developed, and the performance study of LoRa WAN technology in vehicular scenarios has been conducted. A generic Linux interface is used that enables direct and seamless transmission of application-layer data. It can boost the synergy between intelligent transport systems (ITS) and IoT. The major limitations of this work are with respect to the Data compression and data processing techniques implemented for the traffic management system. In order to inspect the sensors available in the market and to study their strengths and weaknesses Padmavathi et al. [3] carried out a research on the vehicle detection and tracking algorithms. Each algorithm used for detection and tracking compares vehicle system existence based on different types of evaluation parameters. In this work, the particle filtering is carried out based on the Bayesian TBD estimator algorithm that sounds the best for vehicle detection and tracking in WSNs. Bayesian TBD estimator algorithm used for vehicle detection and tracking. Finally, the propagation delays in this work are varied based on vehicle tracking. Arbabi et al. [4] focused on combining the best data from the traffic network to provide traffic information, directions and driver services in an intelligent manner. The dynamic traffic monitoring (DTM) system proposed in this work can provide high quality travel time and SMS data in transient flow traffic (which is caused by non-recurring congestion). This method claimed to provide better performance in monitoring message delivery methods. This also provides better penetration rate, message delay and acceptable latency. Measurements may incur some delay caused by vehicles having to carry their data to a nearby TO and low market penetration rates. In a work presented by Ahmad et al. [5] focuses on wireless sensor based road side architecture for intersection traffic data collection and control. It is capable of accurate real time data collection utilizing the IEEE 802.15.4 protocol. It can be leveraged by intelligent transport system (ITS) for implementing numerous decisions depending on the underlying policy. It reduces communication overhead, energy consumption, connectivity delay and interference from neighboring intersections. Hence, the taxonomy of traffic management system is represented in Fig. 1.

Collota et al. [6] proposed a novel approach to dynamically manage the traffic lights using the cycles and phases in an isolated intersection with Fuzzy logic. Fuzzy

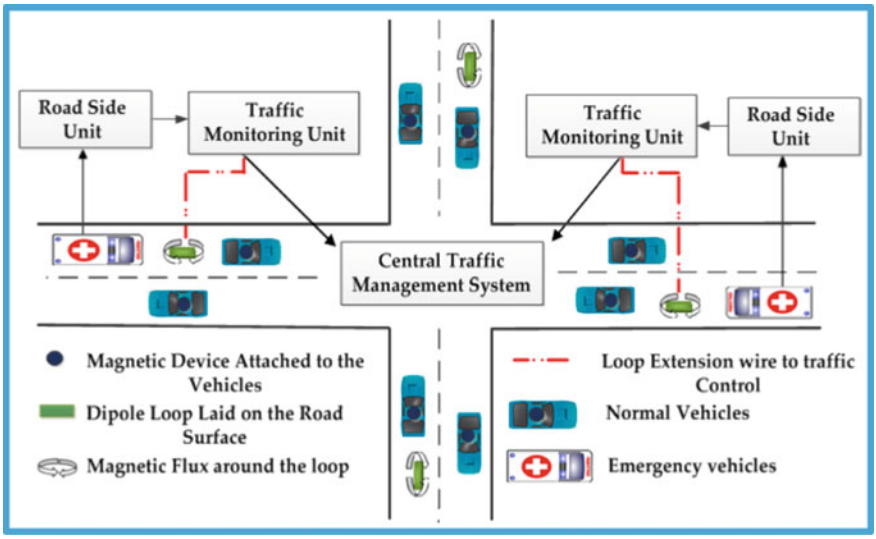



**Fig. 1** Taxonomy of traffic management system

controller addresses vehicles turning movements and dynamically manages both the phase and the green time of traffic lights. It has easy deployment and less maintenance. The advantages of this method claimed with respect to the flexibility, low-cost, non-invasiveness, and scalability. It is limited to certain level of Traffic intensity and RSU design unit is cost effective and complex. The work proposed by Sivanandam et al. [7] focused on designing architecture for advanced traveler information system (ATIS) and data collection techniques. The implementation of data fusion models is based on MATLAB software. This method is used to estimate the dynamic link travel prediction and the response time is varied with the vehicle distance. SEMTECH includes the network layer too so it is possible to send the information to any Base Station already connected to a Cloud platform. LoRa WAN modules may work in different frequencies by just connecting the right antenna to its socket. Long range data transmission License free ISM band ranges. It has complexity in design [8]. Various security concerns were discussed by Oniga et al. [9, 10] showed different data protection and the data privacy methods used in LoRa WAN protocol specifications. An in-depth analysis towards security aspects was proposed by the authors by proposing architecture to protect the data transmission and to prevent any kind of data loss due to unauthorized access. The testing results of this work on the end-nodes on sniffing LoRa traffic revealed that the message will be encrypted and are signed with AES128 keys at the time of communication. This in turn offers strong protection against the intercepting and data reading process during transit times. Similarly, the authors also presented a secured LoRa WAN sensor network architecture to address wide range of security threats and mitigation mechanisms along with variety of security recommendations for a safe and protected data transmission. This architecture also claimed to be helpful for preventing the unauthorized data access

and losses in LoRa WAN networks. Similar attempt was made for LoRaWANv1.1 by Donmez and Nigussie determined various security related problems [11]. The authors revealed that LoRaWANv1.1 does not provide any kind of forward secrecy due to the operations involved without human intervention which deals sometimes with sensitive data. Five security vulnerabilities related with LoRa WAN were discussed by Yang et al. [12] included with i) denial-of-service (DoS) attacks on IoT devices, ii) plaintext recovery, iii) malicious message modifications, iv) falsification of the delivery reports, and v) battery exhaustion attacks/The authors highlighted different types of security features, activation methods, key management systems, counter management systems, cryptography, and message acknowledgment systems related with various securities related issues.

An intelligent transportation system (ITS) proposed by Salazar-Cabrera et al. [13] focused on the urban public transportation to track the public vehicles using IoT systems and their concepts. In this work, the authors used serial monitor in Arduino IDE for reading the messages sent or received to test their method by using the event processing module. Similar work was carried out by Divij et al. [14] for controlling the traffic lights at the time of emergency situations. Different types of acoustic sensors and object detection methods were proposed in this work to alert both commuters and traffic management system to control and manage the traffic on the roads. To address the traffic congestion and monitoring Nor et al. [15] suggested a smart traffic light using LoRa WAN. This system claimed to be sufficient enough to replace the old system by monitoring congestion levels using IoT technology. On the similar lines, Sudeep et al. [16] proposed a quicker ambulance service using LoRa based system after accident detection. Thangam and Vikram [17] proposed a low cost, long range, automated monitoring and communication system using LoRa for emergency services in the rural areas. The authors claimed to support the drivers by reducing the burden of selecting an appropriate route at the time of emergency situations. Technically a sound question was raised by Pop et al. [18] while addressing the bi-directional traffic. The authors highlighted whether the bi- directional traffic is safe or harmful using LoRa WAN based networks. In this work, the authors also highlighted the reliability trade-offs versus energy consumption during retransmission attempt. A detailed energy analysis was carried out on LoRa WAN technology by Mathur et al. [19] for different types of traffic sensing applications. Various features and inadequacies are explored in this work for providing an adaptive data aggression along with a re-transmission algorithm to understand the sensor data of related traffic. From this work, the authors intended to provide a solution due to which the sensor devices are capable of working for more than five years using small batteries. Oh et al. [20] also carried out research on similar lines to propose TRILO, an energy efficient downlink communication mechanism for the LoRa WAN technology. TRILO provides a trade-off that balances the latency and overall energy consumption where it alerts the end-devices periodically using beacon mechanism to alert all those devices which are having pending downlink frames [21]. In the recent times, integration of $5^{th}$ generation (5G) of mobile networks with LoRa WAN was under discussions. The challenges in the existing traffic management system is listed in Table 2.

**Table 2** Challenges in the existing traffic management systems

| Challenges | Definitions |
|---|---|
| Connectivity and coverage | The connectivity of most of the traffic management systems depends on the technology they use and infrastructure available in different locations. For example, the GPS may not be available in remote places due to the shortage of connecting infrastructure and therefore the coverage of a vehicle may not be possible for all the time and duration |
| Communication and energy cost | The communication may lose with the controller stations once the travelling is through a remote area and additional energy cost is needed to maintain the infrastructure. Installing a GPS substation may be an ordeal to for the management |
| Congestion | Traffic congestion is always a challenging task, especially in the busy working hours and managing so many vehicles with different challenges. Slow speed, long trip times and queuing of the vehicles create a serious situation on the roads with poor traffic management systems |
| Traffic incident notification | Alerting the commuters on time with appropriate information related to road accidents or traffic jams is one of the critical jobs of any traffic management system to avoid further road blocks, traffic jams, and accidents. Most of the times, especially during winter time, such type of delays in traffic incident notifications may cause serious accidents on the roads. Probability of losing human life and serious damage of the road properties may further create havocs in commuter's journey |
| Coordination and implementation | A lot of coordination is required by the traffic management systems to ensure a smooth process of events on the roads, without any kind of implementation hurdles. A lot of communication with positive intent and approach needs to be adopted for a good coordination between different departments and traffic management systems. Otherwise, it is going to be a disaster on the roads to maintain a pleasant journey for the commuters |

## 4 Conclusion

In this paper, a systematic literature review is carried out in order to enhance the existing state of research approach regarding smart traffic management design methods in the new technology design areas. This review paper is intended to identify the future design trends and propose key areas for the future research. This systematic literature review was based on a set of criteria developed through the state-of-the-art literature, as well as from best practices in the authentic need of the road Traffic Management design approach. The findings show that while relatively few studies have been conducted in the field thus far, the design approach is definitely garnering more research attention. The study analysis of existing system has

helped to identify the gaps and to define important areas for future research in the field. A deign using the LoRa WAN technology intended to increase the possibility of faster communication between the drivers, BS and Police control rooms for a good advanced smart road traffic management system. To make efficient automatic traffic Chelan fine system for traffic rule breakers and also to design a density based traffic light control system. This reduces road accidents and minimizes the number of traffic rule breakers. It also reduces the longer waiting time at the traffic signals when traffic density is less. Accidents can be minimized and a systematic, safer, and smarter transit can be developed by monitoring the driving behavior through video surveillance. The solution can also be then linked with central traffic system and customized to work in conjunction with other technologies, like face recognition, automatic license plate recognition. Traffic diversion can be planned at peak hours on a daily, monthly or yearly event basis. It should be more accurate, predictive and must be automated using artificial intelligence (AI) and Machine learning (ML) systems along with this solution. It is very important for the local police officials to make the public aware of such events by conducting pilot exercises on roads once in a while in a month/week if possible.

## References


1. Nellore K, Hancke GP (2016) A survey on urban traffic management system using wireless sensor networks. Sensors 16(2):157
2. Santa J, Sanchez-Iborra R, Rodriguez-Rey P, Bernal-Escobedo L, Skarmeta AF (2019) LPWAN-based vehicular monitoring platform with a generic IP network interface. Sensors 19(2):264
3. Padmavathi G, Shanmugapriya D, Kalaivani M (2010) A study on vehicle detection and tracking using wireless sensor networks. Wirel Sens Netw 2(02):173
4. Arbabi H, Weigle CM (2010) Using DTMon to monitor transient flow traffic. In: Proceedings of the IEEE vehicular networking conference (VNC), NJ, USA, 13–15 December 2010, pp 110–117
5. Ahmad F, Basit A, Ahmad H, Mahmud AS, Khan MG, Yousaf ZF (2013) Feasibility of deploying wireless sensor based roadside solutions for intelligent transportation systems. In: Proceedings of the IEEE international conference on connected vehicles and expo (ICCVE), Las Vegas, NV, USA, 2–6 December 2013, pp 320–326
6. Collotta M, Bello LL, Pau G (2015) A novel approach for dynamic traffic light management based on wireless sensor networks and multiple fuzzy logic controllers. Expert Syst Appl 42:5403–5415
7. Sivanandam R, Devi VL, Ravi V, Kumar SK (2015) Advanced traveller information system (ATIS) for Indian Cities. https://coeut.iitm.ac.in/umcsp/pdfweb/v2iitm_ATIS%20For%20indian%20cities_v2.pdf, Accessed 14 Apr 2015
8. SEMTECH, 2019. What is LoRa?
9. https://www.semtech.com/lora/what-is-lora, Accessed 21 Dec 2019
10. Oniga B, Dadarlat V, De Poorter E, Munteanu A (2017) Analysis, design and implementation of secure LoRaWAN sensor networks. In: 2017 13th IEEE international conference on intelligent computer communication and processing (ICCP). IEEE, pp 421–428
11. Oniga B, Dadarlat V, De Poorter E, Munteanu A (2017) A secure LoRaWAN sensor network architecture. In: 2017 IEEE sensors. IEEE, pp 1–3

12. Dönmez TC, Nigussie E (2018) Security of join procedure and its delegation in LoRaWAN v1. 1. Procedia Comput Sci 134:204–211
13. Yang X, Karampatzakis E, Doerr C, Kuipers F (2018) Security vulnerabilities in LoRaWAN. In: 2018 IEEE/ACM third international conference on internet-of-things design and implementation (IoTDI). IEEE, pp 12–140
14. Salazar-Cabrera R, de la Cruz AP, Molina JMM (2019) Proof of concept of an IoT-based public vehicle tracking system, using LoRa (Long Range) and intelligent transportation system (ITS) services. J Comput Netw Commun 2019:1–10
15. Divij N, Divya K, Badage A (2019) IoT based automated traffic light control system for emergency vehicles using LoRa. IJSTE 6(1):23–29
16. Nor RFMA, Zaman FH, Mubdi S (2017) Smart traffic light for congestion monitoring using LoRaWAN. In: 2017 IEEE 8th control and system graduate research colloquium (ICSGRC). IEEE, pp 132–137
17. Sudeepa KR, Rashmi N, Prarthana DR, Sindhu S, Badiger VS (2018) LoRa based network for accident detection and providing quicker ambulance services for medical assistance. IJERT 6(13):1–3
18. Thangam MKS, Vikram VK (2019) Emergency vehicle and health monitoring system using LoRa. https://innovate.mygov.in/innovation/emergency-vehicle-and-health-monitoring-system-using-lora/, Accessed 12 Jan 2019
19. Pop AI, Raza U, Kulkarni P, Sooriyabandara M (2017) Does bidirectional traffic do more harm than good in LoRaWAN based LPWA networks? In: GLOBECOM 2017–2017 IEEE global communications conference. IEEE, pp 1–6
20. Mathur S, Sankar A, Prasan P, Iannucci B (2017) Energy analysis of LoRaWAN technology for traffic sensing applications. In: Intelligent transportation systems (ITS) world congress
21. Oh Y, Lee J, Kim CK (2018) TRILO: a traffic indication-based downlink communication protocol for LoRaWAN. Wirel Commun Mobile Comput 2018:1–14

# A Road Side Unit Based Proxy Signature Scheme for Fast Verification of Messages in Vehicular Ad-Hoc Network


**Farooque Azam, Sunil Kumar, and Neeraj Priyadarshi**

**Abstract** Vehicular Ad-hoc Network (VANET) has become an attraction for both academia and industry because of its application in Intelligent Transportation Systems (ITS) and Vehicular social networks (VSN). VANET facilitate safety as well as non-safety services and hence improve the ride quality. VSN relies on VANET and Mobile Social Network (MSN) where connected vehicle exchange numerous messages. These messages need to be delivered within the time constraint. The idea behind VSN is to develop a robust recommendation system and route planner based on social behaviour of vehicle and its mobility pattern. But vehicles in VANET are prone to security breach. Several researchers have addressed the issues of secured message dissemination using public key cryptography but the schemes lack in message verification time because of high dependence on the trusted authority (TA). Thus, in this research work, road side units (RSU) which are deployed across the city are considered as a Proxy signer by the Road side control centre (RSCC). RSU acts in place of the RSCC as a proxy signer. In order to check the malicious RSU, RSCC uses revocation in consultation with the TA. The proposed scheme is secure because of its reliance on computationally hard Diffie-Hellman problem (DHP). The security analysis shows the effectiveness of the proposed scheme in terms of lowering the overhead on the OBU in verification of signature.

**Keywords** Proxy signature · RSU · RSCC · VSN · VANET

F. Azam (✉) · S. Kumar
Department of Computer Science and Engineering, Sangam University, Bhilwara, Rajasthan, India
e-mail: farooque53786@gmail.com

S. Kumar
e-mail: sunil.kumar@sangamuniversity.ac.in

N. Priyadarshi
CTiF Global Capsule, Department of Business Development and Technology, Aarhus University, 7400 Herning, Denmark
e-mail: neerajrjd@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_6

## 1 Introduction

VANET being a self-organizing network provides vehicle to vehicle (V2V) and vehicle to infrastructure communication (V2I) communication by employing Dedicated Short Range Communication (DSRC) and Wireless Access Vehicular Environment (WAVE). Vehicular networking requires secure vehicle operation, traffic assistance, road safety, fuel station such as electric vehicle (EV) charging station as well as driver safety apart from some infotainment applications [1, 2]. Numerous VANET research and standardization projects are being carried out in several countries across the globe such as PReVENT to bring road safety and preventive safety related application by European automotive industry project co-funded by the European Communication Commission (ECC) in Europe, Vehicle Safety Communications Consortium (VSCC) in USA intelligent transportation system (ITS) Consortium and Advanced Safety Vehicle project in Japan. V2V deployment was demonstrated by General Motors employing Cadillac vehicles in 2006. Automakers such as BMW, Toyota, Daimler, Audi, Volvo, Honda, and Car-to-Car Consortium are also working on V2V communication. The automobile industries, Government organizations, and academia are partnering together for applications viz. vehicle collision warning system, traffic information disseminations, secured EV charging schemes etc. [3–5].

The above mentioned applications of VANET pose challenges in terms of security, privacy, and authenticity of the communicating parties. The communication must guarantee node authentication, message integrity and privacy of the entity. In this regards, numerous research work has been carried out to mitigate and provide secured communication. A vehicle sends safety message every 100–300 ms as per the DSRC protocol [6]. In a PKI based system, a message is signed before sending and the verification for these messages were mitigated in previous research work. But, when a heavily dense traffic where 60–200 vehicles share safety message is considered, the generation and verification of 600–2000 signed messages poses challenges in terms of communication and computational overhead. Also, PKI based system needs the public key and certificate to be appended to the message which may increase the message size and incur message delay. Due to stringent time constraint, even a delay of millisecond may result into catastrophic situation. Hence, there is a need of an algorithm which robust, incur less communication and computational overhead and at the same time proves scalable. Scalability can be defined as an ability to sustain an increase in number of vehicles without compromising the loss in traffic data and performance in the system.

In order to provide integrity and non-repudiation, various digital signature schemes have been used in the recent years viz. identity based signature, proxy signature, ring signature, blind signature etc. [7–10] to mitigate the issues.

In this research work, a modified proxy signature scheme with warrant has been proposed for fast verification of messages to reduce the computational overhead of the OBU and hence increases the throughput.

The paper is summarized as: Sect. 2 discusses the related work. Section 3 give an insight on component of VANET architecture and the system requirement. Section 4, explains the proxy signature scheme by first explaining the basics and then proposes a warrant based proxy scheme. In Sect. 5, security analysis based on the proposed algorithm has been discussed. Finally, Sect. 6 concludes the paper with some future work.

## 2 Related Work

In [7], Mambo et al. discussed various proxy signature schemes, in which the original signer delegates a proxy signer to sign on behalf of an original signer. In this paper, the proxy delegation has been categorized as (i) full delegation, (ii) partial delegation and (iii) delegation with warrant. Full delegation comprises of same secret keys for both the original as well as proxy signer. In partial delegation secret key of both the original signer and proxy signer are different. Proxy signer secret are created from original signer secret key. Delegation by warrant emphasizes that the proxy signer to be entrusted completely. The idea relies on Diffie-Hellman computationally hard problem.

Schnorr et al. [8], have proposed public key approach for signature based on discrete logarithmic problem of sub-group. It shows improvement over Elgamal scheme.

Hua et al. [9], have proposed a hybrid proxy scheme to improve the efficiency and provide less computational and communication overhead as compared to the existing scheme. In this paper, authors have used the concept of identity based PKI based system along with the agent concept. The time to verify a message is $2\,T_{mul} + 2\,T_{mtp} + 6\,T_{pa}$, and overhead byte to verify a message is 126 bytes.

Chaum et al. [11] and Boneh et al. [12], have presented a group signature approach in which a member of the group can sign on behalf of the group head. In this approach, the identity of the signing agent is not known inside the group. Group head also called group manager can check any signed message by the group member through its master key. In order to reveal the identity in case of malicious behavior of a group member, it incurs communication and computational overhead because of the peer discovery procedure.

In [13], authors have discussed an identity based ring signature approach to provide anonymity based on bilinear groups.

Cui et al. [14] have proposed a binary search and cuckoo filter approach to improve the efficiency in batch verification of messages. But in this approach the vehicle signs every message which incurs computational and communication overhead.

In [15], Zhang et al. have discussed a certificate-less proxy ring approach avoiding the use of bilinear pairings and exponentiation operation. The difficulty with the approach is its inefficiency to distribute private keys in the communication environment. Also, membership changed is not allowed as it is predefined. Thus, the scheme fails in VANET requirement.

Zhang et al. [16], have proposed a proxy ring signature approach with t times revocable mechanism. In this approach, the author has shown the effectiveness using the proposed algorithm and claimed completeness, unforgeability, etc. Proxy ring based signature is also discussed in [17], and authors have used additive operation apart from others.

## 3 Preliminaries and System Design

### *3.1 System Model*

Figure 1, shows the typical component of a VANET consisting of Onborad Unit (OBU), Road side control Center (RSCC), Road Side Unit (RSU) and a Trusted Authority (TA).

**OBU.** An OBU is embedded in each vehicle and uses 802.11p protocol to communicate with other vehicle's OBU and nearby RSU. An OBU is considered as a semi-trusted entity in VANET which regularly broadcast it's driving status. OBU has limited computational and storage capability as compared to the RSU.

**RSU.** These are usually deployed alongside the roads and various intersections for V2I communication. Usually, RSU acts as a semi trusted authority and acts a bridge between the TA and vehicle. RSU regular sends a beacon messages regarding road safety and local environment such as a gas station, a parking lot and an infotainment application. As RSU have more computation and storage capability as well as wide communication range, it can acts as a proxy for signing and verifying the critical message to incur better throughput and less overhead.

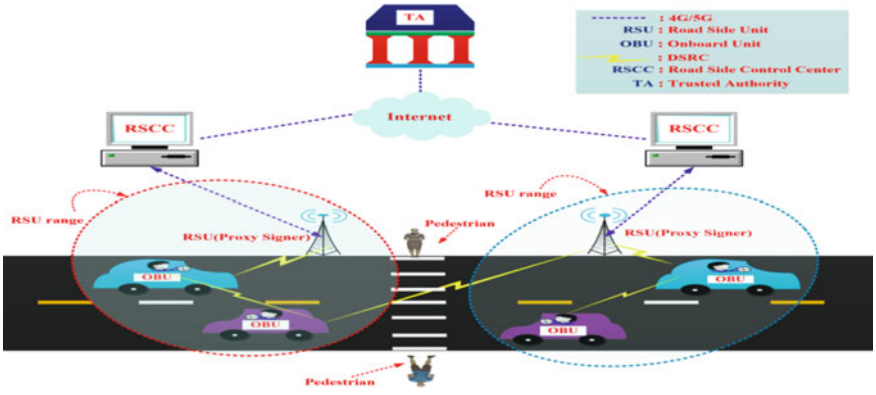



**Fig. 1** System model for a VANET

**TA.** Trusted authority is the most trusted, uncompromised component in VANET. TA holds the responsibility to register, assign public/private key pair and authenticate the vehicle before inclusion in VANET. In order to keep communication safe various efforts have been made by the researchers in the recent past. Usually, a TA sign and verify the signature for secured communication of a vehicle. TA initializes parameters for both the OBU and RSU and can revoke the malicious vehicle as well as RSU. TA can choose the RSU as a proxy signer to offload the burden and facilitate fast exchange of messages.

**RSCC.** Road side control center are spread across the city along with the RSUs. Group of RSUs are associated with the RSCC. RSCC keep tracks of generating the proxy and revoke when found malicious.

### *3.2 Assumptions*

As shown in Fig. 1, RSUs are spread across geographical area at various intersections along the road of a city and group of RSUs are controlled by a Road Side Control Center (RSCC). All RSCC are independent to each other and are connected through the internet and the Trusted Authority (TA). TA maintains information regarding each RSU, RSCC and vehicles. TA issues certificate to each RSU. The public key of the TA is known to all in the VANET. RSU advertises a beacon message as Eq. 1:

$$beacon := \begin{pmatrix} (ID_{RSCC}, Pub\_key_{RSCC}, ID_{RSU}, MAC\_Addr_{RSU} \\ , Loc_{RSU}), H(ID_{RSCC}, ID_{RSU}, Loc_{RSU})_{SK_{TA}} \end{pmatrix} \quad (1)$$

where,

$ID_{RSCC}$ = Identity of Road side control center
$ID_{RSU}$ = Identity of Road side unit
$Loc_{RSU}$ = Location/MAC address of RSU
$MAC\_Addr_{RSU}$ = MAC address of RSU
$Pub\_key_{RSCC}$ = Public key of Road side control center
$H(ID_{RSCC}, ID_{RSU}, Loc_{RSU})_{SK_{TA}}$ = Signature using TA secret key
H(.) denotes a one-way hash function.

Vehicle embedded with OBU come to know about the parameters such as RSCC public key, RSU MAC address, RSU location. The integrity of the beacon is recognized by the signature of TA on the message and hence the RSU is authentic and belongs to a particular RSCC. A vehicle then decides to a particular RSU group or waits for another beacon for the joining process.

## 4 RSU Based Proxy Signature with Warrant

Let $p$ and $q$ are two prime numbers generated in such a way that $q$ is prime factor of $p-1$. Values of $p$ and $q$ are designated to the large geographical space. Then, $g$ as a generator for $Z_P^*$ for covering small area is selected.

### *4.1 Proxy Signing Initial Phase*

RSCC selects a random number $d < q$ which acts a private key for the RSCC for secured communication in VANET. Then, the public key is computed by Eq. (2):

$$Pub\_key_{RSCC} = g^d \bmod p \tag{2}$$

This private/public key is always calculated before the actual start of communication process. The items such as $p, q, d, Pub\_key_{RSCC}$ are made public for all RSUs and OBUs in the communication range. Apart from this, RSCC generates a random number $e \in_R Z_{p-1} \setminus \{0\}$; such that $Z_{p-1} \setminus \{0\}$ represents a finite field. Here, revocation parameter for the scheme is $e$. At the initialization phase, RSCC computes the following Eq. 3:

$$K = g^e \bmod p \tag{3}$$

The value of $K$ is needed to calculate the proxy secret key, $\sigma$ by Eq. (4):

$$\sigma = d + eK \bmod (p-1) \tag{4}$$

From Eq. (5), value of $\sigma$ serves as secret identity of the Road side unit, and is stored in the memory of each RSU which is volatile. Also, the value of $K$ is used to verify the proxy pair $(\sigma, K)$ as well as in verification of delivered message by the RSU and OBU respectively.

An RSU verifies if the following congruence hold on the receipt of a proxy pair $(\sigma, K)$ as Eq. 5:

$$g^{\sigma} \equiv Pub\_key_{RSCC}\, K^K \bmod p \tag{5}$$

If Eq. (6) doesn't holds, the proxy pair $(\sigma, K)$ is discarded and the demand of new proxy is placed before the RSCC. Usually, the value of $\sigma$ and $K$ are valid for long times until unauthorized third party misuse it.

RSU uses $\sigma$ in place of $d$ as a secret key for the signature scheme employed in VANET. Here, Schnorr's scheme is used because of its rapid verification and minimum bandwidth requirement [7, 8]. The modified Signature and verification scheme are detailed below.

### 4.2 RSU Proxy Signature

RSCC chooses random number as a session parameter $m < q$ to compute using Eq. 6:

$$x = g^m \bmod p \tag{6}$$

RSCC generates this session parameter whenever a message is there to be broadcasted. To ensure the freshness of the message, validity time period (VTP) and the sequence number ($Seq\_no$) is appended with the message (M). RSCC calculates the following Eq. 7:

$$h = H(M, x, VTP, Seq\_no) \tag{7}$$

RSCC also calculates the legacy warrant to have a better control over the proxy signer as follows Eq. 8.

$$h_s = (h, VTP), H\,(h, VTP)_s \tag{8}$$

The legacy warrant consists of signature of hash value along with validity time period as shown in Eq. (9). RSU receives the tuple (M, VTP, m) consisting of message (M), validity time period of the message (VTP) and the session parameter (m) through an IPSec tunneling. The process is repeated whenever a message has to be broadcasted. The hash value h, validity information are extracted from legacy warrant $h_s$ and verified using RSCC public key $Pub\_key_{RSCC}$. Next, RSU uses session parameter $m$ and $h$ to compute $y$ as Eq. 9:

$$y = (m + \sigma h) \bmod q \tag{9}$$

Now, the proxy signature $(h_s, y)$ is appended to the message/application resulting in the tuple as $(M, h_s, y, K)$. This tuple is broadcasted by the Road side unit.

### 4.3 Proxy Verification Strategy

The recipient vehicle's Onboad unit utilizes the signature ingredient to verify the proxy signature employed on the safety message $M$ as follows Eq. 10:

$$Pub\_key_{RSCC}{}^{/} = Pub\_key_{RSCC}\, K^K \bmod p \tag{10}$$

The value in Eq. (10) is used by Eq. (11):

$$x^{/} = g^y\, Pub\_key_{RSCC}{}^{/h'} \bmod p \tag{11}$$

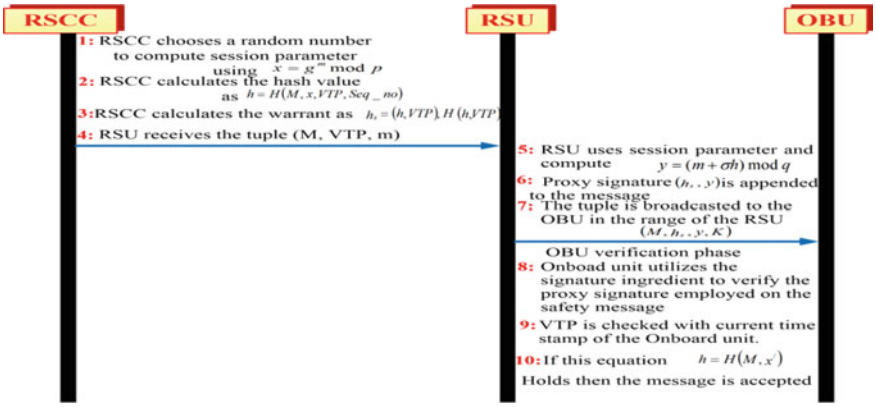



**Fig. 2** Overview of the proxy signature scheme

Upon receiving value of h and VTP from the legacy warrant, h is verified using the public key of RSCC ($Pub\_key_{RSCC}$). VTP is checked with current time stamp of the Onboard unit. If it is outdated, then a resend alert is initiated by the OBU from the RSU for the message. If not, the message is validated through the following Eq. 12.

$$h = H\left(M, x^{/}\right) \tag{12}$$

If Eq. 12 holds then the messages is authenticated and considered valid else OBU either discards the message or generates an alert message to the RSU that the message is corrupted. Figure 2 explains the overview of proposed schemes in terms of generation and verification.

## 5 Security Analysis

The strength of the proposed warrant based proxy signature approach relies on the complexity of solving the DHP (Diffie-Hellman Problem). The proxy signer derives the secret key from the original signer's secret key. An adversary can't generate the secret even by having the unauthorized manipulation of the proxy key because of the intractability of Diffie-Hellman Problem. Below subsection describes the analysis process.

- **False message transmission**
  The original signer RSCC delegates the RSU to sign on behalf of it. In this approach, the session parameter based on the message M and validity time period is utilized to avoid a replay attack. The secret key of the RSCC is not known to anyone in the communication range of the VANET, so only RSCC can generate the session parameter in such a way that vehicle's Onboard unit

can accept the message. Even though an adversary can succeed in verifying a false message from the OBU, but the probability is very less. Since, $p$ and $q$ are large primes with minimum length as 512 bits. The probability calculated for $false/Modified\_message$ to be verified by an Onboard unit is $< {}^{2}/_{(2^{511}-1)} \approx {}^{1}/_{2^{510}}$. Thus, the scheme is robust enough towards thwarting false message attack.

- **Unforgeability**
  RSCC delegates an RSU as a proxy signer by using the proxy pair $(\sigma, K)$ to sign a message. Even though an adversary a valid combination of proxy pair say $(\sigma', K')$, it will not be able to do so because the h value can only be generated by the RSCC. Thus, an adversary will not be successful in creating a proxy. Also, $\sigma$ being computed from a randomly generated secret, computing new $\sigma$ is computationally infeasible because of the hardness of solving a discrete logarithmic problem.
- **Non-repudiation**
  Each RSU is considered as only one proxy and is in full control of the RSCC. Any adversary would not be able to generate the proxy signature even getting access to the public parameter $d$ and $e$ as the $y$ component is unique to the RSU and can be only calculated using Eq. (10).
- **Impersonation Attack**
  Even though an adversary a valid combination of proxy pair say $(\sigma', K')$, it will not be able to calculate the $y$ component and hence can't be able to launch impersonation attack.
- **Revocation**
  An adversary may succeed in compromising the RSU to send a forged message to the OBU. If gone undetected, catastrophe may happen in VANET communication. Hence the RSCC must revoke such compromised RSU. For this, the RSCC regenerate $e$ and then compute the verification parameter. TA will verify through RSCC and revoke such RSU.

## *5.1 Performance Analysis*

In [17] and [18], only time consuming operations such as addition/multiplication have been taken into consideration.

Let $T_{par}$ denotes pairing operation time, $T_{Ga}$ denotes addition operation on the elliptical curve and $T_{Gm}$ denotes the scale multiplication [17].

In the proposed scheme, the calculation of hash value is considered as most time consuming in case of verification of the signature while $T_{\text{mod}}$ is the time for modulus calculation is considered in the generation phase along with the $T_h$(Time to calculate hash value).

The EAAP scheme proposed by Azees et al. [18] $2\,T_p + 4\,T_{ep-1} + T_{ep-1}$ for verifying one certificate and signature, where $T_p$ is the time required for pairing operation while $T_{ep-1}$ and $T_{ep-2}$ are the time required for the exponentiation operation of group $G_1$ and $G_2$. Table 1, shows the signature generation and verification for [17, 18] and proposed scheme.

**Table 1** Performance analysis of different schemes

| Scheme | Signature generation | Signature verification |
|---|---|---|
| [17] | $2(n-1)\,T_{Ga}$ $+2n\,T_{Gm}$ | $2n\,T_{Ga}+n\,T_{Gm}$ |
| [18] | $2\,T_p+4$ $T_{ep-1}+T_{ep-1}$ | $(1+n)\,T_p+$ $4\,T_{ep-1}+n\,T_{ep-1}$ |
| Proposed | $n\,T_h+n\,T_{\text{mod}}$ | $n\,T_h$ |

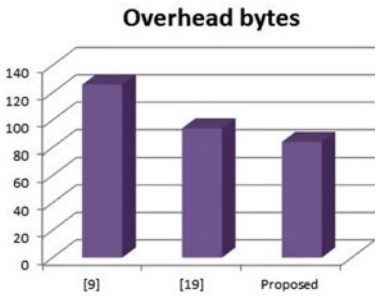



**Fig. 3** Overhead verification bytes for OBU

The proposed scheme has verification time less as compared to schemes mentioned in [17] and [18]. Thus, the proposed scheme reduces the computational burden on the OBU.

## *5.2 Computational Overhead Computation*

In the proposed signature scheme, the broadcasted message by the RSU is the tuple $(M, h_s, y, K)$ Where, M is the message which is of fixed length. The hash calculation is going take 20 bytes, the prime number $p$ and $q$ is of 512 and 140 bits as per the requirement of proxy signature and for every message, the calculation of $x^{/} = g^{y}\,Pub\_key_{RSCC}{}^{/h'} \bmod p$ is done every time. Hence, $x^{/}$ contributes an overhead of 64 bytes. Finally, hash is calculated in the verification process and contributes 20 bytes. Thus, total overhead bytes in the proposed scheme are 84 bytes with standard SHA-I use. In [9], authors have proposed a hybrid approach and overhead bytes for verification of a single message is 126 bytes.

In [19], the transmitting bytes are 161. Excluding the payload, the overhead for a signed message is 94 bytes which is more compared to proposed scheme. Thus, OBU have will be less burdened with verification of messages. Figure 3 shows the overhead bytes plot.

## 6 Conclusion

In this research work, an RSU based proxy signing scheme using warrant to comply with VANET security requirement is proposed. RSU acts as a proxy signer as long as it is not compromised. The parameter $e$, also called revocation parameter to revoke the malicious RSU is used. The session parameter generated by RSCC is unique for each RSU and an adversary will never know the validity time period of a message. Thus, the proposed system is robust enough to resist impersonation attack. In terms of computational burden, the scheme scale well as compared to schemes mentioned and reduces the burden on the OBU.

As a part of future work, a bilinear pairing approach with legacy warrant can be employed and simulation would be added to check the effectiveness of the scheme. Also, vehicle to infrastructure and vehicle to RSU can be worked out for the proposed scheme.

## References

1. Azam F, Yadav S, Priyadarshi N, Padmanaban S, Bansal RC (2021) A comprehensive review of authentication schemes in vehicular Ad-Hoc network. IEEE Access 9:31309–31321. https://doi.org/10.1109/ACCESS.2021.3060046
2. Azam F, Priyadarshi N, Nagar H, Kumar S, Bhoi AK (2021) An overview of solar-powered electric vehicle charging in vehicular adhoc network. In: Electric vehicles. Green energy and technology. Springer, Singapore, pp 95–102 (2021)
3. Azam F, Kumar S, Yadav S, Priyadarshi N, Padmanaban S (2020) An outline of the security challenges in VANET. In: Proceedings of 2020 IEEE 7th Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON), pp 1–6
4. Ma X, Zhang J, Yin X, Trivedi KS (2012) Design and analysis of a robust broadcast scheme for VANET safety-related services. IEEE Trans Veh Technol 61(1):46–61
5. Xu Q, Mak T, Ko J, Sengupta R (2004) Vehicle-to-vehicle safety messaging in DSRC. In: Proceedings of ACM international workshop on vehicular ad hoc network, pp 19–28
6. Raya M, Hubaux JP (2007) Securing vehicular ad hoc networks. J Comput Secur 15(1):39–68
7. Mambo M, Usuda K, Okamoto E (1996) Proxy signatures for delegating signing operation. In: CCS 1996: Proceedings of the 3rd ACM conference on computer and communications security. ACM, New York, pp 48–57
8. Schnorr C-P (1990) Efficient identification and signatures for smart cards. In: CRYPTO 1989: Proceedings of the 9th annual international cryptology conference on advances in cryptology. Springer, London, pp 239–252
9. Liu H, Wang H, Gu H (2020) HPBS: a hybrid proxy based authentication scheme in VANETs. IEEE Access 8:161655–161667
10. Park J-H, Kim Y-S, Chang JH (2007) A proxy blind signature scheme with proxy revocation. In: Computational intelligence and security workshops, international conference on, vol 0, pp 761–764
11. Chaum D, van Heyst E (1991) Group signatures. In: EUROCRYPT, pp 257–265
12. Boneh D, Boyen X, Shacham H (2004) Short group signatures in advances in cryptology. In: CRYPTO 2004, Lecture notes in computer science, vol 3152. Springer, Heidelberg, pp 41–55
13. Wei-min L, Zong-kai Y, Wen-qing C (2005) A new id-based proxy blind signature scheme in Wuhan University. J Nat Sci 10(3):555–558

14. Cui J, Zhang J, Zhong H et al (2017) SPACF: a secure privacy-preserving authentication scheme for VANET with cuckoo filter. IEEE Trans Veh Technol 66(11):10283–10295
15. Zhang C, Su B, Yao S (2013) Certificate-less proxy ring signature scheme without bilinear pairings'. Appl Res Comput 30(12):3797–3799
16. Zhang J, Chen H (2011) An efficient t-times proxy ring signature scheme with revocable anonymity. In: Proceedings of IEEE international conference on communication software and networks, Xi'an, China, pp 6–10
17. Lianhai L, Yujue WJ, Zhang QY (2019) Efficient proxy ring signature for VANET. IET Energy Internet Relat Technol Res 2019:5449–5454
18. Azees M, Vijayakumar P, Deboarh LJ (2017) EAAP: efficient anonymous authentication with conditional privacy-preserving scheme for vehicular Ad Hoc networks. IEEE Trans Intell Transp Syst 18(9):2467–2476
19. Tangade S, Manvi SS, Lorenz P (2018) Decentralized and scalable privacy-preserving authentication scheme in VANETs. IEEE Trans Veh Technol 67(9):8647–8655

# A Secure Data Transmission Using AODV and Hash Function for MANET


**Arudra Annepu and Madalai Jayaprasad**

**Abstract** A Mobile ad hoc network (MANET) is generally a wireless and autonomous network formed using mobile nodes without any infrastructure. Additionally, this MANET is susceptible to security issues due to its dynamic network topology and open medium. In this paper, the hash function is proposed with an ad hoc on-demand distance vector routing protocol for mitigating the blackhole attacks which used to reduce the packet drop. The angle of the arrival-based localization method is used to identify the location of the nodes. The performances analyzed in this proposed method are Average Packet Delivery Rate (APDR), Average Packet Dropped (APD), average throughput, Average Routing Overhead (ARO), and Average Normalized Routing Load (ANRL). The performance of the proposed method is compared with the dynamic threshold-based protocol and trust based probabilistic broadcast method. The APDR of the proposed method is 98% for 10 nodes, which is high when compared to the existing protocols.

**Keywords** Ad Hoc on demand distance vector routing protocol · Angle of arrival based localization method · Blackhole attacks · Hash function · Mobile ad hoc network

## 1 Introduction

Mobile ad hoc network (MANET) is generated by the collection of wireless mobile nodes that are generally organized without using any framework [1]. The nodes in the MANET are operated in three different modes that are communication, computation, and sensing mode [2]. The nodes in the network require intermediate nodes for communication with other nodes that are positioned beyond the transmission range

A. Annepu (✉)
Rajiv Gandhi Institute of Technology, Bangalore 400053, Karnataka, India
e-mail: rgit.csehod@gmail.com

M. Jayaprasad
CMR University School of Engineering, Bangalore 560037, Karnataka, India
e-mail: jayaprasad.m@cmr.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_7

[3, 4]. The MANET is used in various fields such as disaster recovery systems, military applications, forestry, vehicle networks, etc. [5]. The major advantages of the MANET are the design of faster and inexpensive network construction which are obtained by avoiding the definite infrastructure [6]. The data transmission among the nodes is accomplished by using the routing algorithm over the network layer [7].

In general, the MANET is considered homogeneous, where each sensor contributes an equivalent radio capacity. But, the poor scalability affects this homogeneous ad hoc network [8]. The inherent characteristics of the network such as dynamic topology and openness of wireless medium are the reasons that cause MANET to be susceptible to security threats [9, 10]. Without the support of proper infrastructure, the secure data transmission in the MANET gets affected [11]. The threats and malicious nodes present in the network cause the packet drop during the data transmission [12]. To overcome the security and energy-related threats, a routing protocol should be developed to handle the dynamic nature of the network topologies [13]. Therefore, the secure routing accomplishes integrity, authenticity, confidentiality and availability during the communication [14, 15]. The major contributions of the paper are given as follows:

The location of the nodes present in the MANET is identified by using the Angle Of Arrival (AOA) based location method.

The Ad hoc On-Demand Distance Vector (AODV) routing protocol is used for transmitting the data packets from the source to the destination.

The Prevention of Blackhole Attack using Hash Function (PBAHF) is used to secure the data transmission from blackhole attacks.

The overall organization of the paper is given as follows: the literature survey about different AODV routing protocols used to mitigate the blackhole attacks is given in Sect. 2. The proposed AOA localization method and AODV with hash function are described in Sect. 3. Then the results and discussion of the AODV-PBAHF method are described in Sect. 4. Finally, the conclusion is made in Sect. 5.

## 2 Literature Survey

Gurung and Chauhan [16] developed the threshold-based Mitigating Flooding Attack Mechanism (MFAM). The MFAM algorithm utilizes the special nodes known as Flooding Intrusion Detection System (F-IDS) which was used to identify and avoid the flooding attack. The packet delivery ratio of the F-IDS was less than the conventional AODV algorithm. Keerthika and Malarvizhi [17] developed AODV routing with hybrid Weighted Trust-Based Artificial Bee Colony (WTABC) and 2-opt for MANET. The WTABC's global search ability was integrated with the discrete local searching capability of the 2-opt algorithm. This WTABC with 2-opt AODV was minimized the possibility of black hole attacks in MANET.

El-Semary and Diab [18] presented the Blackhole Protected AODV (BP-AODV) protocol based on the chaotic map to detect and protect against the blackhole attack. The developed BP-AODV demonstrated irregular behavior during the forwarding

process. Xu et al. [19] developed the Trust based Probabilistic Broadcast method (TPB) for securing the data transmission using the trustworthiness of the node. The light-weight trust management technique was used to calculate the trust by using the direct and recommended trust. This TPB was considered only trust values during routing, but the trust value was affected because of the lack in position update. Gurung and Chauhan [20] presented the dynamic threshold-based protocol named MBDP-AODV which was used to avoid the black-hole attack. The MBDP-AODV protocol was achieved higher PDR than the conventional AODV. But, the MBDP-AODV protocol has higher routing overhead due to the data transmission of multiple relay packets.

## 3 AODV-PBAHF Method

In the AODV-PBAHF method, the blackhole attack present in the MANET is mitigated by using the hash function-based AODV routing protocol. Since the AOA-based localization is used for identifying the location of the nodes. The overall architecture of the AODV-PBAHF method is shown in Fig. 1.

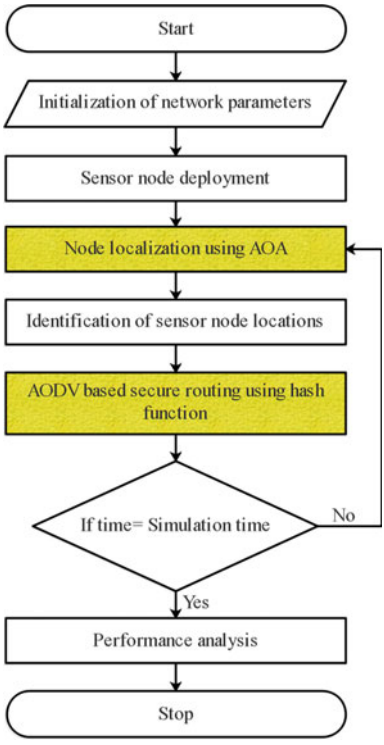



**Fig. 1** The architecture of the AODV-PBAHF method

### 3.1 Node Localization Using an Angle of Arrival

The position of the nodes is identified by using the AOA-based localization method. This AOA localization has two main steps: 1) Node distribution in the interested area, and 2) Calculation of anchor node location. The sample illustration of the AOA-based location is shown in Fig. 2. In Fig. 2, the coordinates of the anchor nodes A and B are $(x_a, y_a)$ and $(x_b, y_b)$ and the location of unknown sensor U is $(x_u, y_u)$. Consider, $x_a \neq x_b$ and $y_a \neq y_b$, here the anchor nodes A and B transmit the signal to the unknown node present in the network.

The position of the unknown sensor node is formulated based on the tangent function shown in Eq. (1).

$$y_a - y_u = (x_a - x_u)\tan\theta_1; \quad y_b - y_u = (x_b - x_u)\tan\theta_2 \tag{1}$$

where, $\alpha_1$ and $\alpha_2$ are the angles from the unknown node to anchor nodes A and B; $\theta_1$ and $\theta_2$ are the angles for anchor node A to an unknown node and anchor node B to an unknown node respectively. Equation (2) defines the solution for Eq. (1).

$$x_u = \frac{y_b - y_a + x_a \tan\theta_1 - x_b \tan\theta_2}{\tan\theta_1 - \tan\theta_2}; \quad y_u = y_b - \frac{(x_b - x_a)\tan\theta_1 - (y_b - y_a)}{\tan\theta_1 - \tan\theta_2}\tan\theta_2 \tag{2}$$

Figure 2 shows that the two pairs of AOA which are $\alpha_1$ and $\alpha_2$ and $\theta_1$ and $\theta_2$ have the following relation as shown in Eq. (3).

$$\theta_1 = \alpha_1 - 180^\circ; \quad \theta_2 = \alpha_2 - 180 \tag{3}$$

Equation (1) is rewritten as shown in Eq. (4), based on the periodicity of the tangent function.

$$x_u = \frac{y_b - y_a + x_a \tan\alpha_1 - x_b \tan\alpha_2}{\tan\alpha_1 - \tan\alpha_2}; \quad y_u = y_b - \frac{(x_b - x_a)\tan\alpha_1 - (y_b - y_a)}{\tan\alpha_1 - \tan\alpha_2}\tan\alpha_2 \tag{4}$$

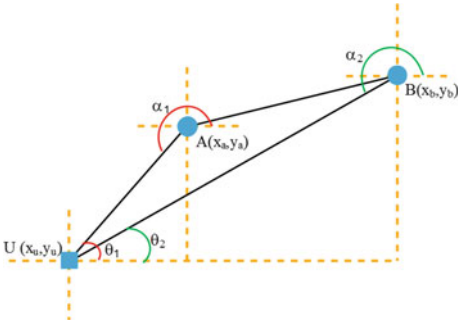



**Fig. 2** AOA based localization with two known anchor nodes

In the location calculation, $y_a = y_b$ when the anchor nodes are deployed in the same line. Equation (5) which used identify the position of the unknown nodes

$$x_u = \frac{x_a \tan\alpha_1 - x_b \tan\alpha_2}{\tan\alpha_1 - \tan\alpha_2}; \quad y_u = y_b - \frac{(x_b - x_a)\tan\alpha_1 \tan\alpha_2}{\tan\alpha_1 - \tan\alpha_2} \tag{5}$$

Then the identified sensor node positions are given to the AODV algorithm for generating the routing path between the nodes.

### *3.2 Prevention of Black Hole Attack Using a Hash Function in AODV Protocol*

In this PBAHF method, the AODV routing protocol is combined with the hash function for avoiding black hole attacks which is described in the following section.

**Hash Function Based RREP Message.** In the AODV-PBAHF method, the one-way hash function is frequently applied to random numbers for generating the hash chain. The Time To Live (TTL) value is set to be identical to the highest value of the hop count field which is shown in Eq. (6).

$$Max\_Hop\_Count = TTL \tag{6}$$

where the $Max\_Hop\_Count$ represents the highest value of the hop count field. Equation (7) shows that the initial value of the hash is equal to $N_s$.

$$Hash = h(N_s) \tag{7}$$

where the random number generated by the sequence number is denoted as $N_s$ and the hash function is denoted as $h$. The hash function is applied for TTL times to compute the maximum hash field which is shown in Eq. (8).

$$Max\_hash = h^{TTL}(N_s) \tag{8}$$

The following condition i.e. Eq. (9) is verified when the node receives the RREQ/RREP from the adjacent node.

$$M_{hash} = h^{TTL-Hop\_count}(Hash) \tag{9}$$

where the $M_{hash}$ represents the condition for verifying the node, $Hop\_count$ specifies the number of hops for the respective node. The computed $M_{hash}$ is integrated into the RREQ structure which is shown in Eq. (10).

$$RREQ \rightarrow \{S, D, ID, Srcnum, Desnum, Hop\_count, M_{hash}\} \quad (10)$$

where, the source node address is denoted as $S$; destination node address is denoted as $D$; RREQ's identification number is $ID$; the source and destination sequence number are represented as the $Srcnum$ and $Desnum$.

**Route Discovery Process.** The route discovery is performed by using 3 message schemes which are: route request (RREQ), route reply (RREP), and routing error (RERR). Here, the route is identified by transmitting the RREQ message throughout the network. The adjacent node which receives the RREQ message sends the RREP message when the respective node doesn't diminish by the black hole attack. The packets are transmitted to the destination node when the source node receives the RREP packet. Additionally, the RRER is transmitted to all nodes in the network, when the link failure is detected in the MANET.

## 4 Results and Discussion

The implementation and simulation of the AODV-PBAHF method are carried out in the Network Simulator (NS)-2.35 with i5 desktop and 8 GB RAM. The important parameters used in the AODV-PBAHF method are given in Table 1. The performance results are compared with the TPB [19] and MBDP-AODV [20]. This TPB [19] and MBDP-AODV [20] are designed and simulated with the same specifications mentioned in the Table 1.

**Table 1** Simulation parameters

| Parameter | Value |
|---|---|
| Area | $750 \times 750\,m^2$ |
| Number of nodes | 10–60 |
| Protocol | AODV with a hash function |
| Maximum mobility speed | 20 m/s |
| Mobility model | The random waypoint mobility model |
| Propagation radio model | Two ray ground |
| MAC layer | IEEE 802.11 |
| Traffic type | CBR-UDP |
| Packet size | 512 bytes |
| Simulation time | 500 |

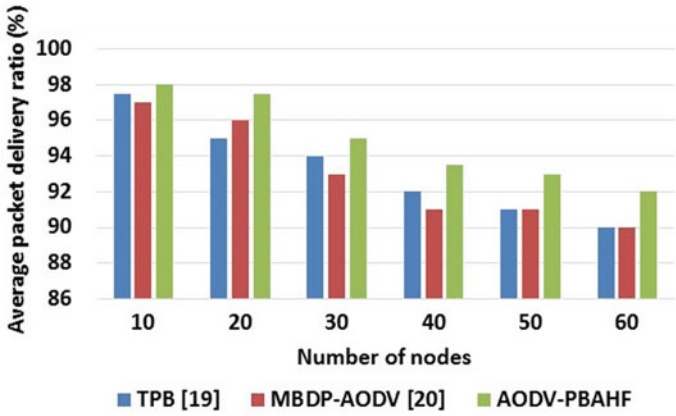



**Fig. 3** Comparison of APDR for AODV-PBAHF with TPB and MBDP-AODV

## *4.1 Average Packet Delivery Rate*

APDR is defined as the ratio of the total data collected ($Q_i$) at the destination to the total data generated ($P_i$) in the source node. The APDR is shown in Eq. (11).

$$APDR = \frac{1}{M} \frac{\sum_{i=1}^{n} Q_i}{\sum_{i=1}^{n} P_i} \times 100\% \tag{11}$$

where $M$ is the number of experiments and $i$ is the amount of source node.

From the APDR comparison of Fig. 3, it is observed that the packet delivery ratio of the AODV-PBAHF method is higher when compared to the TPB [19] and MBDP-AODV method [20].

## *4.2 Average Packet Dropped Rate*

APD is the ratio between the dropped packets and the generated packets in the transmitter which is shown in the following Eq. (12).

$$APD = \frac{1}{M}\left(\sum_{i=1}^{n} P_i - \sum_{i=1}^{n} Q_i\right) \tag{12}$$

From the APD comparison of Fig. 4, it can be concluded that the packet loss of the AODV-PBAHF method is less during the data transmission. The integration of AODV with hash function mainly helps to detect the black hole attack and node failure in the network.

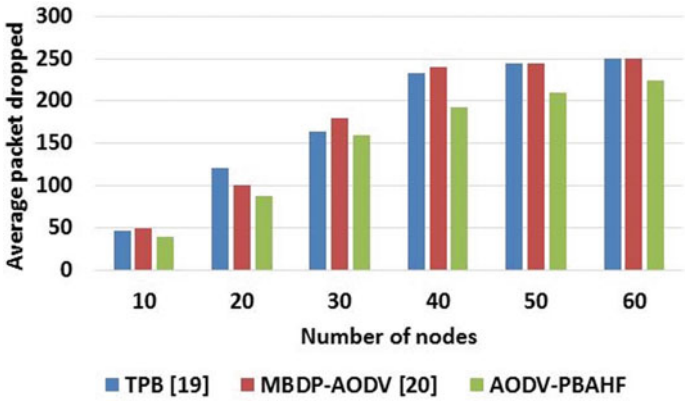



**Fig. 4** Comparison of APD for AODV-PBAHF method with TPB and MBDP-AODV

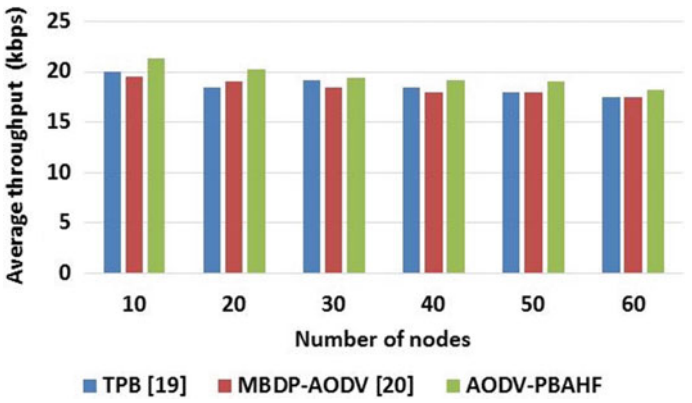



**Fig. 5** Comparison of throughput for AODV-PBAHF with TPB and MBDP-AODV

## *4.3 Average Throughput*

Average throughput is the ratio between the total size of packets ($l$) received by the destination and the difference of start ($t_1$) and stop simulation time ($t_2$) which is expressed in Eq. (13).

$$Average\ throughput = \frac{1}{M} \times \frac{\sum_{i=1}^{n} Q_i \times l}{t_2 - t_1} \tag{13}$$

From the throughput comparison of Fig. 5, it can be concluded that the AODV-PBAHF method achieves higher throughput because the destination receives a high amount of data packets during transmission of the data packets.

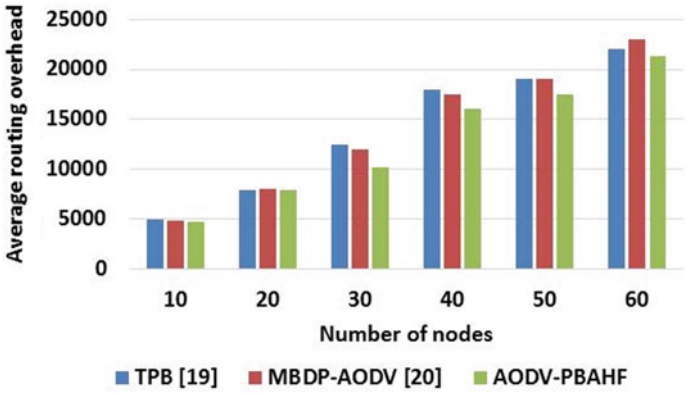



**Fig. 6** Comparison of ARO for AODV-PBAHF with TPB and MBDP-AODV

### *4.4 Average Routing Overhead*

The amount of control packets ($C_i$) created in the MANET is defined as ARO which is expressed in Eq. (14).

$$ARO = \frac{1}{M} \times \sum_{i=1}^{n} C_i \tag{14}$$

From the ARO comparison of Fig. 6, it is concluded that the routing overhead of the AODV-PBAHF method is less during the data transmission. The less routing overhead is achieved by minimizing the amount of RREP messages.

### *4.5 Average Normalized Routing Load*

ANRL is defined as the ratio between the total amount of control packets and packets collected at the destination node which is expressed in Eq. (15).

$$ANRL = \frac{1}{M} \times \frac{\sum_{i=1}^{n} C_i}{\sum_{i=1}^{n} Q_i} \tag{15}$$

From ANRL comparison of Fig. 7, it is observed that the AODV-PBAHF method obtains lesser ANRL by mitigating the node failure and blackhole attack during the data transmission.

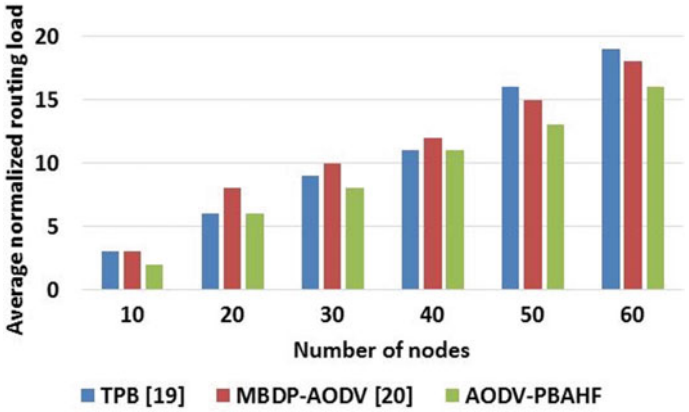



**Fig. 7** Comparison of ANRL for AODV-PBAHF with TPB and MBDP-AODV

## 5 Conclusion

In this research paper, the AODV with hash function routing protocol is used for mitigating the black hole attack in the network. The AOA localization technique is used for identifying the exact location of the nodes. Here, the hash function value is added in the RREP message mechanism of the AODV protocol for achieving reliable data transmission through the network. The mitigation of blackhole attack and node failure in the routing path improves the packets received by the destination node. The throughput of the AODV-PBAHF method is 18.25 kbps for 60 nodes, which is higher when compared to the TPB and MBDP-AODV methods. In the future, a novel optimization algorithm can be used to improve the energy efficiency of the MANET.

## References

1. Kanagasundaram H, Kathirvel A (2018) EIMO-ESOLSR: energy efficient and security-based model for OLSR routing protocol in mobile ad-hoc network. IET Commun 13(5):553–559
2. Merlin RT, Ravi R (2019) Novel trust based energy aware routing mechanism for mitigation of black hole attacks in MANET. Wirel Pers Commun 104(4):1599–1636
3. Babu RL, Balasubramanie P (2020) Fuzzy rule selection using hybrid artificial bee colony with 2-opt algorithm for MANET. Mob Netw Appl 25:585–595
4. Guaya-Delgado L, Pallarès-Segarra E, Mezher AM, Forné J (2019) A novel dynamic reputation-based source routing protocol for mobile ad hoc networks. EURASIP J Wirel Commun Netw 2019(1):1–16
5. Chintalapalli RM, Ananthula VR (2018) M-LionWhale: multi-objective optimisation model for secure routing in mobile ad-hoc network. IET Commun 12(12):1406–1415
6. Satheeshkumar S, Sengottaiyan N (2019) Defending against jellyfish attacks using cluster based routing protocol for secured data transmission in MANET. Clust Comput 22(5):10849–10860
7. Bhuvaneswari R, Ramachandran R (2019) Denial of service attack solution in OLSR based manet by varying number of fictitious nodes. Clust Comput 22(5):12689–12699

8. Sathiamoorthy J, Ramakrishnan B (2017) Design of a proficient hybrid protocol for efficient route discovery and secure data transmission in CEAACK MANETs. J Inf Secur Appl 36:43–58
9. Rajashanthi M, Valarmathi K (2020) A secure trusted multipath routing and optimal fuzzy logic for enhancing QoS in MANETs. Wireless Pers Commun 112:75–90
10. Abirami KR, Sumithra MG (2018) Preventing the impact of selfish behavior under MANET using Neighbor Credit Value based AODV routing algorithm. Sādhanā 43(4):60
11. Ahmed MN, Abdullah AH, Chizari H, Kaiwartya O (2017) F3TM: flooding factor based trust management framework for secure data transmission in MANETs. J King Saud Univ-Comput Inf Sci 29(3):269–280
12. Vanitha K, Rahaman AZ (2019) Preventing malicious packet dropping nodes in MANET using IFHM based SAODV routing protocol. Clust Comput 22(6):13453–13461
13. Veeraiah N, Krishna BT (2020) An approach for optimal-secure multi-path routing and intrusion detection in MANET. Evol Intell 1–15
14. Mukhedkar MM, Kolekar U (2020) E-TDGO: an encrypted trust-based dolphin glowworm optimization for secure routing in mobile ad hoc network. Int J Commun Syst 33(7):e4252
15. Sethuraman P, Kannan N (2017) Refined trust energy-ad hoc on demand distance vector (ReTE-AODV) routing algorithm for secured routing in MANET. Wireless Netw 23(7):2227–2237
16. Gurung S, Chauhan S (2018) A novel approach for mitigating route request flooding attack in MANET. Wireless Netw 24(8):2899–2914
17. Keerthika V, Malarvizhi N (2019) Mitigate black hole attack using hybrid bee optimized weighted trust with 2-Opt AODV in MANET. Wireless Pers Commun 106(2):621–632
18. El-Semary AM, Diab H (2019) BP-AODV: blackhole protected AODV routing protocol for MANETs based on chaotic map. IEEE Access 7:95185–95199
19. Xu H, Si H, Zhang H, Zhang L, Leng Y, Wang J, Li D (2020) Trust-based probabilistic broadcast scheme for mobile ad hoc networks. IEEE Access 8:21380–21392
20. Gurung S, Chauhan S (2019) A dynamic threshold based algorithm for improving security and performance of AODV under black-hole attack in MANET. Wireless Netw 25(4):1685–1695

# A Study on Different Types of Convolutions in Deep Learning in the Area of Lane Detection


**T. S. Rajalakshmi and R. Senthilnathan**

**Abstract** One of the key technologies in autonomous vehicles is image based lane detection algorithm. High performance is detected in modern deep learning methods. But in case of challenging areas like congested roads or poor lighting conditions, it is difficult to accurately detect lanes. Global context information is required which can be extracted from limited visual-cue. Moreover, for automotive driver assisting system, like lane keep, collision avoid etc., it is important to know the position of the vehicle i.e., in which lane it is. Due to large varieties in shape and colour of the lane marking, it becomes difficult to solve this task. For this purpose, an initial step on the input image is the image processing, where the data is processed as per the requirement in pixel level semantic segmentation. Then comes in the creation of the semantic segmentation model which is able to process the data. This model can be of different variant based on the computation ability, as well as the parameter handling capacity.

**Keywords** Convolutional neural network · Deep learning techniques · Evaluation metric · Encoder-decoder · Fully convoluted network

## 1 Introduction

Vehicular automation is the result of technology development. Here, the automation levels increases resulting in the reduced driver intervention. Computers can be made to gain high level understanding from digital images, with the aid of computer vision. It tries to minimize the visual system of human with the help of computer and camera. Deep learning models are inspired from the way in which information is processed in biological nervous system. The scene understanding keeps an important place in

T. S. Rajalakshmi (✉) · R. Senthilnathan
Department of Mechatronics Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamilnadu, India
e-mail: rajalakt1@srmist.edu.in

R. Senthilnathan
e-mail: senthilr4@srmist.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_8

computer vision, because of it real-time perception, as well as its ability to analyses and elaborate an explanation of dynamic scene which leads to more new discoveries. It is affected by the cognitive vision along with the involvement of both computer vision and cognitive engineering. Explicit detection of road features such as lane marking will lead to an improved localization of the activity during detection. With the coming of Convolutional neural network and Alex net [1] winning the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), it can be seen that the deep learning has become the important tool in the image/video processing or computer vision applications. Lane detection is tackled by a range of deep learning methods. The range runs from early CNN-based method to segmentation method.

## 2 Related Works

An application of computer vision is the vision-based lane detection task. It can be classified as classification of image, detection of object or semantic segmentation. In 2012, AlexNet having 5 convolution layers and 3 fully connected layers and an extension of the LeNet [2] won ILSVRC. Many techniques like the ReLU [3] and Dropout [4] are applied in the Alex net. The architecture of AlexNet is simple yet marks a significant success of the network. It is seen that by increasing the depth or width, the CNN's solution space can be amplified [5]. Similar to AlexNet, GoogLeNet [6], VGG [7] with wider and deeper architecture got higher accuracy in ILSVRC2014. In VGG, the depth was increased by 16–19 layers, while in GoogLeNet both depth (22 layers) as well as width was increased. GoogLeNet otherwise known as Inception-v1, through different optimization, it has developed from Inception-v1 to Inception-v4 [8]. The generalization performance of VGG is comparatively better, and hence, it is more often used in order to perform extraction of image features. Deeper the CNN, the issue of exploding gradient or vanishing gradient is a possibility. The detectors based on the already available deep learning models can be grouped into two divisions as two-stage and one-stage methods. Of this, the two-stage method includes the RCNN [9], Fast RCNN [10], Faster RCNN [11], and Light Head RCNN [12]. In the two-stage method, the candidate regions are first generated either by CNN or by traditional methods and then classified into a group. The one-stage method includes YOLO [13], SSD [14, 15] etc. here, without the region proposal stage, the position coordinate and the category probability are directly generated. Semantic segmentation aims to classify every pixel in the image into a category. In 2015, Fully convolutional network (FCN) was proposed by Long et al. [16]. Following the Fully Convolutional Network, encoder-to-decoder architecture as shown in Fig. 3 is used to attend to the issues of image segmentation. GCN [17] adopted large convolution kernels in order to fuse different contextual information, while PSPNet [18] introduced a pyramid pooling module. Deeplab [19, 20] fuses dilated spatial pyramid pooling.

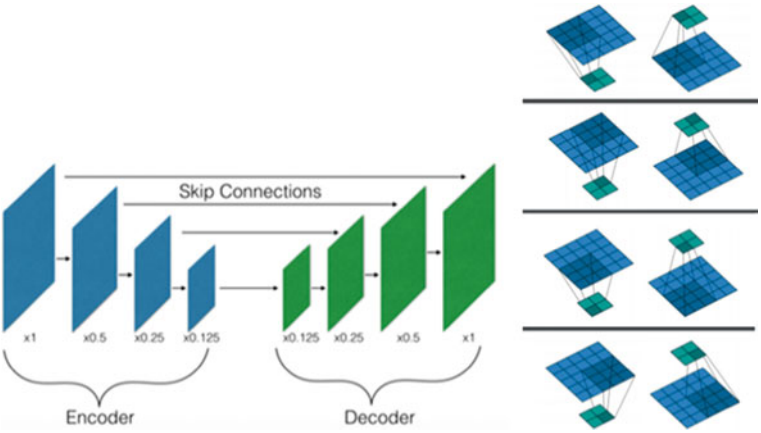



**Fig. 1** Encoder decoder structure with skip connections and transposed convolution

## 3 Methods

### *3.1 Encoder Decoder Architecture*

In the encoder-decoder architecture, the convolution with the pooling block consist of the encoder; while the transposed convolution constitutes of the decoder. The skip connection between the encoder and the decoder is for the purpose of preserving the construct. The encoder generates a high dimensionality feature vector from the input image. The features are aggregated at multiple levels and in that processes, the input image is compressed. The decoder takes the high dimensionality feature vector and generates semantic masks. It up samples the features that are aggregated by encoders at multiple levels. The additional skip connection sends the feature maps directly from an earlier layer of encoder to a later layer of decoder. This aids in the formation of distinctly defined decompressions of input data. Transposed convolution is an up-sampled convolution, which up samples the input feature map. The encoder decoder structure with skip connections and transposed convolution is indicated in Fig. 1.

### *3.2 Dilated Convolutions*

Dilated convolution is also termed as atrous convolution or a hole convolution. These are the type of convolution wherein the kernels are inflated by insertion of holes between the kernel elements. This is an alternate to the down sampling layer. By inflating the kernels, the receptive field is increased while the spatial dimension of feature maps is maintained. The kernel with different dilation rates is represented in Fig. 2.

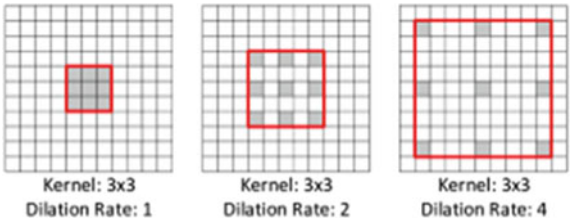



**Fig. 2** Kernel with different dilation rates

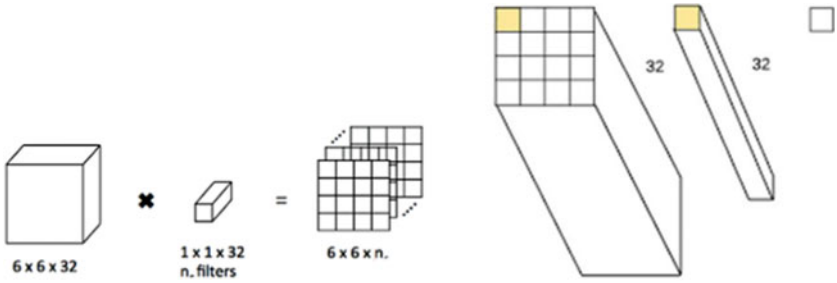



**Fig. 3** Depiction of $1 \times 1$ convolution

## *3.3 1 × 1 Convolution*

As the depth of the network in deep convolutional network increases, the number of feature maps also often increases. Thus, a large filter size results in an increase in the number of parameters as well as the computation. To overcome this issue, a $1 \times 1$ convolutional layer can be applied which offers feature map pooling. This technique results in dimensionality reduction while retaining their important features, by decreasing the number of feature maps. The $1 \times 1$ convolution is denoted in Fig. 3.

## *3.4 Residual Layer*

By going in for skip connections, convolutions learn the residual functions which facilitate training. By the use of 1D factorized convolution, there is a significant reduction in the computational cost, at the same time, retaining accuracy. The residual block allows the convolutional layers to learn residual functions. This results in the reduction of degradation problem which is otherwise present in architectures that stacks more layers, and the residual layer is denoted in Fig. 4.

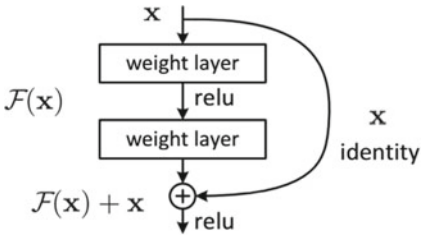



**Fig. 4** Residual layer

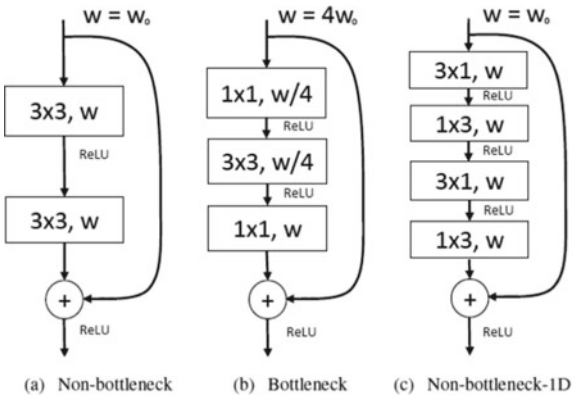



**Fig. 5** Non-bottleneck—1D

## *3.5 Non-bottleneck-1D*

The non-bottleneck in Fig. 5a as well as the bottleneck in Fig. 5b have similar number of parameters as well as accuracy. But the computational resources required by bottleneck is less, while suffering from degradation problem. A non-bottleneck 1D is applied for the purpose of reducing the convolutions on the original residual methods. Large filters are seen to benefit by this decomposition. Also, there is a 33% reduction in the parameters, further resulting in an increased computational efficiency.

## *3.6 Generator Discriminator*

The model consists of generator network which transforms the random input into a data instance. The output of the generator is connected straight to the input of the discriminator. The generator learns to generate likely data, which becomes negative training examples for the discriminator. The discriminator then learns to differentiate the data which is fake from the generator data which is real. The generator is penalized by the discriminator for not producing likely data, as denoted in Fig. 6.

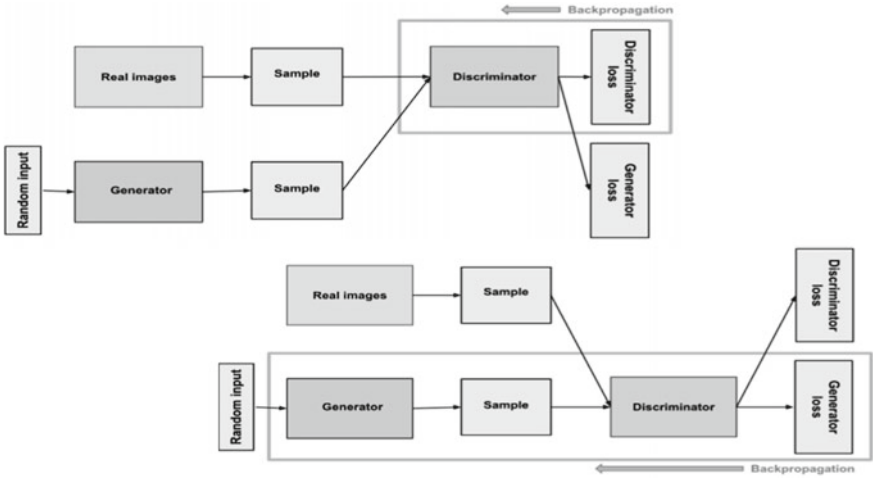



**Fig. 6** Generative adversarial network

# 4 Network Models

## *4.1 Encoder Decoder Models*

Kim [21] proposes a sequential end-to-end transfer learning method without any post processing, to determine both the ego lanes—left and right, directly as well as separately. Here, point-detection is redefined as a region-segmentation problem; as a result of which, the proposed method is not sensitive to occlusion and variations of environmental conditions. The method comprises of two transfer learning steps. Initially, the network's representation domain is changed from a general scene to that of a road scene; then, the target is reduced from road objects in general, to left and right ego lanes in particular. Here, the method proposed determines ego lanes with low sensitivity to the road conditions. Neven [22] casts the lane detection problem as an instance segmentation problem. Here, each lane configures its own instance which can then be trained end-to-end. Initially, a learned perspective transformation is proposed. By doing so, a robust lane fitting is ensured. This fast lane detection algorithm has no fixed number of lanes for detection purpose. Bruls [23] in his paper presents a weakly-supervised learning system for real-time road marking detection. This is done by using images which are incurred from a monocular camera. The expensive manual labelling can be avoided by exploiting additional modes of sensor in order to generate large quantities of images that are annotated in a weakly-supervised manner. Later, these images are used for training a deep semantic segmentation network. The model does not depend on any pre-processing steps. Here, the road markings are detected in real time under various lighting, weather as well as traffic situations.

### *4.2 Fully Convoluted Network Models*

Huval [24] shows how the existing convolutional neural networks (CNNs) can be used for performing lane and vehicle detection at the same time, running at frame rates which are required for a real-time system. For this, sliding window with overfeat features using CN was employed for the purpose of parsing the driving scene. The overfeat feature includes Integrated Recognition, Localization and Detection. He [25] discusses a Dual-View Convolutional Neutral Network (DVCNN) framework for lane detection. For this, to improve the low precision ratio, a novel strategy is designed, wherein, simultaneous optimization of front-view and top-view images are done. In the front-view image, the detections which are false are removed, whereas in the top-view image structures which are non-club-shaped are eliminated. Next, a weighted hat-like filter is used. This is capable of recalling lane lines and reducing false detections. Then, a global optimization function is designed. Here, the lane line probabilities, lengths, widths, orientations are taken into circumstances. Gurghian [26] presents an approach for estimating lane positions directly using a deep neural network. Here, the images are incurred from a laterally-mounted down-facing cameras. In order to create a diverse training set, semi artificial images are generated. Also, for validating the robustness of driver-assist features, the network provides an efficient way. Lee [27] in his paper defines a unified end-to-end trainable multi-task network. This is capable of handling lane as well as road marking detection and recognition. This is guided by a vanishing point under adverse weather conditions. To address this poor weather conditions issue, a lane and road marking benchmark is built. It includes 20,000 images with 17 lane and road marking classes under various scenarios such as, rain, no rain, night and heavy rain. The approach VPGNet is capable of detecting and classifying lanes, road markings, and also predicting a vanishing point with a single forward pass. The VPGNet is able to perform several tasks such as grid regression, detection of object, classification with multi-label, and prediction of vanishing point. Pan [28] proposes Spatial CNN, also known as SCNN. This generalizes the traditional convolutions to slice-by-slice convolutions within the feature maps [30]. This enables passing of message between pixels across rows as well as columns in a layer [31]. The results show that in semantic segmentation, the diffusion effect is found to be beneficial for larger objects. But SCNN is capable of preserving the continuity of long thin structures.

## 5 Discussion

In the previous section different deep learning methods were reviewed. In this section, the numeric discussion of those methods is to be reviewed.

## 5.1 Evaluation Metrics

**Pixel Accuracy.** Pixel accuracy is the percentage of pixels in the image that are classified correctly. The issue of high pixel accuracy, called as the class imbalance does not imply the ability of superior segmentation. Here, a class or some classes dominate the image, while the rest of the other classes make up only a small portion of the image. The pixel accuracy is mathematically defined in Eq. (1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

**True Positive.** *False Negative, False Positive:* In semantic segmentation, when a prediction-target mask pair has an IoU score which exceeds predefined threshold, a true positive is identified. When a predicted object mask has no ground truth object mask associated with it, a false positive is predicted. Similarly, when a ground truth object mask has no predicted object mask associated with it, a false negative is seen.

**Intersection-Over-Union (Jaccard Index).** IoU computes the ratio between intersection of ground truth and predicted, which is the number of true positives and union of ground truth and predicted segmentation, which is the sum of false negatives, true positives and false positives. The IoU is computed for on class by class basis and is then the average is computed. The Jaccard coefficient is defined in Eq. (2).

$$Jaccard\ coefficient = \frac{TP}{TP + FP + FN} \tag{2}$$

**Dice Coefficient.** It is the harmonic mean of the precision and recall. The F1 score is used to measure the classification model. F1 score is biased to the lowest value of both precision and recall. Thus, an increase in F1 score will result in a balanced precision and recall and it is defined in Eq. (3)

$$Dice\ coefficient = \frac{2TP}{2TP + FP + FN} \tag{3}$$

## 5.2 Results

Nguyen et al. [29] developed a hybrid deep learning network based on Gaussian process for segmenting the scene image into background and lane regions. The simulation result showed that the developed model attained 97% of classification accuracy. Further, Muthalagu et al. [30] introduced a new lane detection approach based on histogram analysis and perspective transformations. In this literature study, the developed approach effectively detects both curved and straight lane lines. Hence,

**Table 1** Comparative analysis between the proposed and existing models

| Models | Accuracy (%) |
|---|---|
| Hybrid deep learning network based on Gaussian process [29] | 97 |
| Histogram analysis and perspective transformations [30] | 95.75 |
| **Proposed** | **98.31** |

the developed approach obtained 95.75% of accuracy in lane detection. Compared to these models, the proposed model obtained 98.31% of accuracy in lane detection as in the below Table 1.

## 6 Conclusion

It can be seen that various methods provided results either on non-standard datasets or that they are not real time tested. This makes the comparison slightly difficult. Moreover, information is seen to be lacking on few other metrics such as the time of execution, memory footprint etc. it would be better if there was a standard dataset, open source code for implementation and a report based on the metrics related to semantic segmentation.

## References

1. Krizhevsky IS, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
2. LeCun Y, Bottou L, Bengio Y, Haffner P et al (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
3. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. In: Proceedings of the 27th international conference on machine learning (ICML 2010), vol 560, pp 807–814
4. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov R (2012) Improving neural networks by preventing coadaptation of feature detectors arXiv:1207.0580. 31
5. Montufar GF, Pascanu R, Cho K, Bengio Y (2014) On the number of linear regions of deep neural networks. In: Advances in neural information processing systems, pp 2924–2932
6. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
7. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. Computer Science
8. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-first AAAI conference on artificial intelligence
9. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object 575 detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587

10. Girshick R (2015) Fast R-CNN. In: Proceedings of the IEEE international conference on computer vision, pp 1440–1448
11. Ren S, He K, Girshick R, Sun J (2015) Faster RCNN: towards real-time object detection with region proposal networks. In: Advances in neural information processing systems, pp 91–99
12. Li Z, Peng C, Yu G, Zhang X, Deng Y, Sun J (2017) Light-head R-CNN: in defense of two-stage object detector arXiv:1711.07264
13. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, vol 32, pp 779–788
14. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC (2016) SSD: single shot multibox detector. In: European conference on computer vision. Springer, pp 21–37
15. Wu X, Sahoo D, Hoi SC (2020) Recent advances in deep learning for object detection. Neurocomputing
16. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3431–3440
17. Peng C, Zhang X, Yu G, Luo G, Sun J (2017) Largekernelmatters-improvesemantic segmentation by global convolutional network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4353– 4361
18. Zhao H, Shi J, Qi X, Wang X, Jia J (2017) Pyramid scene parsing network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2881–2890
19. Chen L-C, Papandreou G, Schroff F, Adam H (2017) Rethinking atrous convolution for semantic image segmentation, arXiv preprint arXiv:1706.05587
20. Hu Z, Tang J, Wang Z, Zhang K, Zhang L, Sun Q (2018) Deep learning for image-based cancer detection and diagnosis a survey. Pattern Recogn 83:134–149
21. Kim J, Park C (2017) End-to-end ego lane estimation based on sequential transfer learning for self-driving cars. In: CVPR. IEEE
22. Neven D, De Brabandere B, Georgoulis S, Proesmans M, Van Gool L (2018) Towards end-to-end lane detection: an instance segmentation approach. In: 2018 IEEE intelligent vehicles symposium (IV). IEEE, pp 286–291
23. Bruls T, Maddern W, Morye AA, Newman P (2018) Mark yourself: road marking segmentation via weakly-supervised annotations from multimodal Data. In: 2018 IEEE international conference on robotics and automation (ICRA), Brisbane, Australia, 21–25 May 2018
24. Huval B, Wang T, Tandon S, Kiske J, Song W, Pazhayampallil J, Andriluka M, Rajpurkar P, Migimatsu T, Cheng-Yue R, Mujica F, Coates A, Ng AY (2015) An empirical evaluation of deep learning on highway driving. arXiv:1504.01716v3 [cs.RO], 17 April 2015
25. He B, Ai R, Yan Y, Lang X (2016) Accurate and robust lane detection based on dual-view convolutional neutral network. In: 2016 IEEE intelligent vehicles symposium (IV), Gothenburg, Sweden, 19–22 June 2016
26. Gurghian A, Koduri T, Bailur SV, Carey KJ, Murali VN (2016) DeepLanes: end-to-end lane position estimation using deep neural networks. In: CVPR
27. Lee S, Kim J, Yoon JS, Shin S, Bailo O, Kim N, Lee T-H, Hong HS, Han S-H, So Kweon I (2017) VPGNet: vanishing point guided network for lane and road marking detection and recognition. arXiv:1710.06288v1 [cs.CV], 17 October 2017
28. Pan X, Shi J, Luo P, Wang X, Tang X (2017) Spatial as deep: spatial CNN for traffic scene understanding. arXiv:1712.06080v1 [cs.CV], 17 December 2017
29. Nguyen TNA, Phung SL, Bouzerdoum A (2020) Hybrid deep learning-gaussian process network for pedestrian lane detection in unstructured scenes. IEEE Trans Neural Netw Learn Syst 31(12):5324–5338
30. Muthalagu R, Bolimera A, Kalaichelvi V (2020) Lane detection technique based on perspective transformation and histogram analysis for self-driving cars. Comput Electr Eng 85:106653

# A Study on the Impact of DC Appliances and Direct DC Power System in India


**D. Silas Stephen, T. Muthamizhan, and Jinu Sophia J**

**Abstract** Due to the increase in the load the conservation of energy has become an important aspect in the world. The evolvement of Direct Current (DC) sources like Photo Voltaic (PV) and Battery storage has made an increasing fraction of residential end-use loads to operate on DC. The DC sources have made it possible to connect the DC loads without conversion from AC to DC thus minimizing the conversion loss too. In this paper the compact ability of the DC loads in residential loads, its efficiency and cost effectiveness is discussed. But the major drawback is the in-availability of DC ready appliances in the market. Therefore, the available DC ready appliances in the Indian market are also discussed. The direct DC power system provides energy and cost saving in the residential loads which makes the homes as net zero energy homes. Therefore, the residential load with DC appliances, which generates power through a PV system is considered and a comprehensive study is carried out considering the cost and energy saving. Also the energy efficiency of DC ready appliances considering the power supply losses and saving in cost is analyzed. Based on the analysis it is recommended to step towards the development of DC ready appliances and to make it available in the market which enables the conservation of energy and saving in cost.

**Keywords** Battery systems · Cost assessment · Direct current appliances · Distribution system · Market assessment · Photo voltaic system

D. Silas Stephen (✉)
Panimlar Engineering College, Chennai, India
e-mail: silasstephen@gmail.com

T. Muthamizhan
Sri Sairam Institute of Technology, Chennai, India

J. Sophia J
Ràjalakshmi Engineering College, Chennai, India
e-mail: jinusophia.j@rajalakshmi.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_9

## 1 Introduction

This has occurred due to the dramatically change of the end users since the days when the benefits of AC and DC were debated during the war of the currents. In India, the power is generated, transmitted and distributed as AC as it is the dominant power transmission and distribution technology. But the today's appliances and distributed electricity sources actually favor DC, due to intrusion of photo voltaic (PV), light emitting diode (LED) lighting, and consumer electronics, which are manufactured in local markets. Due to the development of energy efficient DC based motors like permanent magnet DC motors (BLDC) and energy storage devices the scope towards the DC devices and applications has increased abruptly According to Starke et al. [1], power supplied by the DC sources have to be converted into AC by the conversion devices and transmitted through the AC System which adds losses and the further the loss is increases due to the reconversion of AC to DC, so the possible application of a DC distribution system is the DC-DC converter. Also several studies have found that a "Direct-DC" building distribution system with onsite PV and DC appliances could save energy avoids power conversion losses due to the DC to AC converters and back to DC. The saving is about 2–3% to as much as 14% while considering the power consumption of the entire building [2, 3]. Though, the Direct DC distribution system, the losses are reduced and cost is saved, but there are various barriers to the developments. The major barrier is the lack of devices which could readily operate on DC. But in recent days the experts and industrialist have understood the need of DC appliances for the development of DC system in the residential systems.

## 2 Related Works

Vossos et al. [4] study was carried out regarding the market for DC-ready appliances in residential system in India. The market, efficiency, and cost assessment of the current DC-ready appliance market for the residential end-use applications are carried out and were presented. Ginart and Sharifipour [5] analyzed DC fast charging system and DC residential system for the integration of battery storage and solar panels. Siraj and Khan [6] analyzed the impact of several voltages on the residential systems incorporating power electronic losses and distribution losses. In addition to this, author analyzed the efficiency of the system for typical DC home at 380, 220 and 48 V DC and compared the results with 220 V AC using simulation and analytical systems. Sabry et al. [7] investigated the problems of existing works and explained the electrical diagrams by classifying the consumed energy. Sabry and Ker [8] analyzed the power consumption of inverter-driven variable speed controller refrigerator, which was one of the important household loads. Nandini et al. [9] provided an overview of DC Micro-grid with DC distribution system for DC loads.

## 3 End Uses and DC Appliances in India

The usage of electricity in India is drastically increasing and the total power consumed by 2021 is about 379,854 GW/year and it may grow to 482,622 Gw/year in 2026. The electricity demand is 540.9 Millions of tons of oil equivalent (Mtoe) in financial year 2016–20 and the electricity according to various sector. So according to the data residential electrical load contributes about 9.7% in India. The appliance used may be of: lightning, entertainment, kitchen appliances and heating and cooling. The entertainment appliances include radio, CD players, TV, DVDs and Computers, whereas the kitchen appliances are refrigerator, washing machine, Electric cooker, Electric oven and microwave and the heating and cooling appliances are electric water heater, fans, Air coolers and air conditioning. The residential load comprises of 25% cooling and heating appliance, 9% lighting loads, 9% water heating, 9% refrigerators and 8% electronic equipment's. Table 1 shows the average operating per unit power consumption of various appliances.

### *3.1 DC-Ready Product Information by End-Use Application*

Various DC appliances are available in the market which are manufactured by various private manufacturers like SELCO, Simpa, Basil Energetics, Cygni energy and D light. They offer various range of DC appliances such as DC fans, DC mixer grinders, DC Cookers, Refrigerators and coolers. These appliances are used for residential

**Table 1** Average operating per unit power consumption of various applications

| Appliance | Unit | Per unit power consumption |
|---|---|---|
| Lighting | w/hr | 38 |
| Radio | w/hr | 11 |
| CD player | w/hr | 35 |
| TV | w/hr | 135 |
| DVD/ VCR | w/hr | 30 |
| Computer | w/hr | 70 |
| Washing machine | Wh/load | 149 |
| Refrigerator | Kwh/year | 255 |
| Electric oven | w/hr | 1248 |
| Electric water heater | Kwhr/unit | 566 |
| Fans | w/hr | 37 |
| Aircooler | w/hr | 230 |
| Air conditioning | w/hr | 1.811 |
| Water pumps | w/hr | 250–1500 |

**Table 2** DC ready appliances available in market in India

| Appliance | Brand | Wattage | Voltage | Capacity |
|---|---|---|---|---|
| Refrigerator | Sinfin, Prozone [5] | 25 W | 12 V DC | 220 L |
| Air conditioner | Sinfin, Prozone | 800 W | 240 V DC | 1.0 tons |
| Water heater | CSR project | 430 W | 24 V | 10 L |
| Fans | Orient, Atomberg, Havells, Crompton, Superfan [6] | 24 w | 12/24 V | 1200 mm |
| Water pump | ERH India | 1 kw | 12 V | 0.5 Hp |
| | Mach power point pumps India, Sinfin, Prozone | 50–1000 W | 12 V | 0.1–1.0 Hp |
| Lightings (Lamps) | Multiple brands | 3–10 W | 12 V | |
| Lightings (Tube lights) | Multiple brands | 12 W | 12 V/24 V | |
| Mixer | Preethi | 600 W | 12 V DC | |

houses and for productive usage. Many of the DC appliances are in the introduction stage but they are super-efficient and consume less power when compared to AC appliances. The various DC ready products available in market in India by are presented in Table 2.

## *3.2 Market Assessment of DC Appliances*

DC made appliances is available from the early decades in the form of radios, telecommunication and railways. Now days the DC products are available in data centers, lightings and data transfer. The DC grid connected application has also come into existence due to the development of power electronic based devices. In the residential loads the DC appliances are developed such that they could be operated by directly connecting to the solar panel. Various appliances which could be connected are lightings, chandeliers, water pumps, air conditioners, water heaters, cookers, mixers etc.

## *3.3 End-Use*

In heating and cooling system many DC based appliances have come into existence. Refrigerators at a rating of 25 W which could operate at 12 V DC is available. It can be connected directly to the 100 W solar panel and 30 Ah battery. Similarly, air conditioners are available at the rating of 1.0 to 1.5 ton which could be operated in 240 V with a rating of 800 W. This air conditioner has MG compressor, and can be connected directly to PV panel with 2500 Ah batter. It consumes about 5–10 units if it runs for 10 h. In India 24 V DC water heater has been developed under the

CSR project with a capacity of 10 L. DC fans are equipped with highly efficient BLDC motors, and are typically marketed for off-grid applications. The DC fans Consumes less power, have intelligent motion control and Variable flow. They are available in various brand names designed for 12 or 24 V with a rating of 24 W. Now a day, DC LED lamps are available in most of the brand with various rating in local markets at reasonable cost. DC Water pumps are commercialized now days at a capacity of 0.5 to 1.5 Hp which are designed such that they can be operated by directly connected to solar panel with a battery system. They are operated at 12 or 24 V with wattage of 50 to 1000 W. Even mixers, grinders which can operate in DC are available n market. While considering the electronic equipment's, though all products are DC-internal, they are supplied by rectifying the AC voltage into DC. Some of the exempted devices which could operate directly are AC-DC power adapters, such as cell phones (typically powered by USB), laptop computers, a small number of computer monitors, and other small electronics. So at current scenario there is a possibility of usage of all appliances at a home in DC.

## 4 Efficiency Assessment of DC Appliances

According to the literatures DC appliances and DC internal devices are more efficient than the AC appliances [10]. The efficiency of the DC appliances is enabled by various technologies. In air conditioners the efficiency and energy saving is improved by replacing the single speed compressor which run by AC by variable speed compressor and fans which run by brushless DC motor. The air conditioner efficiencies are improved with the variable speed compressor. In case of lightning system, the energy is saved with the LED lamps and tube lights. In refrigerators VSD compressor run by BLDC motor is uses and the energy saving is about 53% [11]. The LED lamps consume much lesser energy by 75% than the incandescent bulbs. For example, a LED flood lamp use only 11 W while creating light output comparable to 50 W' incandescent lamp. The commercially available LED lamps have efficiencies of 200 lumens per watt. Unlike AC fans, DC fans are more efficient. DC fans consume less than 70% power than the AC ceiling fans. For example, if an AC fan consumes 100 W, the DC fan may consume only 30 W for the same output. The major disadvantage of the AC water pumps is the running time of the pump, whereas the operation time is not limited in DC. As the water pumps uses a brushless, electronically communicated motor the efficiency is very high at most speed ranges. The efficiency of the PWM brushless DC motor when compared with the 240 AC motor: 1 kW 240 V capacitor start capacitor run motor has an average efficiency of 69%, when three phase the efficiency is 74%, but the efficiency of 7 kW brushless DC motor is about 93% [12].

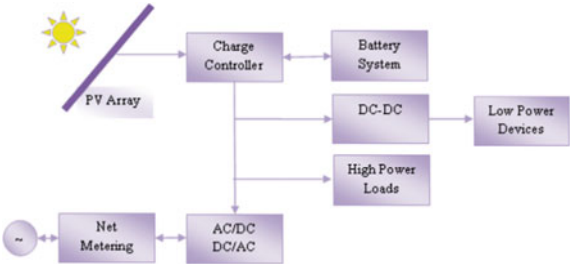



**Fig. 1** Taxonomy of a direct DC system connected to the grid with DC ready appliances

## *4.1 Results of Redirect-DC Building Distribution Systems*

Direct DC building is possible due to the various DC market ready appliances and the PV based system. It consists of the PV system, AC to DC converter (if connected to the grid) and the DC appliances. The elements of the system are connected by DC cabling. The direct DC building distribution system is more efficient and energy saving. Figure 1 shows the schematic diagram of a direct DC system connected to the grid with DC ready appliances. In direct DC system the energy saving is very higher when the load is fed by the PV and battery system. The saving can be maximized by operating the loads according to the DC generation. It is based on the relevant AC and DC supplies which can be enhanced by using by using direct DC supply.

The saving in energy is high when the loads are operated according to the DC generation. The residential house which utilizes AC in India commonly comprises of 3 Fluorescent lamps of 40 W, 1 Fluorescent lamp of 20 W, incandescent lamp of 40 W, 2 ceiling fans of 75 W, desktop computer or television with 200 W, Juice mixer grinder of 800 W, Refrigerator of 200 L, washing machine of 1000 W and water heater of 1000 W. Therefore, the total load on an average in a residential house is about 5 kW. But the average running load will be 20%. Although the actual daily load curve for an average day have greater variability than smooth average load shapes. Proper Load shifting techniques will be helpful in correlating the PV output with the load curve by which energy saving can be further enhanced.

## *4.2 Cost Assessment of DC Appliances in India*

DC appliances are very limited in the residential and commercial loads in India. The limited usage of DC appliances limits the number of manufacturer and as the demand of DC appliances is less, the manufacturing cost is high due to which the price of the appliance increases. The various factors which influence the rise in cost of DC appliances in India are limited usage, manufactured by small organizations with less buying power, less partnership among the manufacturers, harmonization of standards in alignment with the global world. The cost of the appliances also depends on the

**Table 3** DC appliance market in India

| Appliance type | Price range | Units (2018) | Unit (2023) |
|---|---|---|---|
| Fans [8] | 2000–2500 | 56,112 | 188,220–243,301 |
| Entertainment [9] | 15,000–2000 | 8042 | 55,226–154,634 |
| Refrigeration [9] | 6000–12,000 | 3883 | 26,708–53,416 |

**Table 4** Average price of various DC appliances in India

| Appliance | Rating | Price range in INR |
|---|---|---|
| Ceiling fans | 35 W, 12 V | 2800–3200 |
| Lightings [12] | 12–20 W, 12 V | 180–350 |
| Water pumps | 0.5 to 1.5 Hp<br>90 feet to 250 feet | 15,000–20,000 |
| Refrigerators | 150–200 L, 24 V | 15,000–22,000 |
| Solar cooker [14] | 400 W, 12 V | 5500–6000 |

quality and robustness which depends on the seal, protection, enhanced electromagnetic shielding and certification. This means that the cost of the DC appliance cannot be compared directly. Table 3 gives the details of the present estimated market size of DC appliances in India during 2018 and 2023 [13]. From the table it is inferred there is a scope for the DC appliance industry to boom in coming decades which will lead to the reduction in the cost of the DC appliances.

.

The cost of AC LED lamps varies from INR 100 to INR 500 based on the wattage and energy saving technology whereas the DC lamps are available in the market at a cost of INR 250–750 for the same wattage range. The cost of DC water pump of centrifugal type, 0.1 to 1 Hp costs about INR 15,000. Similarly, DC powered refrigerators 200 L are available at cost of INR 15,000. Table 4 gives an average price of various DC appliances in India and the cost is an average of the cost quoted by various private players like prozone systems, sinfin, Cygni Energy private limited etc. [14]. It is noticed that the cost of some of the DC appliances increases due to the presence of DC-DC converters in some DC products. Based on the analysis the cost of the DC appliance may degrade as the market is booming due to its higher efficiency and due to emerging technologies like distributed energy systems with PV and battery energy storage systems [15].

The DC appliances still may grow due to availability of Solar based bundled appliances in the market which could operate in dedicated off grid power system. Normally they are presented as Solar Hybrid System (SHS) and Energy Storage System (ESS) with load capacity of 500 W to 10 kW along with the Li-ion battery. In order to improve the performance of the DC solar system the systems are provided with motion sensors Remote control wireless touch panel systems, MPPT based solar charge controllers. To compensate the voltage and current drops after long wiring DC-DC boosters are used with conversion efficiency of about 96%. The SHS and

ESS do have the option to drive both AC and DC loads. The demand of appliance may further grow due to the consumer sensitivity, consumer preferences and availability.

## 5 Challenges and Way Forward to DC Appliances in India

The major factors which influence the implementation of DC appliances at domestic applications are: DC ready products Cost of the appliance, Limited manufactures, operating voltages, Standards and awareness among consumers. One of the factors which hinders is the in-availability of DC-ready products in the market. The manufacturer's manufactures a product only of there is a demand. As the demand over DC appliances are very less the manufacturer's interest in investment and R&D over the product is negligible. In order to move forward, greater availability of DC-ready appliances in bulk procurement programs has to be organized that match large purchasers of appliances with potential vendors. To make the product viable, it can be designed that it works both on AC and DC. This makes the consumers to ensure DC compatibility in the future, even if the distribution infrastructure is not changed. The next major challenge is the price that is higher when compared with AC appliances. For example, the cost of BLDC motor in India is INR 2000–2500 which is 40–50% higher than an AC fan. Though the DC appliances have greater advantages like efficiency and better life, the cost affects the sale of DC appliances. Limited manufacturers make the availability of the product volume low and prevent the economies of scale to emerge. As discussed earlier as the demand is very less, the manufactures are not keen to enter the DC appliance market and the willing to invest in Research and Development is limited. The DC appliances can be operated in various operating voltages like 12, 24, 48 V etc. Different player in the market introduce the equipment's in different rating, therefore a standard operating voltage for power has to be enforced among the private players. Harmonization of standards in alignment with the global standards allows the manufacturers to cater the international market too. In order to break through it, IEC has prepared draft road map proposal for LVDC (Low Voltage DC) electrical distribution for 48 V. It has standards which enable to compute specifically for fans, Lightings, TVs etc. The BIS (Bureau of Indian Standards) also have published guidelines to standardize the 48 V ELVDC distribution systems in 2017. It provides the essential requirements for an extra low 48 V DC source. The standard acts as a step to increase the potential of DC systems.

The consumer awareness is one of the key point while introducing a technology. As the awareness among the Indian consumers regarding DC is very less, it lacks the improvement in the market of DC appliances. Though several awareness programs were organized by national bodies, still more are required to educate the consumers. Also DC resource center may be established with DC product database, Design methodologies etc. As the cost of DC appliances is high and are available in bundles, financing will uptake the DC appliances going forward. Some of the MFI like Muthoot Microfin and Fullerton have product loan portfolio to include DC TV and fans. In India, both the Government and the private players have worked to make

the electrical power available to every home. The Soubhagya scheme has fueled the intensive electrification which increases the electrical energy demand. In the scheme, the government has sanctioned standalone PV solar system for 0.35 million households, which make the consumers to utilize the DC appliances. The solar home system ranges from 200 to 300 W' power which has a connected of 5 LED lights, 1 DC fan and a TV terminal. The Remote Village electrification program and off-grid and centralized solar PV application program have also paved the way for solar homes which can utilize DC appliances. From the above discussion it is evidently understood that the scope of DC appliance in India will take-up in mere future. The comparative analysis of the conventional methods is explained in Table 5. In this table, latest DC distributed systems have been explained with limitations.

**Table 5** Comparative analysis

| Author | Year | Methodology | Advantage | Limitation | Performance analysis |
|---|---|---|---|---|---|
| Vossos et al. [4] | 2017 | DC distributed system | Provides cost assessment to the current DC ready appliance | Unwanted heat generation | Power consumption in uW—184 Frequency in MHz—214 |
| Ginart and Sharifipour [5] | 2021 | Cloud change system | The power generation is improved and the power loss is reduced by integrating solar panels and battery storage | Cost expensive in DC appliances | Power consumption in uW—194 Frequency in MHz—204 |
| Siraj and Khan [6] | 2020 | Energy efficient DC distributed system | The developed system significantly reduced distribution and power electronic losses | DC transmission setup is cost expensive and complex | Power consumption in uW—214 Frequency in MHz—254 |
| Sabry and Ker [8] | 2020 | DC environment compressors | By analyzing the experimental results, speed of the power transmission is improved | DC shunt motors size is too large | Power consumption in uW—247 Frequency in MHz—264 |
| Nandini et al. [9] | 2021 | DC micro wind system | The power distribution is high, even while the load is high | Installation cost is high | Power consumption in uW—114 Frequency in MHz—218 |

## 6 Conclusion

From the above review, the conclusion is that the energy sector in India is increasing day by day, so energy conservation becomes unavoidable. In order to ensure the conservation of electrical energy DC appliances plays a major role. The Indian market do have DC appliances for domestic application but the DC ready products are limited due to less demand and less awareness among consumers. However, there is an increasing opportunity for the DC products due to the various government programs and Off-grid electrification schemes. The DC appliances will evolve in the market in mere future due to its extra efficiency and as the power generated from solar is DC. As the power generated from Solar is DC, it has enabled direct DC distribution system due to which DC appliances are available in bundles along with the Solar PV systems. Also, the energy loss is reduced due to the addition of battery energy storage systems which help the consumers to utilize the power when the solar power is not available. The increasing PV based homes increases the number of consumers makes more private players to evolve in the DC appliance market which will eventually decreases the cost of the DC appliances which in turn further enhances the utilization of DC appliances among the consumers.

## References

1. Starke M, Tolbert LM, Ozpinecei B (2008) AC Vs DC distribution: loss comparison. In: 2008 IEEE/PES transmission and distribution conference and exposition, pp 1–7
2. Backhaus S et al. (2015) DC microgrids scoping study—estimate of technical and economic benefits, Los Alamos National Laboratory, LA-UR-15-22097
3. Vossos V, Garbesi K, Shen H (2014) Energy savings from direct-DC in U.S. residential buildings. Energy Build 68(Part A):223–231
4. Vossos V, Johnson K, Kloss M, Khattar M, Brown R (2017) A market assessment for DC distribution systems in buildings. Lawrence Berkeley National Laboratory, Berkeley, CA
5. Ginart A, Sharifipour B (2021) High penetration of electric vehicles could change the residential power system: public dc fast chargers will not be enough. IEEE Electrific Mag 9(2):34–42
6. Siraj K, Khan HA (2020) DC distribution for residential power networks-a framework to analyze the impact of voltage levels on energy efficiency. Energy Rep 6:944–951
7. Sabry AH, Shallal AH, Hameed HS, Ker PJ (2020) Compatibility of household appliances with DC microgrid for PV systems. Heliyon 6(12):e05699
8. Sabry AH, Ker PJ (2020) DC environment for a refrigerator with variable speed compressor; power consumption profile and performance comparison. IEEE Access 8:147973–147982
9. Nandini KK, Jayalakshmi NS, Jadoun VK (2021) An overview of DC Microgrid with DC distribution system for DC loads. Mater Today Proc
10. Garbesi K, Vossos V, Shen H (2011) Catalog of DC appliances and power system. Energy Analysis Department, Lawrence Berkeley National Laboraroty, United States
11. Rasheed A, Khan S, Erteza Gelani H, Dastgeer F (2019) AC vs DC homes: an efficiency comparison. In: International symposium on recent advances in electrical engineering, pp 1–6
12. Nordman B, Christensen K (2015) The need for communications to enable DC power to be successful. In: IEEE first international conference on DC microgrids (ICDCM), pp 108–112
13. Hoque SN, Das BK (2013) Analysis of cost, energy and emission of solar home systems in Bangladesh. Int J Renew Energy Res IJRER 3(2):347–352

14. Rathore KS, Kalla UK, Palwalia DK, Singh B, Mishra AK (2020) Voltage-controlled power factor corrected CSC derived DC–DC converter for PMBLDC driven home appliances. IET Power Electron 13(15):3407–3418
15. Marahatta A, Rajbhandari Y, Shrestha A, Singh A, Gachhadar A, Thapa A (2021) Priority-based low voltage DC microgrid system for rural electrification. Energy Rep 7:43–51

# A Survey on Vehicle Detection and Classification for Electronic Toll Collection Applications


**N. Sathyanarayana**

**Abstract** One of the primary concerns of an electronic toll collection (ETC) system is to automate vehicle detection and classification which is essential so that the toll is charged at the toll plaza according to the vehicle type. Automated vehicle classification has received a lot of attention due to its applicability in a wide variety of applications. Although the challenges of image based solutions in less constrained environments are known, they often have advantages of relatively easy installation and low set up cost. This justifies the time and effort in devising an image based solution either in addition to other components or as a main component. This paper presents a survey of various techniques for vehicle detection and classification along with their advantages and disadvantages, with more emphasis on electronic toll collection. This study deals with various approaches used for vehicle detection and classification but emphasized more on computer vision and deep neural networks based techniques.

**Keywords** Computer vision · Electronic toll collection system · Toll plaza

## 1 Introduction

Vehicle identification is a common task in most applications, such as traffic monitoring, video surveillance etc., related sub tasks include vehicle localization, vehicle counting, license plate recognition and so on. Vehicle classification has many practical applications such as traffic density estimation and video surveillance. Another application is for toll collection on highways. Electronic toll collection (ETC) is becoming an increasingly common trend in road pricing system all over the world, the main aim being eliminating delay on toll roads. Different approaches have been adopted in different countries, these are the result of a combination of the policy and the underlying technology. A number of existing solutions are based on radio frequency identification or inductive loop based technologies involving many sensors. These sensors include magnetic sensor, laser sensor, strain gauges, multiple

N. Sathyanarayana (✉)
Vemana Institute of Technology, Bengaluru, India
e-mail: sathyanarayana28@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_10

calibrated cameras etc. In many existing systems an image based system is either present only as an add-on or not at all. A computer vision-based solution would be useful even if the main component of the solution is not image-based, it serves as an additional cue for improving system performance as well as an independent auditing mechanism [21].

Computer vision has been widely used in various vehicle detection and classification applications such as surveillance, vehicle counting and traffic density estimation. In this work varieties of approaches used for vehicle classification near the toll plazas emphasizing on computer vision-based solutions are reviewed. Rest of the paper contains the following, Sect. 2 provides an overview of various approaches and existing solutions for automated toll collection, Sect. 3 overviewed different computer vision based techniques that have been employed in applications involving vehicle classification and Sect. 4 has discussions and conclusions of the work.

## 2 Related Work

There are many approaches used for vehicle classification near the toll plaza for automatic charging of the appropriate toll charge based on vehicle type. Some of the classical methods used for vehicle classification and their limitations are discussed in this section.

### *2.1 Axle Based Approaches*

Many axle based methods for vehicle type identification employ sensors which are used for counting the number of axles and also determine heights of various axles. Most such systems employ a classification back-end system, typically a simple classification mechanism that classifies the vehicle based on the above features such as number of axles and height of axles. Ma et al. [1] developed a system which uses accelerometer sensors to find out the axle locations and magnetometer sensors to estimate the speed of the vehicle by combining both, Axle spacing and number of axles are obtained which are latter fed to a classification scheme which classifies the vehicles. But, the sensors used in this system were battery operated and therefore had a very less efficiency and also this system was not suited for all traffic environments. Bitar et al. [2] developed an approach for solving the problem of assignment in obtaining the optimum thresholds for axel-based vehicle classifiers. Gaussian distribution fitting and histograms were individually done for each class, for each axle spacing, and from the data gathered. The derived distributions were used to compute the optimal class boundary thresholds and a new algorithm for classification was constructed. As a result, the error rate was reduced for all the classes. But, comparing the original algorithm with newly developed one found challenging

due to the impossibility of using one dataset in all stations as all the classification Stations in Oklahoma incorporate the same classifier.

## 2.2 RFID Based Approaches

There are numerous works based on Radio Frequency Identification (RFID) for vehicles. The applications are varied, ranging from vehicle type identification for security and surveillance applications, vehicle counting, intelligent vehicle speed control and so on. Sliwa et al. [3] proposed a radio-finger printing-based system for vehicle classification. They claim that their proposed system is able to provide real-time and precise vehicle classification. And they also claim that their system is weather independent, good in privacy preservation, cost-efficient in installation and maintenance. This system is formulated for a binary classification for only truck and car, using a feature vector with six signal-based features. The highest accuracy for classifying trucks and cars was more than 99%. This system was implemented for binary classification so multi-class classification was not possible here. Jain et al. [4] developed a system for toll collection with 2 levels of security, one is OCR (Optical Character Recognition) and the other is RFID. The OCR is used to read the number plate and RFID reader near the toll will be used to read the RFID tag inside the Vehicle and obtains the details such as owner name, vehicle number, type and a 16 digit unique ID provided to each vehicle at the time of registration. Later the vehicle number obtained from OCR system and the RFID system are compared if does not match then the vehicle will stop and verified for stolen vehicle or any error due to the reading system, if it matches then the toll amount will be deducted automatically from the customers e-wallet and the vehicle is allowed to pass through the toll or if the funds are insufficient then the vehicle is stopped.

## 2.3 Limitations of Existing Technologies

Although some of these systems have good accuracy, one disadvantage is that the devices such as sensors are expensive to install and maintain. Any miss-alignment of the sensors will lead to a wrong interpretation. Again re-laying the sensors is a tedious task and requires manpower and the system requires a large amount of start-up expenses. Since the vehicle has to stop for a certain time to acquire the data, this directly introduces delays on toll roads. Among other technologies loop-based technologies will exhibit poor performance under congestion. Vehicle length based methods can distinguish only trucks and cars, and therefore they are not suited for applications that needs more detailed classification. Non-intrusive technologies such as acoustic and infrared sensing are prone to weather and lighting conditions. Usually some of these systems also incorporate an additional computer vision based component. Methods based on RFID tags are having several drawbacks such as

i) High cost (Active RFID is costly due to the use of batteries). ii) RFID can be easily tapped or intercepted. iii) RFID devices need to be programmed which is time consuming. iv) External electromagnetic interference. v) Limited range (only 3 m). vi) The problem of Misusing of RFID tags and number plates still exists.

## 3 An Overview of Different Computer Vision Based Approaches for Vehicle Classification

Over the years' number of computer vision techniques are being proposed for vehicle detection, classification and tracking. This section outline different types of computer vision approaches and their representative examples. The taxonomy of various computer vision approaches used for vehicle classification are shown in Fig. 1. The interested reader may consult these for a further overview of vehicle detection and classification.

### *3.1 Feature Engineering Approaches*

Some of the earlier works on vehicle detection followed the paradigm of detecting key features from the images and using these as inputs to one or more classifiers. Early work on vision based systems for detecting vehicles focused mainly on simple features with clear physical meanings, such as edges, symmetry, textures, underbody shadows, and corners. Morphology based features are extensively used for various computer vision tasks. Khanipov et al. [5] developed a system in which various computer vision techniques have been applied to identify the features of the vehicle like number of axles, height of the vehicle on the first axel and vehicle length etc. to classify the vehicle types. But this system is capable of operating only on video

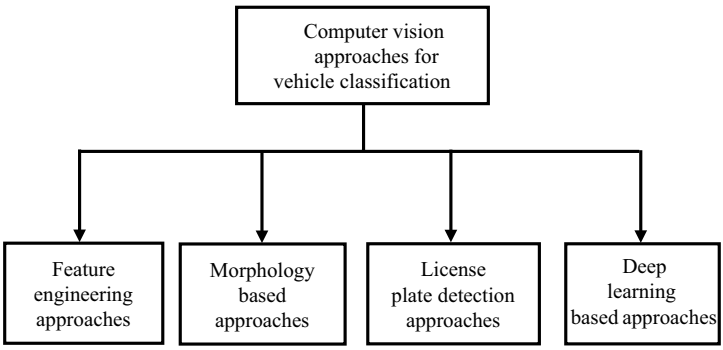



**Fig. 1** Taxonomy of various computer vision approaches used for vehicle classification

stream and needs induction loops for its working. Suryakala et al. proposed a method [6] which uses Haar cascade classifier for vehicle detection and background separation using KNN algorithm for pedestrian detection and this method could detect only 196 vehicles out of 290 vehicles. Yuan et al. [7] implemented a method in which the vehicle detection is carried out in two parts. In the first part target location is done by using the feature-based method of maximum classes square error and in the second part vehicle detection is done by using machine learning based method, Haar-based Adaboost classifier. These types of approaches fail frequently in complex backgrounds or under variations in illumination. Subsequent works used more sophisticated features and techniques.

### *3.2 Morphology Based Approaches*

The approaches based on morphology uses operators like erosion, dilation, closing, opening are very popular for applications like Character recognition, license plate recognition and it can be used even for detecting vehicles and classifying them. Ajmal et al. [8] has come up with a system in which areal images taken from a camera on highways is preprocessed by mask operation, multiple thresholding, image differencing and filtering then these images are given for vehicle detection which performs series of operations like edge detection, binary dilation, binary filling, opening and boundary vehicle removal and so on later its sent for vehicle classification and they achieved a success rate of 85%. Some of the methods based on morphology and their challenges are outlined in Table 1. As we can observe from the challenges faced by many researchers, morphological methods are not well suited for vehicle detection and classification.

### *3.3 License Plate Detection Approaches*

There is a significant amount of work related to license plate recognition. License plate recognition is especially relevant in applications which involve vehicle identification and where the license plate is clearly visible. Some of these methods are outlined in Table 2.

### *3.4 Deep Learning Based Approaches*

The year 2012 can be considered a landmark in image classification and in many ways heralded the deep learning age. In that year a deep convolutional neural network named AlexNet, designed by Krizhevsky et al. [15] was used for winning the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). Convolutional

**Table 1** Morphology based approaches for vehicle classification

| Year | Brief description | Challenges |
|---|---|---|
| 2010 | Pandu Ranga et al. [9] Developed a system based on Morphological technique in which the image of a road captured from top view is taken and it is compared with the reference image of the road without vehicles. The difference image will give the vehicle information, edge detection of which gives the boundaries of the vehicles on which morphological operations are performed to make it in the form of square boxes, later based on the total number of pixels present in that area of the boxes vehicles are classified into three categories like small (100–400 pixels), medium (401–700 pixels) and big (701–1000 pixels) | 1) There is no finer classification of vehicles<br>2) This system is more prone for inter class classification errors |
| 2015 | In this system [10] the image of a vehicle passing through the toll is captured by the still camera then based on the area occupied by the vehicle in the image they are classified as light or heavy. This information is fed to the raspberry pi which intern connects to the web server and charges the appropriate toll tax | This type of classification method may not suit for some vehicle types |
| 2017 | Sarikan et al. [11] developed an automated method of vehicle classification to distinguish different types of vehicles. In this approach computer vision and machine learning techniques are employed in classification of vehicles in the dedicated lanes. They specifically address the problem of classification between motorcycles and sedans. Image morphological features serve as inputs to a classifier, they used a KNN and decision tree classifiers | This method is focused only on motor cycle detection |

Neural Networks (CNNs) are the most frequently employed architectures for Deep Networks in image classification. Some of the well-known older architectures include AlexNet, LeNet-5 and VGG 16. Newer architectures include Inception, ResNet, ResNeXt and DenseNet.

Deep learning based methods have demonstrated significant successes on various image classification tasks such as object detection (including multi object detection), image captioning and so on. A significant advantage offered by deep learning is that it combines feature extraction and classification. Deep learning methods have been used in detection of vehicles and classifying them, an overview of some of the recent deep learning approaches are presented in Table 3.

**Table 2** License plate detection based approaches for vehicle classification

| Year | Brief description | Challenges |
|---|---|---|
| 2012 | In this system [12] the vehicle is detected by the proximity sensor and signals the PC on the arrival of vehicle at the toll plaza. The camera captures the front view image of the vehicle and the algorithm recognizes the license number of a vehicle, based on which the toll is charged | 1) It's unable to detect the targets which are fast changing, because of the low quality camera used in this system<br>2) Since it used template matching for number identification it is unable to distinguish some of the characters like 8 & B or 0 & O |
| 2013 | In this approach [13] the camera captures the image of a License number plate which is then given to the ANPR to identify the vehicle number in the form of text. The toll will be deducted automatically from the owners account then the toll gate will be opened. This approach also gives a buzzer if the vehicle is stolen and the vehicle details are updated in the database as stolen vehicle | 1) This system needs the font size of a number plate to be fixed<br>2) For this system to work all the vehicle details must be updated in the database<br>3) Customer should have an active account linked with the license number |
| 2020 | Tamboli et al. [14] proposed a system for number plate recognition using IOT and machine learning. In this approach The camera captures the color image of a vehicle which is then preprocessed by converted to greyscale, eliminating noise and performing then binarization. Later the number plate is segmented by using sobels edge detection algorithm and the characters are extracted by horizontal scanning, these characters are fed to the Trained CNN for character recognition. Once the license number is identified toll will be charged from the owners account and the intimation will be sent through SMS | 1) For this system to work all the vehicle details must be updated in the database<br>2) Customer should have an active account linked with the license number |

## 4 Discussions and Conclusions

This work summarizes the work related to vehicle classification using computer vision. The motivating application was automatic vehicle type identification in toll plaza, however most of the sub tasks are applicable to other applications that involve vehicle recognition. Some of the challenges specific to Indian context have been highlighted. A representative set of work utilizing different technologies was outlined. Paid particular attention to computer vision based work i.e. where the input was a video or an image. Deep learning based methods have been extensively used for various computer vision tasks such as object detection and classification, we expect the trend of increased use of deep learning based methods for computer vision tasks. Accordingly, this paper summarizes the recent works related to deep learning used in vehicle detection and classification.

**Table 3** Deep learning based approaches for vehicle classification

| Year | Brief description | Challenges |
|---|---|---|
| 2018 | In this technique to enhance the classification performance, the authors Han et al. [16] have used an unsupervised pre-training approach to better initialize CNNs parameters. In addition, a trained object detection model was used to remove unrelated background as much as possible. They evaluated their approach on about 2000 labelled images for each category consisting of 4 categories, the authors claim an accuracy of 93.50% obtained using their approach | This approach was made to classify vehicles into only four classes |
| 2020 | Wong et al. [17] developed a model in which the author used CNN implemented using Tensorflow written in python for vehicle classification for classifying the vehicles into three classes namely Buses, Cars and Trucks. They used 500 images for each class of the vehicles, out of which 400 images were used for model training and 100 images for model validation purpose and achieved an validation accuracy of 93.8%.In the tests they used 10 images for each class out of which 9 were classified correctly in each class yielding an accuracy of 90% | This model had difficulties in differentiating vehicles of one class with the other class with similar visualization |
| 2020 | In this system the authors Alessio et al. [18] used CNN for detection and classification of vehicles in videos with different conditions such as low resolution, low illumination, blurred and inclement weather. They have used 107,000 images for training and 17,000 images for validation and classified the vehicles into five classes such as Motorbike, Bus, Truck, Car and Van with an accuracy of 92% | This system was designed only for single vehicle images and will not suit for mages with multiple vehicles |
| 2021 | Atif Butt et al. [19] developed a CNN based system for vehicle classification which uses a self-constructed local dataset of 10,000 images and 50,000 VeRi dataset consisting of six vehicle classes for fine tuning an improved ResNet-152 by adding a new block for classification and the system achieved an accuracy of 99.68% in classifying the vehicles into six classes namely Car, Truck, Van, Rickshaw, Motorbike and Minivan | Data set is limited and the system is not tested on a practical scenario |

(continued)

**Table 3** (continued)

| Year | Brief description | Challenges |
|---|---|---|
| 2021 | Karungaru et al. [20] developed a system which involves two parts. The first part involves vehicle detection for which they proposed Yolov2-tiny for target extraction, and performed K-means clustering and network parameter adjustment for network training and for training the detection part Darknet was used. The second part of the work is vehicle type classification which uses improved CNN for feature extraction and AlexNet with Spatial Pyramid Pooling for classification. After the CNN training they performed SVM training to reduce network overfitting also to improve the network accuracy. As a result they could achieve the classification accuracy of 98.87% for Car, 98.02% for Bus, 76.50% for Van and 74.03% for Other vehicles | 1) They could not combine both the parts of the work together<br>2) They could classify the vehicles into only four classes namely Car, Bus, Van and Others<br>3) Though the accuracy for Bus and Car was satisfactory for Van and Others it showed a poor response |

## References


1. Ma W et al (2014) A wireless accelerometer-based automatic vehicle classification prototype system. IEEE Trans Intell Transp Syst 15(1):104–111
2. Bitar N, Refai HH (2017) A probabilistic approach to improve the accuracy of axle-based automatic vehicle classifiers. IEEE Trans Intell Transp syst 18(3):537–544
3. Sliwa B et al (2018) A radio-fingerprinting-based vehicle classification system for intelligent traffic control in smart cities. In: Annual IEEE international systems conference (SysCon)
4. Jain P et al (2018) A unique identity based automated toll collection system using RFID and image processing. international conference on computing, power and communication technologies (GUCON) Galgotias University, Greater Noida, UP, India, 28–29 September
5. Khanipov T et al (2015) Vision-based industrial automatic vehicle classifier. In: Seventh international conference on machine vision (ICMV 2014), Proceedings of SPIE, vol. 9445. 944511-1
6. Suryakala S, Muthumeenakshi K, Joseph Gladwin S (2019) Vision based vehicle/pedestrian detection in traffic surveillance system. In: International conference on communication and signal processing, 4–6 April India
7. Yuan1 C et al (2019) Research on vehicle detection algorithm of driver assistance system based on vision. 978-1-7281-0106-4/19/$31.00c IEEE
8. Ajmal A et al (2010) Vehicle detection using morphological image processing technique. In: International conference on multimedia computing and information technology (MCIT), Sharjah, United Arab Emirates, 2–4 March
9. Pandu Ranga, HT et al (2010) Vehicle detection and classification based on morphological technique. IEEE. 978-1-4244-8594-9/10
10. Suryatali A, Dharmadhikari VB (2015) Computer vision based vehicle detection for toll collection system using embedded Linux. In: International conference on circuit, power and computing technologies
11. Sarikan SS et al (2017) Automated vehicle classification with image processing and computational intelligence. Procedia Comput Sci 114:515–522. CAS October 30–November 1, Chicago, Illinois, USA

12. Soomro SR et al (2012) Vehicle number recognition system for automatic Toll tax collection. TE-08, Sukkur IBA
13. Chhoriya P et al (2013) Image processing based automatic toll booth in Indian conditions. IJETAE 3(4). ISSN 2250-2459, ISO 9001:2008 Certified Journal
14. Tamboli A et al (2020) Vehicle number plate recognition for automatic toll tax collection using IoT and machine learning. Int J Future Gener Commun Netw 13(3s):1685–1694. ISSN 2233-7857 IJFGCN
15. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems (NIPS), pp 1097–1105
16. Han Y et al (2018) Pretraining convolutional neural networks for image-based vehicle9 classification. In: Advances in multimedia
17. Wong ZJ et al (2020) Vehicle classification using convolutional neural network for electronic toll collection. ISBN 978-981-15-0058-9
18. Alessio et al (2020) Vehicle detection and classification in difficult environmental conditions using deep learning. ISBN 978-3-030-55180-3
19. Butt MA et al (2021) Convolutional neural network based vehicle classification in adverse illuminous conditions for intelligent transportation systems. https://doi.org/10.1155/2021/6644861
20. Karungaru S et al (2021) Vehicle detection and type classification based on CNN-SVM. https://doi.org/10.18178/ijmlc.2021.11.4.1052
21. Satyanarayana GSR et al (2021) A vehicle detection technique using binary images for heterogeneous and lane-less traffic. IEEE Trans Instrum Meas 70:1–14. https://doi.org/10.1109/TIM.2021.3062412. Art no. 5007514

# A Systematic Study of Sign Language Recognition Systems Employing Machine Learning Algorithms


**Pranav and Rahul Katarya**

**Abstract** The deaf and mute population struggles a lot in expressing their thoughts and ideas to others; Sign Language is the most expressive means of communication for them, but a majority of the general population is callow of sign language, hence the mute and deaf experience difficulties while communicating to the rest of the world. To overcome this communication barrier, a device that can accurately translate sign language gestures to speech and vice-versa in real-time is needed. There exist solutions for converting verbal or written language to sign language in real-time reliably and accurately, however the same cannot be said about translating sign language to textual and/or vocal format. The currently existing systems either do not support communication in both directions, are not real-time, have low recognition accuracy, or require static surrounding conditions. Some systems require additional hardware components like expensive sensors, which tend to increase the cost. In this survey, we have reviewed numerous existing solutions and have categorized them depending on the method used. We hope that the results obtained from this study may serve as a road map to guide future study in the domain of Sign Language Recognition (SLR).

**Keywords** Inertial Measurement Unit (IMU) · Neural Networks (NN) · Surface Electromyogram (sEMG)

## 1 Introduction

As of March 2020, World Health Organization reported that almost 430 million people in the world suffer from severe hearing problems [1]. Also going by the 2011 population census of the government of India [2], there were approximately 1.64 million people who suffer from speech disability, and 1.26 million people suffer from hearing impairment. The deaf and mute population resorts to using sign language

Pranav (✉) · R. Katarya
Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science, Delhi Technological University, New Delhi, India
e-mail: pranav_2k20cse16@dtu.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_11

for communicating, which involves hand gestures, body posture, and facial expressions. However, as the mass population is uninformed of the meaning of signs in sign language, the Deaf-Mute population experiences difficulties in expressing their thoughts. There are three possible solutions to this problem, the first of them being teaching sign language to the masses, which is not possible, as most of the population might not be willing to do so. The other two methods being vision-based SLR systems, and sensors-based SLR systems. Some of the existing solutions in the field are discussed in Sect. 2. The taxonomy of signs and some sign languages are discussed in Sect. 3. Section 4 describes the working of SLR. Sections 5 and 6 provides the conclusion of the paper and aims to provide a roadmap for future development in the field.

## 2 Some Existing Sign Language Recognition Systems

SLR systems can be broadly classified into two categories: 1) Sensor-based systems; 2) Vision-based systems.

### *2.1 Sign Language Recognition Systems Based upon Sensors*

For data collection, this method uses a data glove with sensors embedded on the fingers, palm, and arm of the signer. Some of the existing sensor-based systems in the field are demonstrated in Table 1.

In [3], they used a Myo armband [23] worn on the signer's forearm, which consists of sEMG sensors to recognize muscle activity. This data was passed on to a custom Artificial Neural Network (ANN) for recognizing some basic words of the Chinese sign language (CSL). In [4], they used a Photoplethysmography (PPG) sensor (to detect contractions of the forearm muscles) with IMU (to detect orientation and

**Table 1** Some existing sensor-based SLR systems that use Machine Learning Techniques

| Ref. | Sensors | Classifier | Type | No. of signs | Accuracy (%) |
|---|---|---|---|---|---|
| [3] | sEMG | ANN | Static | 15 | 88.7 |
| [4] | IMU, PPG | ResNet, GBT | Static | 9 | 98 |
| [5] | IMU | CapsNet | Dynamic only | 60* | 94 |
| [6] | Flex, IMU | kNN, DTW; CNN | Dynamic only | 10 | 96.6, 98 |
| [7] | sEMG, IMU | DBN | Dynamic | 150 | 95.1 |
| [8] | IMU | LSTM | Dynamic | 28 | 99.89 |
| [9] | EMG, IMU | ANN; SVM | Dynamic | 13 | 93.8, 85.6 |

**Note.**—Static = Static signs only; Dynamic = Both Static and Dynamic signs unless explicitly specified; * = approximately

motion of hands) on the signer's wrist. The collected data were pre-processed and passed on to a Gradient Boost Tree (GBT) and deep Residual Network (ResNet) for classifying signs for numbers one to nine of the ASL. In [5], they used a deep capsule network (CapsNet) for recognizing gesture data of approximately 40–80 (they used 20 sentences, with each sentence consisting of 2–4 words) words of the ISL, captured using an IMU. In [6], they used multiple flex sensors attached to a glove, in conjunction with three IMUs on the signer's arm. They used two methods for classifying the gestures: 1) k-Nearest neighbors (kNN) with Dynamic time warping (DTW) method, 2) Convolutional neural networks (CNN). They tested their models on 10 different gestures of the Italian sign language. In [7], they used an sEMG sensor along with IMU to capture muscle activity data along with hand motion data to capture gesture features and employed a Deep Belief Net (DBN) deep learning model for recognizing 81 single-handed and 69 two-handed words of the CSL. They achieved recognition accuracy of up to 88% for user-independent tests and 95% for user-dependent tests. In [8], six IMUs were used on fingers and back of the palm of the signer's hand to acquire hand motion and finger movement data for gestures. They used Long-term short memory (LSTM) deep learning algorithm for recognizing some commonly used sign words of the ASL. In [9], they used a Myo armband consisting of EMG sensors and IMU to capture features of sign gestures. The system uses Artificial Neural Networks (ANN), and support vector machines (SVM) for gesture recognition. They tested their system using a set of 13 commonly used ASL sign gestures.

## *2.2 Sign Language Recognition Systems Based on Vision-Based Techniques*

For data collection, this method uses a camera module directed towards the signer to capture images/video of the gestures. Some of the existing solutions utilizing computer vision techniques for SLR are demonstrated in Table 2.

In [10], RGB gesture images were initially cropped down to reduce processing load and then converted to grayscale. They used multiple images of the same gesture for training and testing the performance of the You Look Only Once (YOLO) system, which is a CNN based object detection method. Their systems achieve 100 hundred percent accuracy on English alphabet fingerspelling images of the Indonesian sign language, however, accuracy drops significantly when testing on videos. In [11], they used RGB images along with a depth map (obtained using Kinect sensor [24]) of fingerspelling of English alphabets of the ASL. They used CNN for recognizing the signs and achieved up to 99.79% accuracy. In [12], they used VGGnet and ResNet architectures of CNN for gesture recognition. Their dataset includes 32 static Arabic sign language gestures, with 800 images of each. In [13], they used a Leap motion sensor [25] to acquire the spatial orientation of the signer's hand and fingers. They used a Hidden Markov classifier (HMC) for recognizing the gestures. In [14], they

**Table 2** Some existing vision-based SLR systems

| Ref. | Classifier | Type | No. of signs | Accuracy (%) |
|---|---|---|---|---|
| [10] | YOLO | Static | 24 | 73 |
| [11] | CNN | Static | 24 | 99.79 |
| [12] | CNN | Static | 32 | 99 |
| [13] | HMC | Static | 24 | 86 |
| [14] | Multi-class SVM | Dynamic | 35 | 87.6 |
| [15] | I3D CNN | Dynamic | 249 | 64.44 |
| [16] | PCA, SVM | Static | 35 | 95.31 |
| [17] | CNN | Static | 71 | 93.04 |

**Note.**—Static = Static signs only; Dynamic = Both Static and Dynamic signs

used a Kinect sensor to capture skeletal images of the signer. However, the Kinect does not capture does not capture finger joints and cannot distinguish signs involving finger articulations and orientation, which greatly reduces the number of signs that can be recognized using it. They used "Multi-class SVM with radial basis function kernel" on a dataset of 35 ISL gestures. In [15], they used a "two-stream model of inflated 3D (I3D) ConvNets". The first stream was fed RGB data and the second was fed optical flow data of sign videos. Using only RGB, or optical flow data, recognition accuracy was significantly reduced compared to when using both of them together. In [16], they used static images of ISL gestures English alphabets and numbers 1 to 9. They used Grayscale conversion followed by segmentation, and noise removal before subjecting the images to the skin thresholding process. The output image of this step is a small-sized image including only the hands of the signer, from which PCA was used for extracting the features of the gesture, which were then fed to an SVM classifier for recognizing the gestures. In [17], they used a dataset of English alphabets, numbers, and 35 common static gestures of the ASL. They used skin-color based segmentation to extract hand region from images, which were then fed to a Keras and CNN classifier for recognizing the signs. The problem with this system is that it takes a lot of time to recognize the gestures, averaging 2.6 s for each sign in the dataset, which is not acceptable for real-time SLR systems.

## 3 Sign Languages and Taxonomy of Signs

### *3.1 Types of Signs*

SL gestures can be performed using one hand or both hands, they can be static or dynamic, manual (include hand gestures only) or non-manual (include other physical features like mouth movements, facial expressions, and body orientation). Dynamic signs can further be classified as type 0 and type 1, as shown in Fig. 1. The type 0 sign


**Fig. 1** Taxonomy of sign language gestures [18]

indicates a two-handed dynamic sign, in which both hands are performing motion, whereas in a type 1 sign is a two-handed dynamic sign with only the primary hand performing motion [18].

## *3.2 Problem with Universal Sign Language Recognition System*

There is no universally accepted sign language. Just like spoken languages, over time as people communicated with each other, sign languages also spread out and took multiple forms. There are more than a hundred different forms of sign language in use today all over the world. [19] Even countries that have a common spoken language do not necessarily share a common sign language, e.g. America, Britain have English as their main language, yet both nations have their own sign language viz. American sign language (ASL), and British sign language (BSL).

Fingerspelling for English alphabets in the ASL is done using a single hand, and signs for most of the alphabets are static in nature, as shown in Fig. 2; while Fingerspelling of English alphabets in the Indian sign language (ISL) mostly involves using both hands, and the signs are static in nature, as shown in Fig. 3. Also, other countries having their own sign language (e.g. Brazilian, Chinese, etc.) makes the existence of a single universal solution for SLR impossible.

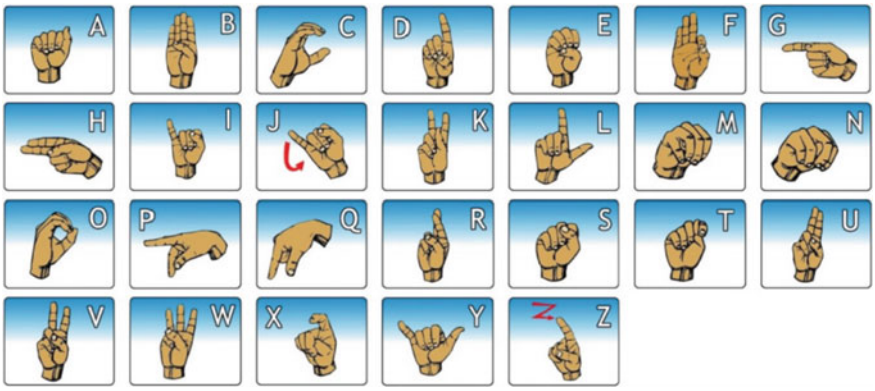



**Fig. 2** ASL fingerspelling dataset [20]

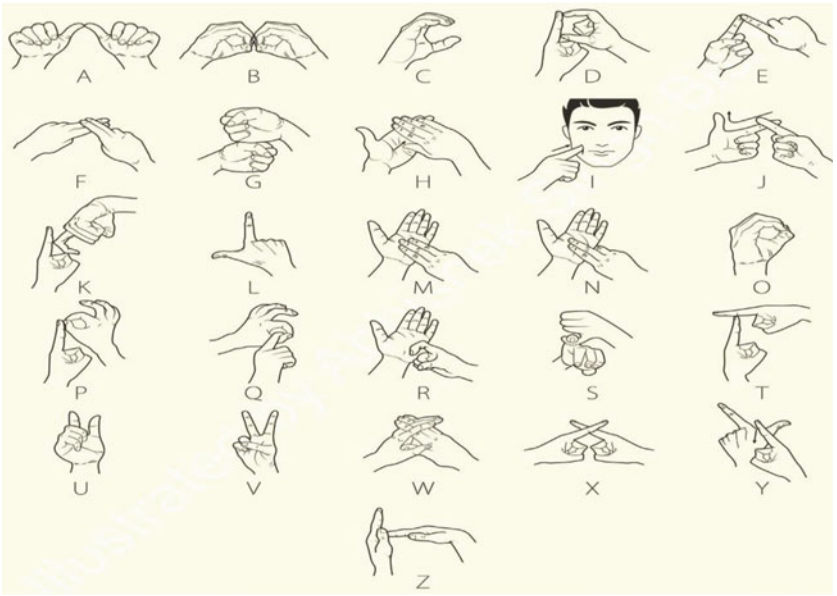



**Fig. 3** ISL fingerspelling dataset [21]

## 4 Sign Language Recognition System Overview

Gestures in sign language are a combination of the following elements:

- Articulation points (finger joints, wrist, elbow, shoulder),
- hand/palm orientation,
- Facial expressions,

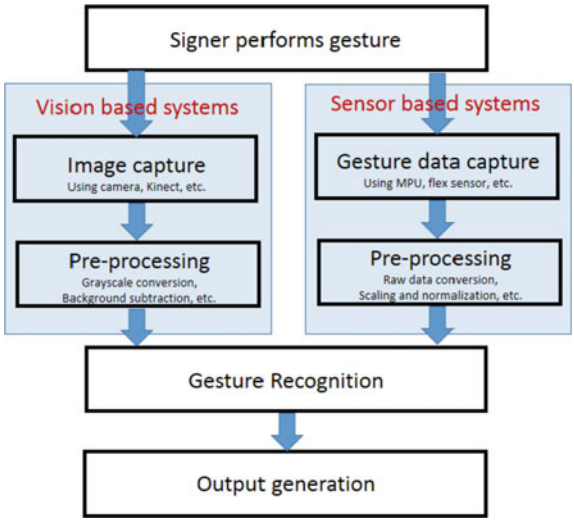



**Fig. 4** Overview of the sign language recognition process

- Motion (of fingers, hands, and head).

The process of SLR can be divided into the following stages as shown in Fig. 4:

## *4.1 Data Acquisition*

To detect the motion, position, and orientation of hands, fingers, etc., the available systems can be predominantly categorized into two classes: sensor-based and vision-based.

The vision-based systems make use of a camera unit directed toward the subject and captures images or videos as the person gestures in sign language. The camera can capture RGB or monochrome image sensors, as well as IR image sensors also.

Sensor-based systems majorly make use of some or all of Flex sensors, IMU, Pressure/contact sensors, sEMG sensors, etc. attached to a glove.

## *4.2 Pre-processing and Features Extraction*

The elements of sign language gesture are acquired in this step. In vision-based systems, the captured image is pre-processed to reduce noise and improve image characteristics to make recognition easier. It may include background subtraction, image segmentation, subject isolation, noise removal (data cleaning), face detection and facial expression detection, grayscale conversion, binarization, tracking of hand movements for dynamic gestures, etc.

For the vision-based systems, lighting conditions, camera resolution, frame rate (for dynamic signs), distance from the subject, skin color and background color play a significant role in sign recognition accuracy.

In sensor-based systems, data from the appropriate sensors are collected in raw form and some pre-processing is applied on them to make the data more suitable for use, e.g., orientation data from IMU can vary from a few hundred to thousands [22], however, the desired output data should be in the range from 0 to 360 indicating the orientation angle; also, different sensors have different baud-rate, which must be factored in during the pre-processing stage.

### *4.3 Recognition*

In this step, the features extracted in the previous section are matched against a pre-defined dataset or passed on to a machine learning classification model (which has been prior trained with reference data) to recognize the signs. The output of this step can be in form of audio from a speaker, or text display on a screen.

## 5 Discussion

Flex sensors provide data about flexion of the fingers, however, they do not describe the relative position of the fingers, due to this, using only flex sensors only a small number of signs can be differentiated due to the limited resolution/range of the sensors and difficulties in differentiating similar signs such as 'M' and 'N'. PPG sensors [4] suffer from body motion artifacts and hence cannot be employed in a practical device. Another way of recognizing gestures is using an EMG sensor [3, 7, 9], however, an EMG sensor alone cannot differentiate signs having similar hand orientation and different position like 'mother' and 'father' signs in children's sign language have the exact same hand orientation and differ only in position. The inclusion of IMU [4, 6–9] sensors enables for better differentiation of dynamic signs, however, the problem with similar static signs as in the previous case is still present. Pressure/contact sensors provide data about whether the fingers touch each other, and how firmly, which can be used for discerning many similar gestures and improve the accuracy of the system. Also, as only a few [7] of the aforementioned systems have used a large dataset of other than basic signs (usually 26 alphabets, 10 numbers, and some static gestures), it is difficult to make a statement about the scalability of the other sensor-based systems, because as the number of signs increases, it gets more likely for the signs to be similar to each other and interfere in accurately recognizing the signs.

In the vision-based systems also, apart from [15] who used a dataset for 249 signs for their system and [17] whose dataset contains 71 different signs, all other systems have been tested for a significantly fewer number of signs and hence it is difficult

to comment about their performance with a significantly large dataset for the same reasons as in the case of sensor-based systems.

## 6 Conclusions

We have seen in the previous sections that most of the systems are merely a proof of concept and not a real solution to the problem, as scalability remains a major concern. For the sensor-based systems, the accuracy of the gesture recognition increases as the number and type of sensors used is increased, hence, future research should continue to explore the accuracy of SLR by using flex sensors on the wrist of the signer also, to capture the relative angle of the hand with respect to the arm, along with using flex sensors on all 10 fingers, and IMUs on each hand to capture as much data as possible to make a decision. The inclusion of contact sensors and sEMG sensors will only improve the accuracy of the system.

Also, it can be concluded from the previous sections that vision-based systems cover the areas missed by sensor-based systems, like facial expressions, and sensor-based systems do not need perfect line of sight communication, unlike vision-based systems. This can be extremely useful in discerning signs that are similar to each other. Hence, it will be fascinating to see the two methods combined to make a hybrid SLR system, which should eliminate most of the drawbacks of the current solutions and can make for a real-world solution to sign language recognition, however, the cost constraint of such a system might pose as a roadblock for availability to the masses.

## References

1. Who.int. Deafness and hearing loss. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss. Accessed 26 Apr 2021
2. Censusindia.gov.in. Census of India: Disabled Population. https://censusindia.gov.in/census_and_you/disabled_population.aspx. Accessed 26 Apr 2021
3. Zhang Z, Su Z, Yang G (2019) Real-time Chinese Sign Language Recognition based on artificial neural networks. In: 2019 IEEE international conference on robotics and biomimetics (ROBIO). IEEE, pp 1413–1417
4. Zhao T, Liu J, Wang Y, Liu H, Chen Y (2019) Towards low-cost sign language gesture recognition leveraging wearables. IEEE Trans Mob Comput
5. Suri K, Gupta R (2019) Continuous sign language recognition from wearable IMUs using deep capsule networks and game theory. Comput Electr Eng 78:493–503
6. Saggio G, Cavallo P, Ricci M, Errico V, Zea J, Benalcázar ME (2020) Sign language recognition using wearable electronics: implementing k-nearest neighbors with dynamic time warping and convolutional neural network algorithms. Sensors 20(14):3879
7. Yu Y, Chen X, Cao S, Zhang X, Chen X (2019) Exploration of Chinese sign language recognition using wearable sensors based on deep belief net. IEEE J Biomed Health Inform 24(5):1310–1320

8. Chong TW, Kim BJ (2020) American sign language recognition system using wearable sensors with deep learning approach. J Korea Inst Electron Commun Sci 15(2):291–298
9. Fatmi R, Rashad S, Integlia R (2019) Comparing ANN, SVM, and HMM based machine learning methods for American sign language recognition using wearable motion sensors. In: 2019 IEEE 9th annual computing and communication workshop and conference (CCWC). IEEE, pp 0290–0297
10. Daniels S, Suciati N, Fathichah C (2021) Indonesian sign language recognition using YOLO method. In: IOP Conference Series: Materials Science and Engineering, vol 1077, no 1. IOP Publishing, p 012029
11. Wang CC, Chiu CT, Huang CT, Ding YC, Wang LW (2020) Fast and accurate embedded DCNN for RGB-D based sign language recognition. In: ICASSP 2020–2020 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 1568–1572
12. Saleh Y, Issa G (2020) Arabic sign language recognition through deep neural networks fine-tuning. Int Assoc Online Eng. 71–83
13. Vaitkevičius A, Taroza M, Blažauskas T, Damaševičius R, Maskeliūnas R, Woźniak M (2019) Recognition of American sign language gestures in a virtual reality using leap motion. Appl Sci 9(3):445
14. Gangrade J, Bharti J (2019) Real time sign language recognition using depth sensor. Int J Comput Vis Robot 9(4):329–339
15. Sarhan N, Frintrop S (2020) Transfer learning for videos: from action recognition to sign language recognition. In: 2020 IEEE international conference on image processing (ICIP). IEEE, pp 1811–1815
16. Mali D, Limkar N, Mali S (2019) Indian sign language recognition using SVM classifier. In: Proceedings of international conference on communication and information processing (ICCIP)
17. 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 9(6):821–827
18. Wadhawan A, Kumar P (2019) Sign language recognition systems: a decade systematic literature review. Archives of computational methods in engineering
19. Ethnologue. Sign language. https://www.ethnologue.com/subgroups/sign-language. Accessed 5 June 2021
20. American Sign Language (ASL) Discussion Board. ASL Fingerspelling Alphabet.https://www.fingerspellingalphabet.com/. Accessed 5 June 2021
21. Islrtc.nic.in. Poster of the Manual Alphabet in ISL | Indian Sign Language Research and Training Center (ISLRTC), Government of India. http://www.islrtc.nic.in/poster-manual-alphabet-isl. Accessed 26 Apr 2021
22. Arduino Forum. Understanding MPU-6050 register raw data? Gyro + Accelerometer. https://forum.arduino.cc/index.php?topic=643304.0. Accessed 26 Apr 2021
23. Amazon.com. https://www.amazon.com/Thalmic-Labs-Gesture-Control-Presentations/dp/B00VHWBH02. Accessed 4 June 2021
24. Flipkart.com. MICROSOFT Kinect Sensor for Xbox 360 Motion Controller - MICROSOFT: Flipkart.com. https://www.flipkart.com/microsoft-kinect-sensor-xbox-360-motion-controller/p/itmew4fyxx3gpypn. Accessed 4 June 2021
25. Ultraleap.com. Tracking | Leap Motion Controller | Ultraleap. https://www.ultraleap.com/product/leap-motion-controller/. Accessed 4 June 2021

# ACS Fed Coplanar Monopole Antenna with Complementary Split Ring Resonator for WLAN and Satellite Communication Applications


**Ch. Ramakrishna and Bandi Geervani**

**Abstract** An antenna designed for the two-band purpose is deliberated with a compact asymmetrical co-planner. The novelty of the work is asymmetric coplanar feeding with tapered radiating patch for reducing antenna dimension by incorporating with the defected structure together. The rectangular complementary split ring resonator is located on the ground plane and the radiating patch is accommodated on the substrate with FR4 epoxy with a dielectric constant of 4.4 designed and aimed at two-band setup. The dimensions of the proposed antenna 17 × 7.75 $mm^2$ is designed and verified. The proposed ACS feed monopole antenna E-plane and H-plane radiation patterns are simulated at 5, 8 and 10 GHz correspondingly. Simulated results showed that the proposed antenna takes a 1 GHz (5.3–6.3 GHz), 3.3 GHz (7–10.3 GHz) impedance bandwidths which are essential aimed at a WLAN and for Satellite communications applications entire antenna bandwidth the S-parameters and VSWR are below −10 dB and less than 2.

**Keywords** Asymmetrical coplanar strip · Compact size · Defective structure · Impedance bandwidth · Microstrip line · Rectangular complementary ring resonator

## 1 Introduction

In the field of radio technology, the development of antennas for various wireless applications has taken on great importance. A compact double-band antenna where the antenna has metamaterial characteristics and defective ground structure is incorporated is demonstrated. It can be used by means of a CSRR [1] structure, which delivers the metamaterial property, for mobile applications. In micro strip line feed, effects can be analyzed that can be transported to an antenna using the CSRR. A study be situated conducted also is well described in this work on the CSRR in the ground

Ch. Ramakrishna (✉) · B. Geervani
Vardhaman College of Engineering, ECE, JNTU Hyderabad, Hyderabad, India
e-mail: ramakrishna@vardhaman.org



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_12

plane. In a U-shaped antenna [2] which can be used for certain band applications was discussed.

The effect of the defective ground structure and the influence of a CSRR were discussed in and appearances the efficient miniaturization method used for wireless applications with enriched fractional bandwidth [3] also minimum reflection coefficient were presented in the bottom plane. For the four-band of antenna is presented in with an octagonal shaped radiation element. This antenna is used for the compact and more impedance matching applications of C, L, X and Ku band. With the intention of decrease, the complete size of the antenna, ACS was established. In, a small arrangement for mobile devices, an antenna through the Chinese tai-shaped patch is established. There is discussed in a compact uniplanar antenna [4] composed of L shaped patch and bottom plane. In a small antenna through reversed L slot in addition to L shaped radiation patch, used for wireless and WIMAX [5] applications. A small Asymmetrical coplanar strip [6] antenna feeding with an L type mirrored radiation component, encumbered with the ring resonator mentioned in. In antennas with another alphabetic radiant element are designed and developed for good impedance and enhanced bandwidth. In for triple-band [7] applications, antennas with two step slits in open-ended shape were presented. The antenna takes a constant gain on desired frequency bands, making it ideal for wireless use. In, a miniaturized asymmetric coplanar antenna [8] resonates with WLAN and WLAN frequencies. In, a compact and reconfigurable antenna, motivated by metamaterial for antenna reconfiguration in order to operate on the required frequency bands, is presented. The ACS feed antenna is presented as two polarized stubs loaded with the frequency adjustable, consisting of L-stubs, U stubs and U-slot [9] allied to the PIN diodes located in various locations, which are used for reconfiguration of the antenna. In, a small design contains a coplanar [10], circular and SSRR [11], formed by patch loading with a microstrip line; in antenna with the tapered patch was discussed in addition to the ground plane with a CSRR. Numerical simulations in proposed an antenna with a WiMAX [12] application through an enhanced gain. The antenna can be built into several small devices due to its compact size. A miniaturized version of a U-shaped antenna is available in that have a 140 MHz frequency band. A low-profile planar antenna [13], which offers acceptable performance in operating bands, is discussed in for WIMAX, Wi-Fi, RFID, Bluetooth, and other wireless Communication Applications.

## 2 Related Works

A rectangular patch, a rectangle ground and a feeding network are the antenna. On two sides of the single feed network was the folding microstrip line and the coplanar waveguide [14] (CPW) structure. Two slots of the new CPW hence are able to create electric fields in and out of phase correspondingly to enable the patch to be stimulated in the uniform and odd mode of the lower common strip and in the upper common strip higher order modes. For a compact satellite payload communication system,

a Compact, two-band, circular, polarized antenna was proposed. The antenna has a pentagonal radiator [15] and a stacked rectangular parasite element for a lower and greater axial ratio. The antenna shows a bandwidth of 380 MHz (2.34 to 2.82 GHz) and a bandwidth of 770 MHz (3.05 to 3.83 GHz) per 10 dB impedance, with a small size of 0.33 pound 0:33 pound 0:03 pound The Antenna shows two (with respect to lower end frequency). The article proposed a flat, low-profile, dual-band and double-polar antenna [16] on a semi-substrate. The antenna is produced on a 3.04 mm thick Rogers substrate. The suggested antenna may be employed in an indoor-outdoor wearable device with this simultaneous, dual-band and dual polarization operation.

The work contains a wearable textile antenna with dual-band and dual-sense features. It works with a WBAN and WLAN band of 2.45 GHz and with a GPS band of 1.575 GHz. This may be achieved in the field. It is composed of a 3–3 square patch cell array, where each cell has a unit Four square slots and a square ring [17] were combined. On top of the substrate as its radiator is then a square patch. Two corners of this radiator are truncated to facilitate dual band operation and all four angles are included with a rectangular slit to permit its circular polarization in the GPS band.

In order to improve antenna radiation characteristics, the non-existing qualities of metamaterial surfaces can be used. In this study, a design and performance study for LTE 46/WLAN and Ka-band applications is conducted on the basis of a Single Negative (SNG). The antenna based on the SNG [18] proposal spans the 0.35–5.69 GHz (LTE 46/WLAN) and 17.81–20.67 GHz bandwidth = 1.00 dB. (Ka-band). The antenna element's size is 20.2/28.4 $mm^2$, while the antenna element has a 2/3 SNG metamaterial surface that increases gain.

As the trend of reduction of devices is in full swing, antennas cannot stay as separate devices. Compact designs must be applied to meet industrial requirements. The designers' major objective is the radiation patch. Multiple processes such as folding, meandering etc. usually reduce the total antenna size. While designers emphasise the radiator, even the most compact designs are left undisturbed in the feeding region. The antenna's entire area relies on both radiation and feed size. The thesis aims to create tiny antennas by an effective minimization of both feed and the radiation structure. The overall design complexity of the antenna also needs to be reduced for its effective use in a device. In comparison with a coplanar waveguide fed monopoly, an in depth analysis will be conducted utilizing an ACS fed monopole. Following thorough simulation and experimental research, all the benefits of the CPW-fed antenna plus additional smallness remain with an ACS-fed antenna. A double-band antenna with the F shape and an inverted C shaped multi segment antenna [19] are constructed and researched to demonstrate the ability to use ACS as an effective feed for dual-band and multi-band antennas. All studies are successful and promising.

The problem is to create ultra-compact broadband antennas meet operators' standards. In literature several existing strategies have been offered, for example the usage of metamaterials, PBGs, fractals, meandering etc. All of the following strategies increase the complexity, expense or size of the equipment, making it virtually difficult to implement. The problem has been thoroughly investigated by employing the antenna space efficiently and totally.

## 3 Method

This paper focuses primarily on a radiating element for the two band operation through the small ACS feed. The primary of the defective construction together through asymmetrical coplanar strip dissipation effectively decreases the antenna structure. The complementary coplanar strip is on the bottom plane and CSRR on the patch is etched to achieve two-band setup and removal of existing wireless communications. Defaults on the bottom plane are called the defective ground structure (DGS). The DGS techniques discussed in [20] improve various parameters of the antenna and miniaturization. The CRR can have the same effects as CSRRs is observed in this work. The antenna geometry and simulation findings are discussed in the following section.

### *3.1 Design and Development of the Proposed Antenna*

Figure 1 shows the size of the antenna that is proposed with two bands inspired by a ground plane CSRR defect and I-slit on the radiating patch. Table 1 shows the optimized set of measures.

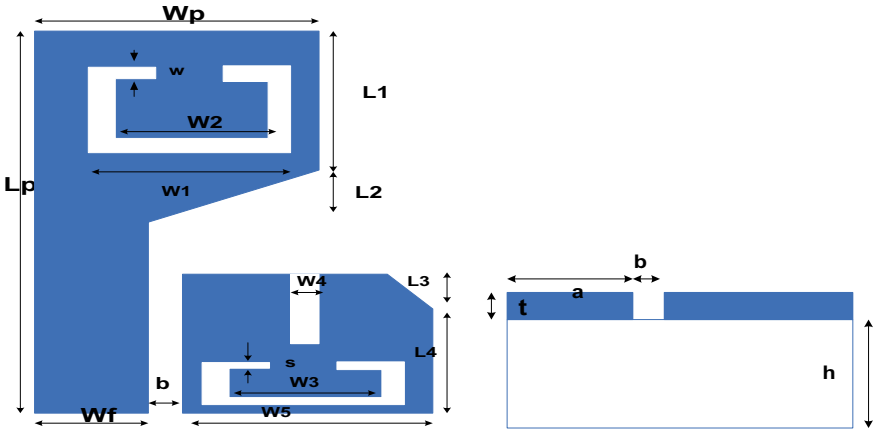



**Fig. 1** Structure of the proposed antenna **a** Top view **b** Side view

**Table 1** Parameters of the proposed antenna

| Parameter | Lp | Wp | Wg | Wf = a | W1 | W2 | L1 | L2 | W3 | W4 | W5 | w | L3 | L4 | h | b |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Values (mm) | 17 | 6.95 | 3.9 | 3.6 | 2.8 | 1.8 | 6.6 | 3.8 | 0.5 | 1.8 | 2.8 | 0.5 | 4.7 | 1 | 1.6 | 0.25 |

The antenna proposed has a total size of 17 × 7.75 mm placed on a FR4 epoxy $\varepsilon_r = 4.4$ and h = 1.6 mm in height. The longer the wavelength is, the longer the antenna must be adequately matched, because the wavelength is inversely proportionate to the frequency. A longer antenna is more theoretically, but it is far more necessary that the length of an antenna fits the frequency when it comes to best antenna performance. The measurements of the offered ground plane are enhanced for better impedance.

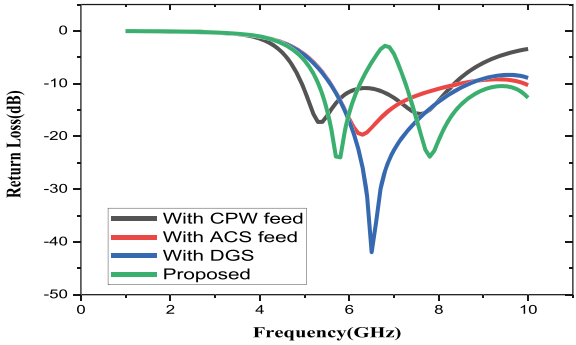
## 4 Results and Discussion

The Figs. 1 and 2 displays the progress of the antenna proposed for two band applications, respectively, in the development of its reflection coefficient and the antenna vswr is less than 2 at operating frequency bands. The design began with the basic configuration CPW fed which provides a large 5 GHz resonance mode. Basic mode match was poor and improved by changing the feed to ACS. ACS feed line is characterized by impedance. The simulated results of the proposed two-band antenna are shown in Figs. 3 and 4.

Figures 5 and 6 shows parametric studies of the projected structure aimed at various sizes of complementary split ring resonator defects. By changing the CRR's slot width on the ground, which has an impact on impedance bandwidth and gains

**Fig. 2** Design development of the ACS Monopole antenna

**Fig. 3** Return loss of the design evolution



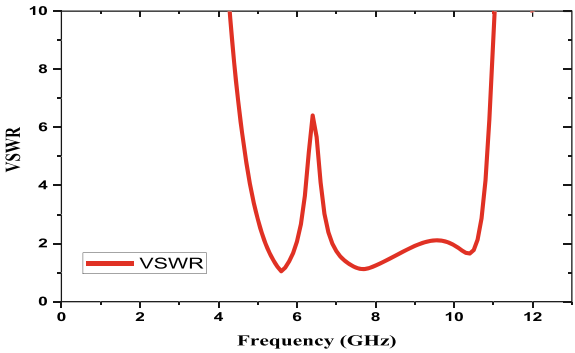



**Fig. 4** VSWR characteristics of the proposed antenna

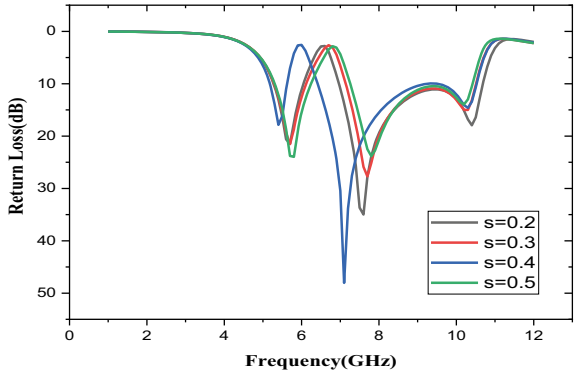



**Fig. 5** Return loss characteristics by varying the slot width 's'

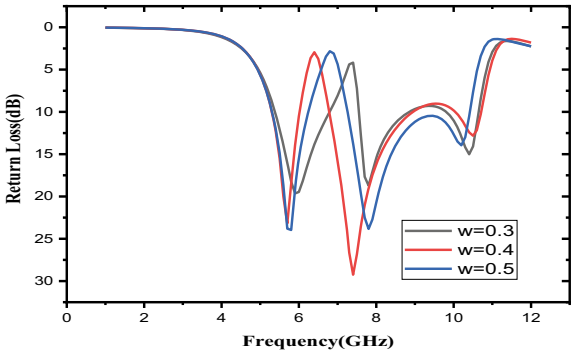



**Fig. 6** Return loss characteristics by varying the slot width 'w'

the resonance band at 5 GHz. Optimal values are achieved if s = 0.5 mm and other sizes are sufficient and resonances be apt to move to additional frequencies. This results in the lowest resonance with better matching mode by creating a rectangular defect known as an I-slit. The parametrical study illustrations that through changing the slot width 'w' the correspondence of the basic operation in the lower order and that the optimized results for w = 0.5 mm have been achieved.

In Fig. 7, the simulated 2D E and H-principle radiation patterns for 5, 8 and 10 GHz are shown. This indicates that the omnidirectional and bidirectional radiation patterns in H and E planes. Irregularity in the offered antenna feeding alignment is responsible for the small asymmetry in the patterns.

The current distribution in the offered antenna in lieu of the three resonance modes shown in Fig. 8. The low resonance mode i.e. the current of 5 GHz is more disturbing in the radiation strip in the opposite direction the I-slit. The Table 2 presents the

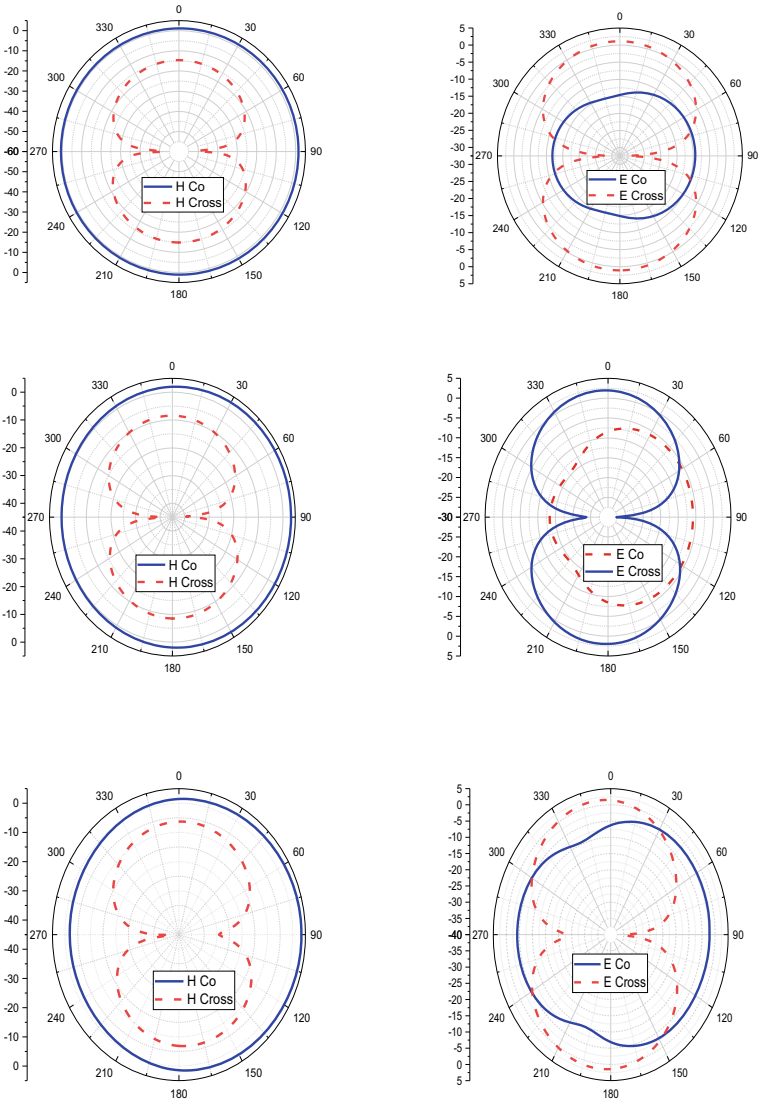



**Fig. 7** Simulated E-plane and H- plane radiation patterns at **a** 5 GHz **b** 8 GHz **c** 10 GHz

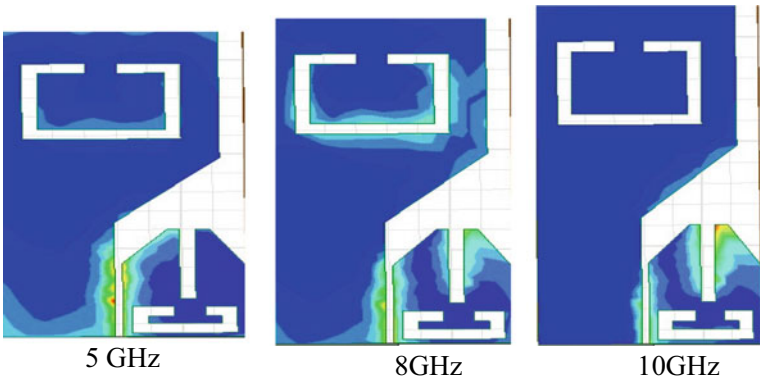



**Fig. 8** Current distributions on proposed antenna

**Table 2** Comparison between the performances of proposed and existing antennas

| S.No | Reference | Antenna size | Resonance frequency (GHz) | Total Area ($mm^2$) | Purpose of antenna | Bandwidth (MHz) | Isolation (dB) | Gain (dBi) |
|---|---|---|---|---|---|---|---|---|
| 1 | [1] | 20.2 × 28.4 | 5.35–5.69 17.81–20.67 | 573.68 | Dual band | 340/2860 | <-32 | 5.04 |
| 2 | [2] | 40 × 40 | 1.575/2.45 | 1600 | Dual band | 120/172 | <-10 | 1.94 |
| 3 | [3] | 37 × 40 | 1.563- 1.594 2.434–2.451 | 1480 | Dual band | 31/17 | <-10 | 3 |
| 4 | [4] | 33 × 33 | 2.44–2.82 3.05–3.83 | 1089 | Dual band | 380/770 | <-10 | 4.4 |
| 5 | [5] | 29 × 12 | 2.453–2.821 5.876–6.892 | 348 | Dual band | 368/1016 | <-20 | 4 |
| 6 | Proposed | 17 × 7.75 | 5.5/7.5 | 131.75 | Double band | 1000/3300 | −45/−18 | 4.1 |

relative study of the antennas and size, operating frequency, volume, S11, fractional bandwidth and the use of the radiating element, made on the existing antennas.

The dipole radiates in phase or in synchronization by providing an equal amount of power that reaches each dipole at the same moment for an increased gain by virtue of its design. Figure 9 compares the gain of the antenna being offered in lieu of defected structures with and without a CRR structure. It is clear from the figure that the antenna's gain is enhanced by etching the CRR defect to the side of the ground and providing a maximum gain of approximately 2.5 dBi in the upper resonance mode.

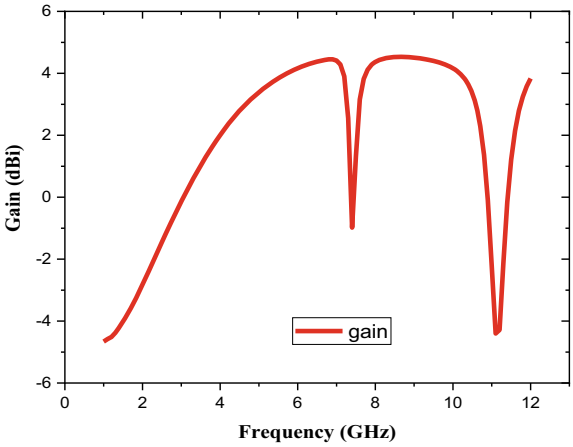



**Fig. 9** Gain performance of the ACS antenna

## 5 Conclusion

A novel antenna design designed for the achievement of desired frequencies, the Asymmetrical coplanar strip fed and complementary split ring resonator features are presented. This antenna works on two frequency bands named 5.7 and 8.6 GHz. The structure of the antenna is imported into the FR4 substrate and measures 17 × 7.75 $mm^2$. The antenna remains changed to work on WLAN and Satellite communication appliances by incorporating CSRR resonators in Asymmetrical coplanar strip structures. The simulations and parameter studies are performed using the IE3D electromagnetic solver method based on MOM. Ansys' HFSS extracts the dual negative property of metamaterials. Simulated results have shown that can be finest used for multiband wireless applications.

## 6 Future Scope

The loading of metamaterial is suitable to decrease the radiation towards the user's head when employing the antenna on practical mobile telephones. In other words, new and intriguing venues for designing simple Miniature Antennas for a designer are introduced with the ACS feed antennas, which are cheap and easy to fabricate.

## References

1. Asif M et al (2021) Design of a dual band SNG metamaterial based antenna for LTE 46/WLAN and Ka-band applications. IEEE Access 9:71553–71562. https://doi.org/10.1109/ACCESS.2021.3077844

2. Joshi R et al (2020) Dual-band, dual-sense textile antenna with AMC backing for localization using GPS and WBAN/WLAN. IEEE Access 8:89468–89478. https://doi.org/10.1109/ACCESS.2020.2993371
3. Paracha KN et al (2019) A low profile, dual-band, dual polarized antenna for indoor/outdoor wearable application. IEEE Access 7:33277–33288. https://doi.org/10.1109/ACCESS.2019.2894330
4. Alam T, Islam MT (2018) A dual-band antenna with dual-circular polarization for nanosatellite payload application. IEEE Access 6:78521–78529. https://doi.org/10.1109/ACCESS.2018.2885067
5. Liu X, Wu Y, Zhuang Z, Wang W, Liu Y (2018) A dual-band patch antenna for pattern diversity application. IEEE Access 6:51986–51993. https://doi.org/10.1109/ACCESS.2018.2870142
6. Jan NA, Saleh AM, Lashab M, Abdussalam FM, Djouablia L, Abd-Alhameed RA (2017) A compact CSRR loaded monopole antenna with defected ground structure for mobile WLAN and WiMAX applications. In: Otung I, Pillai P, Eleftherakis G, Giambene G (eds) Wireless and satellite systems. WiSATS 2016. Lecture notes of the institute for computer sciences, social informatics and telecommunications engineering, vol 186. Springer, Cham
7. Kunwar A, Gautam AK, Rambabu K (2017) Design of a compact Ushaped slot triple band antenna for WLAN/WiMAX applications. AEU Int J Electron Commun 71:82–88. ISSN 1434-8411
8. Kumar MN, Yogaprasad K, Anitha VR (2019) A quad-band Sierpenski based fractal antenna fed by co-planar waveguide. Microw Opt Technol Lett 1–6
9. Niu X, Liang J, Wu G, Lin Y (2016) Compact dual-band ACSfed monopole omnidirectional antenna for WLAN/WiMAX applications. In: 2016 progress in electromagnetic research symposium (PIERS), Shanghai, pp 2014–2018
10. Arvind K, Vummadisetty N, Praveen, Vinay K (2017) A compact uniplanar ACS fed multi band low cost printed antenna for modern 2.4/3.5/5 GHz applications. Microsyst Technol 24
11. Rajkumar R, Kiran KU (2016) A compact ACS-fed mirrored L-shaped monopole antenna with SRR loaded for multiband operation. Progr Electromagnet Res C 64:159–167. 9p.
12. Naidu PV (2017) Printed V-shape ACS-fed compact dual band antenna for bluetooth, LTE and WLAN/WiMAX applications. Microsyst Technol 23:1005–1015
13. Naidu PV, Kumar A, Kumar V (2018) Uniplanar Y and L shaped ACS fed multiband and wideband compact printed antenna for advanced wireless communication systems. Microsyst Technol 24:2051–2061
14. Vummadisetty PN, Kumar A (2018) Multi feed multi band uniplanar ACS fed antenna with N shape and inverted L shape radiating branches for wireless applications. Microsyst Technol 24:1863–1873
15. Hodoodi M, Majidzadeh M, Ghobadi C, Mohammadi B, Nourinia J (2019) A novel compact triple band ACS-Fed antenna. In: 2019 27th Iranian conference on electrical engineering (ICEE), Yazd, Iran, pp 1606–1610
16. Naik K (2018) Asymmetric, CPW-fed SRR patch antenna for WLAN/WiMAX applications. AEU-Int J Electron Commun 93:103–108
17. Rajeshkumar V, Rengasamy R, Naidu PV, Kumar A (2019) A compact meta-atom loaded asymmetric coplanar strip-fed monopole antenna for multiband operation. AEU Int J Electron Commun 98: 241–247. ISSN 1434-8411
18. Morshed KM, Esselle KP, Heimlich M, Mueck MD, Hay SG (2015) A simple planar monopole antenna for mobile handset applications. In: International symposium on antennas and propagation (ISAP)
19. Morshed K, Karmokar D, Esselle K (2017) Low profile single-layer Uslot loaded shorted-patch antenna for wireless communications: MORSHED et al. Microwave Opt Technol Lett 59:2224–2226. https://doi.org/10.1002/mop.30712
20. Khandelwal MK, Kanaujia BK, Kumar S (2017) Defected ground structure: fundamentals, analysis, and applications in modern wireless trends. Int J Antennas Propag 2017:1–22

# Advance the Energy Usage in Cloud Centers Utilizing Hybrid Approach


**D. Jayakumar and Talluri Lakshmi Siva Rama Krishna**

**Abstract** Cloud Computing is one of the new arising advances that offers types of assistance to buyers in compensation as you go model. Distributed computing offers ITC based administrations over the web and the utilization of virtualization permits it to give figuring assets. Server farms are the center of cloud processing, which comprises of: arranged workers, links, power sources, and so forth which have the running applications and store Business data. Elite has consistently been the most basic worry in cloud server farms, which comes at the cost of energy utilization. The imperative test is adjusting between framework execution and force utilization by lessening energy utilization without biased effect on the execution and nature of administrations conveyed. There are numerous procedures and calculations proposed to accomplish effective energy usage in distributed computing, these methods include: VM Relocation, Consolidation and Resources coordination in cloud figuring. This paper gives an overview of approaches and procedures for energy effectiveness in Cloud computing.

**Keywords** Cloud registering · Data center · Energy effectiveness · Energy saving and virtualization

## 1 Introduction

The advancement of innovation and joining organizations, capacity and handling power prompted new time of processing, called distributed computing or usually known as the cloud. Distributed computing is characterized as a mechanical world-view that permits on-request access through the web to a typical common registering assets. It is viewed as a model for management, putting away and preparing information online through the web [1]. Some distributed computing attributes incorporate

D. Jayakumar (✉) · T. L. S. R. Krishna
Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India
e-mail: djayakumar@kluniversity.in

T. L. S. R. Krishna
e-mail: sivamca.mtech@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_13

on-request benefits, network access by utilizing web as a medium, shared assets by pooling assets together to be utilized by different customers and adaptability by keeping up versatility of assets. Distributed computing offers extraordinary administrations dependent on three conveyance models; (1) Software as an assistance (SaaS): It permits client of cloud to access the supplier's applications (PA) over the web. (2) Platform as a Service (PaaS): It permits clients to send their applications on a stage which specialist co-op of cloud (SPC) gives. (3) Infrastructure as a Service (IaaS): It permits clients to lease, store, and measure in a framework given by SPC. In this paper, we endeavor to limit make span and energy utilization in light of the significance factor dictated by the end user and the supplier, by the use of asset allocation in a line made by four heuristics strategies and the grasshopper calculation. Most fitting strategy to carry out each assignment is then picked utilizing the roulette wheel. To tackle the above issues, this paper makes five primary commitments:

- Energy-mindful run-time scheduler for booking of work process.
- Provides a procedure to naturally create the necessary force utilization profile.
- Selects the best arranging strategy to play out any undertaking between a heuristic strategy and grasshopper algorithm to build usefulness.
- Offers a model to assess the energy utilization also, makespan of the assignment.
- Evaluates the compromise between energy saving and makespan for various situations.

To meet the quick changing business and association needs, associations need to give financial plan and time to speed up their IT foundation like programming, equipment and organization administrations. Despite the usage of on location IT structure, scaling the framework could be troublesome. And furthermore the associations are regularly unequipped for accomplishing an ideal utilization of IT establishment. Accordingly, the distributed computing is the proposed arrangement. As per National Institute of Guidelines and Technology (NIST), distributed computing is the conveyance of IT assets on-request use by giving a pay more only as costs arise model for the buyers, while you can self-serve for the administrations that you need to your own application or any IT foundation that you need [2]. A cloud processing administration comprises of exceptionally used assets counting programming applications or virtual stockpiles that can be utilized upon client demand, buyers can just associate with the cloud and utilize the accessible assets. This causes associations to avoid capital utilization for on-premises system resources and increasing or cutting back as indicated by business necessities [3]. Distributed computing administrations can be conveyed utilizing three unique models a private cloud, public cloud or a half breed cloud. Private cloud work exclusively for one association on a private organization and is its exceptionally secure. Public cloud is possessed by the cloud specialist co-op and offers the most elevated level of proficiency and shared assets and half and half cloud is viewed as a mix of private and public sending models. In a half breed cloud, explicit assets are run or utilized in a public cloud and others are run or utilized on-premises in a private cloud this gives expanded effectiveness.

Data center offer an IT spine to distributed computing. A Data center is a specialized office that houses associations IT activities and hardware where it stores, oversees and spreads its information. A Data center houses and organizations most basic frameworks and are fundamental to business congruity and activities. A Data community is viewed as the core of distributed computing which contains all the cloud assets counting workers, organization, links, and so forth on which business data is put away and applications run. As of not long ago, elite has been the sole worry in server farm dispersions and this interest has been fulfilled without giving a lot of consideration to energy utilization. In any case, a normal Data center devours as much as the utilization of 25,000 houses. As the energy accessibility diminishes and energy cost relatively expands the requirement for moving the center for using server farm asset the board to upgrade energy execution while looking after assistance execution is turning into a need. In this manner cloud administration suppliers need to change their energy measures to guarantee that their overall revenue isn't drastically diminished because of high energy costs [4].

## 2 Related Works

Today, distributed computing assumes a fundamental part in scholarly world furthermore, industry. Be that as it may, because of high client interest furthermore, restricted assets; it is important to move a little measure of work to Data centers. Since there are such countless errands on cloud workers and assets in the cloud are heterogeneous, planning errands in a cloud climate is a steady test. In this part, we momentarily depict some current exploration in the field of booking on the subject of dispensing assets to assignments in cloud conditions. Stavrinides and Karatza [5] presented a time and energy-mindful booking calculation called EATSDCD. presents a technique that uses a blend of two bunching and duplication methodologies and the utilization of slack time, to plan DAGs on server farm processors with dynamic voltage and recurrence scaling capacity. In the first stage, by astutely joining the two systems of duplication and bunching, the emphasis is set on diminishing the makespan and energy devoured by processors with an end goal to execute DAG while fulfilling the throughput limitation. After ascertaining the leeway time for each undertaking and afterward deciding the basic way, the leeway time is dispersed between the set of noncritical ward undertakings in a group, and the recurrence of processors with dynamic scale adaptability at that point performs the reliant errand set. The recurrence of DVFS-enabled processors is downsized to execute noncritical errands just as inactive and correspondence stages, without increasing the makespan of the assignments. Juarez et al. [4] introduced an energy-mindful planning calculation for task-based applications called the multiheuristic asset designation calculation (MHRA). The objective of this planning algorithm is to limit the capacity of the objective that consolidates energy utilization and makespan. These models are joined with a significance factor decided end-clients furthermore, specialist co-ops to characterize which is more significant for their destinations: energy saving or makespan. The creators moreover proposed

a model for assessing the energy utilization needed to execute productively applications on set of resources. Zhu et al. [7] explored the energy aware plan calculation for a work process with restricted time stretches in a homogeneous cloud climate. they were ready to diminish makespan, especially when the quantity of errands was huge. The fundamental goal was NSGA-II. In request to control the energy utilization adequately, the DVFS framework was fused into the improvement procedure, and a bunch of non-domination arrangements are gotten also, utilized. Moreover, the fake neural organization (ANN), which is a fruitful AI calculation, was utilized in light of the highlights of the assets accessible in it. Mandal et al. [11] proposed a productive metaheuristic approach called energy-productive work process planning (EEWS) for minimizing makespan and expanding energy preservation while booking work processes. (e aftereffects of EEWS were better than both HCRO and MPQGA, as far as makespan, moderated energy, and wellness esteem. Belgacem et al. proposed a unique asset designation model to meet client interest for assets with improved responsiveness. Barzegar et al. [12] proposed a multi objective pursuit calculation called the separating multi objective subterranean insect lion algorithm (S-MOAL) to limit both the makespan and the cost of utilizing virtual machines. Its effect on energy consumption with the unique administration of assets was concentrated using unused virtual machines (VMs), which demonstrated that their methodology would save energy.

## 3 Energy Efficiency Computing

Energy saving strategies in figuring gear have been delegated static force the executives (SPM) and Dynamic force the executives (DPM). SPM and DPM are totally unique in order, SPM are more energy effective at single framework and expected to be under the class of equipment level strategies, and since SPM strategies are identified with equipment level effectiveness, low power utilization circuit planning is an illustration of this strategy. On the opposite side, DPM are more energy proficient in enormous frameworks and expected to be under the classification of level asset the executive's techniques. Additionally, DPM procedures are generally executed in programming or on network layer, for model convention plan and calculations [6].

Energy mindful planning, energy effective directing, load adjusting, virtualization, asset union and relocation. Since high accessibility also as nature of administration and execution ensure are still overlooked which is generally needed in such appropriated conditions as the clients pay for their provisioned assets. The clients would not compensation or may change to other comparable specialist organizations if either nature of administration or expected execution level isn't reachable. Energy issues should be basic and furthermore should be overseen appropriately in some climate where portable cloud registering is included. Decreasing the measure of energy utilized by applications through green compilers and hearty programming can be accomplished through application/programming level strategies [8]. In next areas, Application level and undeniable level asset the board methods are talked

about to accomplish energy proficiency in single framework, groups, matrices and cloud datacenters.

### *3.1 Resource Scheduling Model for Energy Saving in Cloud Computing*

Fundamentally the asset model of the cloud Datacenter and the unique force model of the actual machine are both fabricated, and a short time later a three-dimensional virtual asset booking technique (TVRSM) is proposed alongside related calculations. The interaction of virtual asset booking in TVRSM is partitioned into three phases, these stages are virtual asset designation stage, virtual asset booking stage what's more, virtual asset enhancement stage. With respect to, three calculations are planned relating to the referenced phases of the virtual asset booking. These calculations are MVBPP based heuristic virtual asset distribution calculation (HVRAA), multi-dimensional force Mindful Based Virtual Asset Booking Calculation (MBVRSA) and virtual asset improvement calculation (VROA). At first, the main stage in TVRSM which is virtual asset portion stage is fundamentally responsible for designating the mentioned VMs by the client to the reasonable hosts. This stage is treated as multi-dimensional vector canister pressing issue (MVBPP) and the MVBPP based heuristics virtual asset portion calculation (HVRAA) is proposed to tackle it. Also, the second stage which is virtual asset planning stage is liable for moving the VMs from the over-burden hosts to different hosts with lower asset use by utilizing the VM movement innovation to accomplish load adjusting of the cloud server farms and furthermore to limit the measure of infringement of Service Level Understanding. The multidimensional force mindful based virtual asset planning calculation (MPVRSA) is proposed in this stage. Moreover, the third stage which is virtual asset advancement stage is responsible for moving the VMs from the hosts with the least asset use to other has and change the first has to rest mode, this cycle can additionally decrease the energy utilization of the cloud information focuses by planning the virtual asset enhancement calculation (VROA). At long last, the creators checked the viability of the proposed strategy through experimentation. The outcomes demonstrate that the TVRSM can proficiently assign and deal with the virtual assets in the cloud server farm. Furthermore, an examination is made between the proposed strategies with other conventional calculations. The results showed that the TVRSM can adequately diminish the energy utilization of the cloud server farm and limit the measure of infringement of Service Level Agreement [9].

### 3.2 Resource Model of Cloud Data Center

In [8], the authors proposed an asset model of the cloud server farm, it comprised of M bunches, and each group contains N actual machines. A few virtual machines are conveyed on each physical machine. As indicated by the assets claimed by the virtual machine, each virtual machine can run various applications. So the heap of each virtual machine results from the applications running on the virtual machine. The hub regulator runs on each actual machine is answerable for checking the asset usage of each actual machine furthermore, control the actual machines status, for example, setting the actual machine to rest mode or initiating the dozing actual machine. Additionally, the hub regulator sends the board orders to the Hypervisor to change the assets possessed by the neighborhood virtual machines. The worldwide assets the executive hub is liable for booking furthermore, distributing every one of the assets possessed by the cloud information focus. It can oversee and screen every one of the assets and execute the heap equilibrium of the cloud server farm. Each actual machine is described by the CPU execution, measure of RAM and organization transmission capacity, and each actual machine can run various virtual machines, and the actual asset claimed by each virtual machine comprises of CPU, memory limit and network transmission capacity. The actual machines use Network Joined Storage (NAS) rather than having neighborhood plates. It employments NAS to save information, which can facilitate the information sharing between every actual machine and empower live relocation of virtual machines rapidly.

### 3.3 Dynamic Power Model

In [7], the creators have shown that the force utilization by PMs can be depicted by a straight connection between the power utilization and CPU use. Thus, the force utilization $Pi(t)$ of PM $i$ running on time $t$ can be communicated by the connection between the CPU usage $ui(t)$ of PM $i$ on time $t$ and the greatest force utilization PiMax of PM I, as demonstrated in Eq. (1).

$$p_i(t) = k \times p_{imax} + (1 - k) \times p_{imax} \times u_i(t) \tag{1}$$

As demonstrated in Eq. (1), $p_{imax}$ is viewed as the greatest force utilization of host, $p_i(t)$ is considered to be the force utilization of PM I running on time $t$, $k$ viewed as the negligible portion of force utilization when the have is out of gear state and $u_i(t)$ is the CPU asset usage of the PM on time $t$. In the beneath recipe, the CPU usage of the PM I is characterized as the proportion of the all-out CPU assets mentioned by the all VMs running on PMi on time $t$ to the all Central processor assets possessed by the PM I as demonstrated in Eq. (2).

$$u_i(t) = \sum_{j=1}^{VMS_i} \frac{CPU_{rj(t)}}{CPU_{itotal}} \quad (2)$$

### 3.4 Energy Aware Virtual Machine Relocation

The proposed procedure proposes another strategy for keeping up energy effectiveness in distributed computing, by relocating the maximally stacked virtual machines to the least stacked dynamic machine, while looking after framework execution by playing out a live movement of the virtual machines to guarantee that every one of the running applications won't get separated during relocation. The proposed strategy presents another system for improving asset use levels dependent on the bio-enlivened Firefly enhancement method to accomplish energy proficiency in cloud server farms. The attainability of the proposed procedure has been appeared by executing the outcomes by utilizing the Cloud Sim test system [10, 15].

### 3.5 Firefly Optimization (FFO) Algorithm

The Firefly Optimization (FFO) calculation has been presented by Xin-She Yang in the late 2007 and 2008 at Cambridge University. It was carried out upon the fireflies blazing attributes and conduct, the attributes have been presented as follows:

1) One firefly is drawn to different fireflies paying little mind to their sex as all fireflies are unisex
2) The appeal is proportionate to the splendor, subsequently the two of them decline as their distance increments and for any two blazing fireflies, the less brilliant one will be drawn in close to the more brilliant one
3) The brilliance of a firefly is determined utilizing the target capacity to be improved The (FFO-EVMM) Technique presents the possibility of moving the most stacked VM from a functioning hub which fulfills least standards for energy utilization, to another dynamic hub that burns-through the least energy. The method is executed in four primary parts as demonstrated in Fig. 1.

## 4 Result and Analysis

The factual outcomes for the proposed FFO-EVMM calculation were contrasted and the ACO-based and Worldwide Journal of Machine Learning and Computing, Vol. 9, No. 1, February 2019 101FFD-based calculations, utilizing the CloudSim toolbox test system which has caught that that FFO-EVMM method runs less number of dynamic has and performs less number of virtual machines relocation in contrast

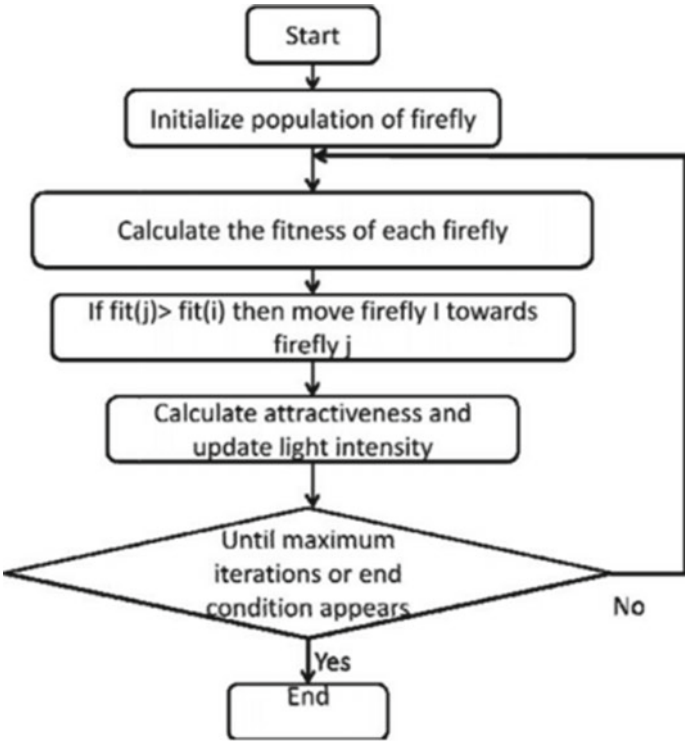



**Fig. 1** Flow chart for FFO algorithm

with ACO [13] and FFD-based calculations [14]. As it's appeared with the less number of running hosts and live relocations, FFO-EVMM requires lesser energy interest contrasting with FFD [14] and ACO [13] calculations as it has been taken note from Table 1 and Fig. 2.

## 5 Conclusion

In this paper, we have introduced a double target scheduling calculation that knows about make span and energy utilization to dispense assets to assignments and to sort undertakings, in light of heuristic strategies and grasshopper calculation. Besides, the most proper strategy to execute each undertaking is chosen dependent on the significance factor dictated by the end-client or specialist organization, set by utilizing the roulette wheel in the cloud. The proposed BHS calculation limits the objective capacity as indicated by this significance factor, considering the beginning occasions, arrangement times, and end times for virtual machines and energy profiles. Datacenter are the center of cloud figuring that store business data and host the

**Table 1** Comparative analysis

| Energy consumption (joule) | | | |
|---|---|---|---|
| Number of nodes | ACO [13] | FFD [14] | Proposed |
| 20 | 23 | 22 | **13** |
| 30 | 24 | 22 | **12** |
| 40 | 45 | 43 | **11** |
| 50 | 46 | 45 | **24** |
| 60 | 50 | 50 | **34** |
| 70 | 56 | 54 | **46** |
| 80 | 57 | 54 | **47** |
| 90 | 58 | 60 | **56** |
| 100 | 60 | 61 | **60** |
| 110 | 62 | 61 | **62** |
| 120 | 63 | 62 | **62** |
| 130 | 70 | 67 | **64** |
| 140 | 72 | 67 | **59** |
| 150 | 80 | 78 | **75** |
| 160 | 82 | 78 | **72** |
| 170 | 92 | 87 | **79** |
| 180 | 92 | 86 | **80** |
| 190 | 93 | 89 | **81** |
| 200 | 93 | 90 | **83** |

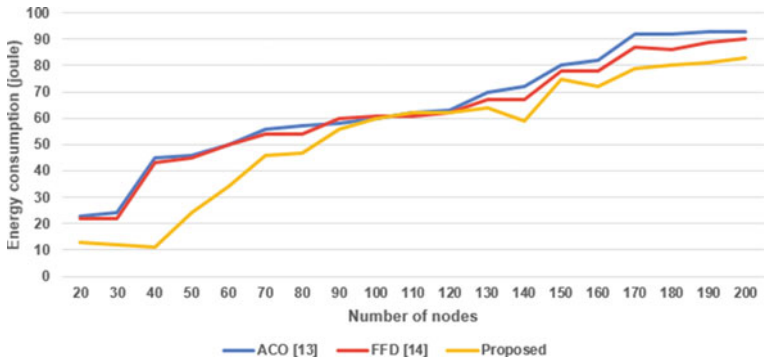



**Fig. 2** Virtual machine vs energy consumption

running applications. Superior has consistently been the sole worry of all in server farms. This worry has been overseen without thinking about energy utilization and execution. The test is to adjust between power utilization and framework execution. Numerous methods and calculations have been proposed to accomplish sufficient energy use in cloud server farms.

## References


1. Buyya R, Yeo CS, Venugopal S (2008) Market-oriented cloud computing: vision, hype, and reality for delivering it services as computing utilities. In: Proceedings of the 10th IEEE international conference on high performance computing and communications, pp 5–13
2. Beloglazov A, Buyya R (2012) Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Concurr Comput Pract Exp 24(13):1397–1420
3. Beloglazov A, Buyya R, Lee YC, Zomaya A (2011) A taxonomy and survey of energy-efficient data centers and cloud computing systems. Adv Comput 82:47–111
4. Juarez F, Ejarque J, Rosa MB (2018) Dynamic energy-aware scheduling for parallel task-based application in cloud computing. Futur Gener Comput Syst 78(1):257–271
5. Stavrinides GL, Karatza HD (2019) An energy-efficient, QoS-aware and cost-effective scheduling approach for realtime workflow applications in cloud computing systems utilizing DVFS and approximate computations. Futur Gener Comput Syst 96:216–226
6. Fakhar F, Javed B, Rasool R, Malik O, Zulfiqar K (2012) Software level green computing for large scale systems. J Cloud Comput Adv Syst Appl 1(1):4
7. Zhu W, Zhuang Y, Zhang L (2017) A three-dimensional virtual resource scheduling method for energy saving in cloud computing. Futur Gener Comput Syst 69:66–74
8. Monisha N, Jaya Kumar D, Deepika V, Ishwarya A (2016) An efficient technique for data compression and convergent encryption in the hybrid cloud. Int J Comput Sci Mob Comput 3:330–335
9. Jayakumar D (2013) A resourceful allocation of data storage in cloud computing. Int J Adv Eng Emerg Technol 1:45–50
10. Kansal NJ, Chana I (2016) Energy-aware virtual machine migration for cloud computing-a firefly optimization approach. J Grid Comput 14(2):327–345
11. Mandal R, Mondal MK, Banerjee S, Biswas U (2020) An approach toward design and development of an energy-aware VM selection policy with improved SLA violation in the domain of green cloud computing. e J Supercomput 76(9):7374–7393
12. Barzegar B, Motameni H, Movaghar A (2019) EATSDCD: a green energy-aware scheduling algorithm for parallel taskbased application using clustering, duplication and DVFS technique in cloud datacenters. J Intell Fuzzy Syst 36(6):5135–5152
13. Alharbi F, Tian YC, Tang M, Zhang WZ, Peng C, Fei M (2019) An ant colony system for energy-efficient dynamic virtual machine placement in data centers. Expert Syst Appl 120:228–238
14. Jangiti S, Ram ES, Sriram VS (2019) Aggregated rank in first-fit-decreasing for green cloud computing. In: Cognitive informatics and soft computing, pp 545–555. Springer, Singapore
15. Jangiti S, Shankar Sriram SS (2020) EMC2: energy-efficient and multi-resource-fairness virtual machine consolidation in cloud data centres. Sustain Comput Inform Syst 27:100414

# Advanced Architecture of Analog to Digital Converter Derived from Half Flash ADC


**Tejaswini Jayawant Kutre, Sujata N. Patil, Sheela Kore, and V. M. Aparanji**

**Abstract** In all digital systems, data conversion is the main process that is done with the help of an Analog to digital converter. There are many types of ADC's available for this conversion but among those Flash, ADCs are more advantageous because of their high speed. Flash ADC has a high speed due to the simultaneous conversion. The addition of the comparator delays and the logic delays (logic delays are negligible) gives the conversion speed of flash ADC. So, Flash ADC is one of the most commonly used ADC in high-speed applications. But even though it has high speed its area and power requirement are also more if we consider conventional Flash ADC. To enhance the performance of the Flash ADC the advanced architecture of ADC is proposed using the information of the half Flash ADC in which conversion happens in two steps. For 8-Bit half Flash ADC, the first higher 4-Bits are calculated. And then in the second step remaining 4 -Bits are calculated. Flash ADC also requires a large no of components so it is difficult to use it under any microcontroller. Hence in the proposed work, ADC is designed to reduce the complexity of conventional Flash ADC and to achieve less area and less power consumption. This model is designed based on half Flash ADC only difference is that here Bitwise conversion happens. It uses $2N - 1$ operational amplifiers with the elimination of DAC and encoders which makes the circuit suitable for use inside the micro controller. The proposed work explains the design and implementation of 3-Bit Advanced Flash ADC architecture with Error correction.

**Keywords** Analog to digital converter · Advanced architecture · Half flash

T. J. Kutre · S. N. Patil (✉) · S. Kore
Department of Electronics and Communication Engineering, KLE Dr. M S Sheshgiri College of Engineering and Technology, Belgaum, Karnataka, India
e-mail: sujata.patil@klescet.ac.in

V. M. Aparanji
Siddaganga Institute of Technology, Tumkur, Karnataka, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_14

## 1 Introduction

In the Flash ADC main component is the comparator and its design is very challenging. This comparator compares the input signal with reference signal [1]. The rise in resolution bits increases the number of Flash ADC comparators. So, the design should be done to reduce the number of comparators because a huge count of comparators requires large power and a larger area. The second important component is the thermometer to a binary encoder which gives the final digital output [2]. There are many types of encoders are available such as fat-tree, binary-ROM, folded Wallace tree and Wallace encoder, etc. But these all encoders have their limitations like low speed, more complexity, large area, bubble error occurrence, large delay, and unbalanced signal paths [3]. So there a need to design the ADC with a fewer number of components which reduces the power and area requirement and also reduces the circuit complexity. Hence to reduce the circuit complexity and die area of the Flash ADC Half Flash ADC was introduced [4, 5]. Suppose 4 Bit conversion needs to be done then the first 2 Bits are converted with step size VREF/4. Then the output is fed to a 2 Bit DAC in which four voltage levels are generated (0, VREF/4, VREF/2, 3VREF/4) which are quantizing levels of the converter. Then a differential amplifier is used to subtract the input voltage from the output DAC. The differential amplifier output must be within VREF/4 and this reference voltage is used for the remaining two-bit conversion having a step size of (VREF/4)/4. Hence a Bit result is obtained with fewer comparators and smaller encoding circuit [6].

The model proposed in this paper follows almost the same working technique as Half Flash ADC. The only difference is that here bitwise operation accurses with the elimination of DAC. In this proposed work first input voltage is compared with VREF/2. The output of the first comparator is the MSB bit of the final digital output is fed to differential amplifier along with input voltage. The differential amplifier used here is designed such that it should give the output within 0 to VREF/2. This output is compared with VREF/4 and comparator output with differential amplifier output is fed to the next differential amplifier which is designed to give the output within 0 to VREF/4. This process continues till the LSB bit is obtained. The proposed model has a conversion time almost the same as half flash because here the encoder circuit and DAC are eliminated. Due to its large speed of conversion and smaller size, it can be used in digital applications [7].

## 2 Related Works

This section gives brief idea of the work carried out related to Flash type analog to digital converters. Here we discuss the different architectures of ADCs proposed by using different methods. Islam et al. 2019 proposed the simplified architecture of Analog to Digital Converter by referring the method used in the Half Flash ADC architecture. In this work 3-Bit ADC is designed by eliminating the components

such as encoder and DAC and simulated by using Proteus simulation tool. In this proposed work N-Bit ADC requires 2N − 1 differential amplifiers. For 3-bit ADC implementation 5 differential amplifiers are used. This work reduces the complexity of the ADC circuit by using a smaller number of components compared to other architectures. Also, it requires less power and area. Modi et al. 2018 in this work author has proposed Flash ADC architecture with reduction in count of comparators.

Tan et al. 2004 designed the simplified architecture of Half Flash ADC by using tanner tool. In this architecture conversion happens in two cycles. For 8-bit ADC first 4 bits are calculated in first cycle and in next cycle remaining 4 bits are calculated. For 8-Bit resolution 80% comparators reduction and for higher resolution more than 80% achieved. The proposed work requires less area and power due to less no of comparators. The seed of conversion of this ADC will be similar like conventional Flash ADC. Sivakumar et al. 2019 introduced a new digital built in self-test approach to test an ADC in time domain. The developed testing approach comprises of ADC test circuit, linear analog ramp generator, BIST controller and an output response analyzer. Kumar et al. 2020 designed a new ADC with limited power consumption and operated voltage. In this research article, a flash type ADC was designed with less transistors and comparator with encoder using tanner EDA analog environment 180 nm technology. N.H. Patel, 2021 designed a power efficient 4-bit flash ADC for medical applications. The power efficient 4-bit flash ADC has comparator, encoder, and hold circuit for obtaining a power effective ADC.

Tan et al. 2000 in this work author has proposed CMOS ADC architecture with 8-bit resolution and high efficiency. In this design two improved full flash ADCs with voltage estimator with 4-bit resolution are used. The minimization with area and power consumption is achieved with flash ADC. Senthil Sivakumar and Sowmya Priya (2019) analyzed the power utilization, and area of different comparators for solving the area utilization issue in the flash ADC. The comparator circuits were simulated by using CMOS 180 mn technology. In this literature, the delay, power and area of the dissimilar comparators were compared for best usage in the flash ADC.

## 3 Proposed Architectures

### *3.1 Advanced Architecture of Flash ADC*

In this proposed architecture we consider the input voltage $V_{in}$ is the signal which should be transformed into digital output. Here the IC biasing and reference voltage $V_{in} <= V_{CC}$ and $V_{CC}$. Here we are designing the 3-bit Flash ADC so the output varies from 000 to 111. Here the MSB is calculated first. The comparison and done between $V_{CC}/2$ and analog input voltage (i.e., binary 100). In Fig. 1 if $R_1 = R_2$ we use the voltage divider law the voltage at the inverting terminal of the operational amplifier $U_1$ will be $V_{CC}/2$ [8]. Hence if the input voltage is greater than $V_{CC}/2$ then the yield

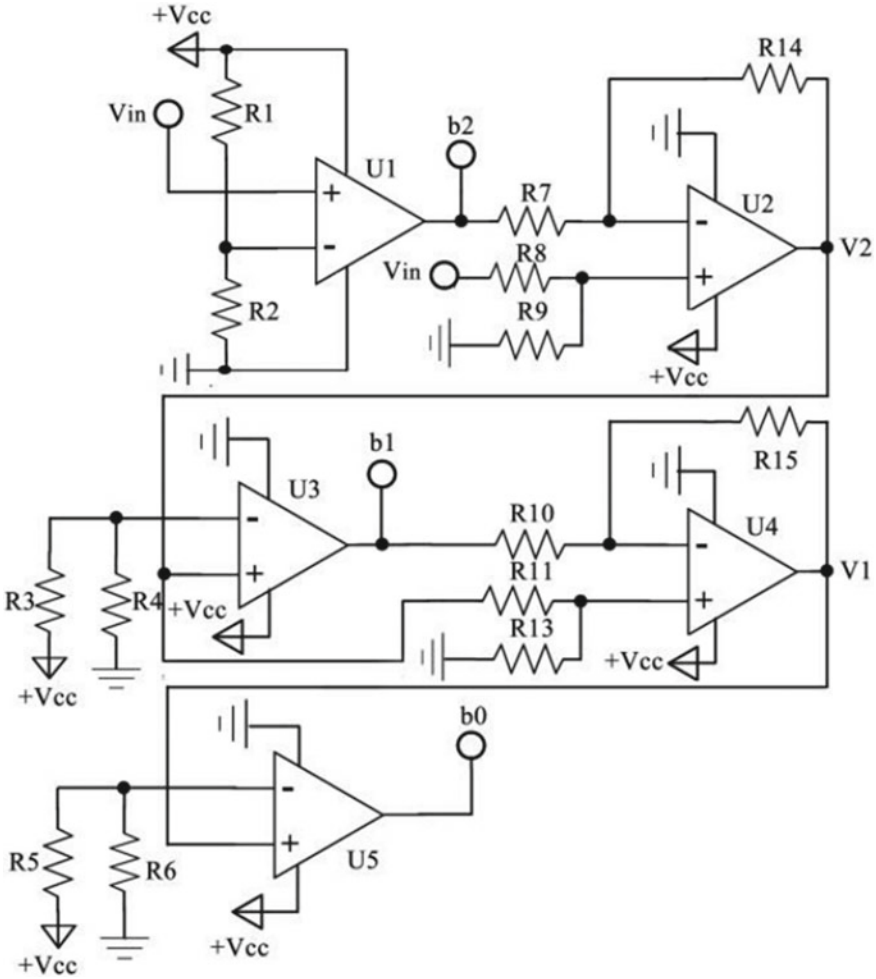



**Fig. 1** Advanced flash ADC circuit

of $U_1$ will be equal to $V_{CC}$ (i.e., Bit 2 will be made as high or logic 1 otherwise the output will be zero volts (i.e., $b2 = 0$).

At the first stage, the $V_{CC}/2$ were the step size. It will be reduced to $V_{CC}/4$ to determine the second MSB Bit. If applied input is more than $V_{CC}/2$ then subtraction of $V_{CC}/2$ is done from input voltage by using an op-amp $U_2$ differential amplifier. The superposing theorem is used to calculate the output of differential amplifier $U_2$ which is calculated theoretically by Eq. 1. Each step process is explained in Eqs. 2, 3, and 4.

$$V_2 = -(R_{14}/R_7)(V_{CC.}\, b_2) + \{R_9\,(R_{14} + R_7)/R_7(R_8 + R_9)\}.V_{in} \qquad (1)$$

Now if, $R_8 = R_{14}$, $R_7 = R_9$ and $R_9 = 2R_8$ then Eq. (1) becomes

$$V_2 = -(R_8/R_7)(V_{CC.}\, b_2) + \{R_7\,(R_8 + R_7)/R_7(R_8 + R_7)\}.V_{in} \tag{2}$$

$$V_2 = -(R_8/2R_8)(V_{CC.}\, b_2) + \{(R_8 + 2R_8)/(R_8 + 2R_8)\}.V_{in} \tag{3}$$

$$V_2 = V_{in} - (V_{CC}/2).\, b_2 \tag{4}$$

So, initial step size is subtracted from input if MSB is set otherwise second comparator is fed with the original input voltage. If $V_2$ is greater than Vcc/4 then $V_{CC}/4$ is subtracted from it by using $U_4$. The output of $U_4$ can be calculated using below equations (Eqs. 5, 6, 7, and 8),

$$V_1 = -(R_{15}/R_{10})(V_{CC.}b_1) + \{R_{13}(R_{15} + R_{10})/R_{10}(R_{11} + R_{13})\}.V_{in} \tag{5}$$

If, $R_{11} = R_{15}$, $R_{10} = R_{13}$, $R_{13} = 4R_{11}$ then Eq. (5) becomes

$$V_1 = -(R_{11}/R_{13})(V_{CC.}b_1) + \{R_{13}(R_{11} + R_{13})/R_{13}(R_{11} + R_{13})\}.V_2 \tag{6}$$

$$V_1 = -(R_{11}/4R_{13})(V_{CC.}b_1) + \{(R_{11} + 4R_{11})/(R_{11} + 4R_{13})\}.V_2 \tag{7}$$

$$V_1 = V_2 - (V_{CC}/4).b_1 \tag{8}$$

Finally, $V_1$ is with Vcc/8, which calculates the LSB (i.e., b0) value.

This calculation uses below Fig. 2 to find the reference voltage position and corresponding binary value can be known before conversion.

For example, we have the Vcc as 5v and an input voltage of 2.24Vcc/8 which lies between 2Vcc/8 and 3Vcc/8. So, the predicted output by using Fig. 2 is 010 after conversion. Thus, we can see that the input is in the $6^{th}$ quantization level.

First, we compare the input voltage with Vcc/2 the comparator $U_1$ will have the output as logic 0. This sets the MSB ($b_2 = 0$). And the output of the differential amplifier is 1.40 V ($V_2$) by using Eq. 4. This is again compared with Vcc/4 by using $U_3$ this will give the second MSB as logic 1. (i.e., $b1 = 1$). And the second amplifier $U_4$ output will be 0.15 V ($V_1$) by using Eq. 8.

Finally, $V_1$ is compared with Vcc/8 by using comparator $U_5$ which gives the LSB bit as logic 0 ($b_0 = 0$). Thus, we obtained the conversion of 1.40 as 010 by using this 3-Bit advanced Flash ADC architecture.

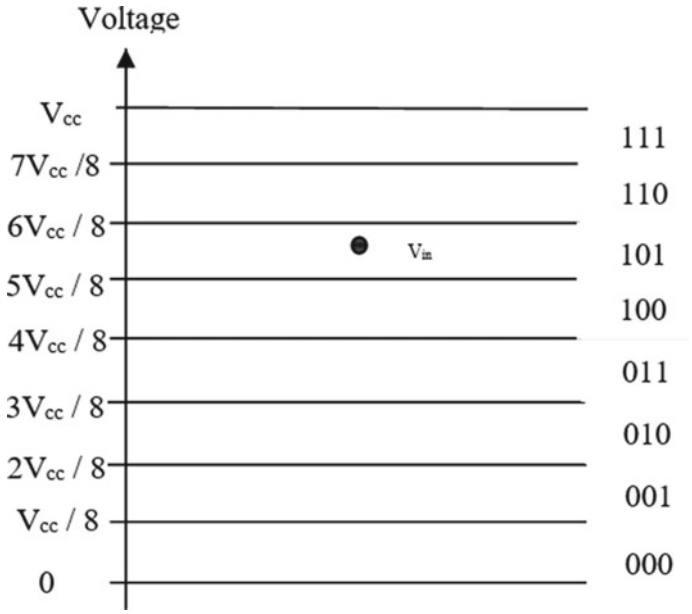



**Fig. 2** Reference voltage levels with corresponding binary values

## 4 Implementation of Advanced Flash ADC

In this work 3-bit advanced converter is implemented on Proteus 8.1 using a general comparator model with an output of 5 V and an LM358 operational amplifier as a differential amplifier. The Fig. 3 shows the implementation of advanced Flash ADC. An input 2.75 V is given as by varying the potentiometer, reference voltage of 5, and 12 V of biasing voltage, compared with Vcc/2 by using a $U_2$ comparator. Here we get the comparator as logic "1" indicated by $b_2$. This is the first MSB bit obtained. The voltage at the output of differential amplifier $U_5$: B is +0.25 V which fulfil the Eq. 4. The same voltage is given to the $U_3$ comparator which compares $V_2$ voltage with a step size of $V_{CC}/4$. The output of comparator $U_3$ gives the second MSB bit $b_1$ logic "0". The output of the differential amplifier $U_6$: A as +0.26 V ($V_1$) which fulfils the Eq. 8 with a small error voltage (0.01 V). Finally, $V_1$ is compared with step size $V_{CC}/8$ by using a $U_1$ comparator which gives LSB bit $b_0$ as 0. Thus, we obtain the conversion of 2.75 as 100.

From the above Table 1, we can observe the analog to digital conversion of different values. This model produces a small amount of error when the output voltage of the operational amplifier is greater than its input. It ranges from 1–3 mv in simulation but in the practical case when input is near to zero it will be 16 mv offset value. This causes difficulty in achieving higher resolution.

The simulation includes the ideal comparator by using a general comparator model but in the case of practical implementation, the output will be smaller than 5 V with

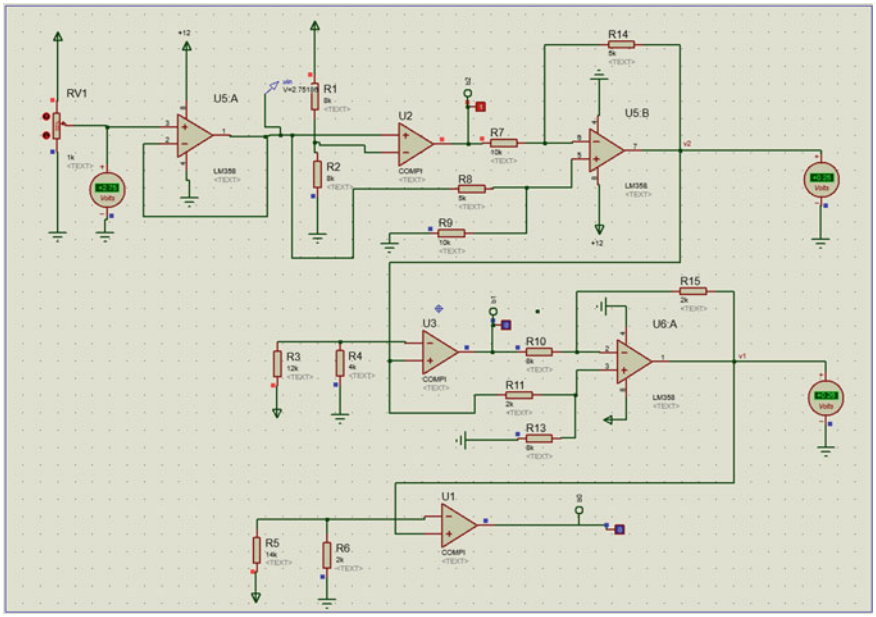



**Fig. 3** Implementation of advanced flash ADC architecture

**Table 1** Simulation results of A/D conversion

| Analog input | Circuit response | | | | |
|---|---|---|---|---|---|
| | Digital output | Differential amplifier output due to 2nd MSB (V2) | | Differential amplifier output due to 1st MSB (V1) | |
| | | Desired voltage | Obtained voltage | Desired voltage | Obtained voltage |
| 0 | 000 | 0.00 | 0.03 | 0.00 | 0.04 |
| 0.80 | 001 | 0.80 | 0.80 | 0.80 | 0.81 |
| 1.40 | 010 | 1.40 | 1.40 | 0.15 | 0.16 |
| 2.15 | 011 | 2.15 | 2.15 | 0.9 | 0.91 |
| 2.75 | 100 | 0.25 | 0.25 | 0.25 | 0.26 |
| 3.00 | 100 | 0.50 | 0.50 | 0.50 | 0.51 |
| 3.55 | 101 | 1.05 | 1.05 | 1.05 | 1.06 |
| 4.00 | 110 | 1.50 | 1.50 | 0.25 | 0.26 |

5 V biasing. LM358 gives the maximum output of 4.86 with 5 V reference voltage. We can see a large amount of error is occurring here.

This error can be eliminated by using the amplifier which works similarly to an ideal op-amp that is a rail-to-rail operational amplifier that can be a solution for

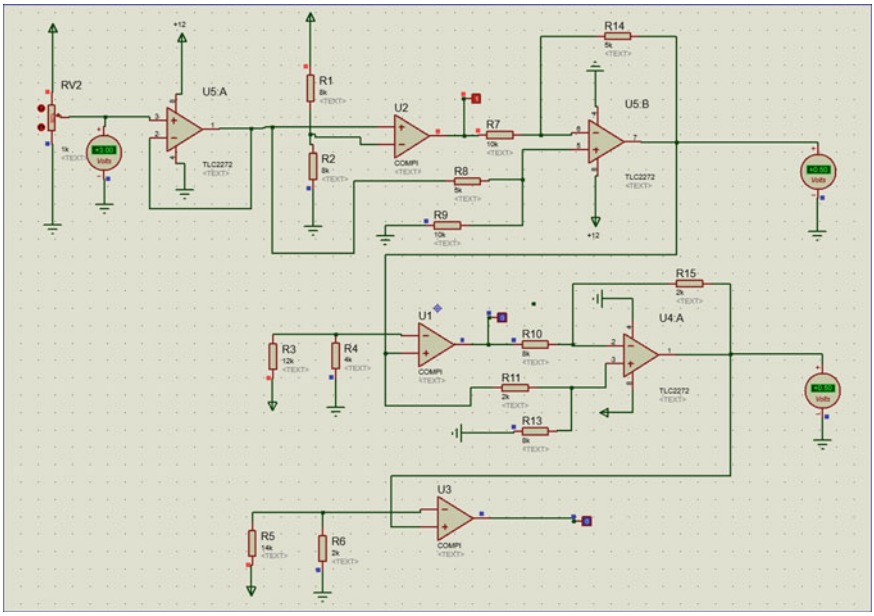



**Fig. 4** Error correction by using TLC2272 dual-rail to rail operational amplifier

this problem [9]. Hence, we have designed the proposed circuit using TLC2272 rail to rail operational amplifier as differential amplifier eliminates the error occurred. It provides 0.5 pA offset current and 300 uv offset voltage of input hence the biasing voltage and reference voltage will be kept the same (i.e., 5 V). Now the comparator output ranges from 01–4.99 V. Implementation of the improved circuit using TLC2272 is shown in the Fig. 4.

Above Table 2 shows the simulation results of the proposed work after error correction by implementing the TLC2272 rail-to-rail operational amplifier.

## 5 Comparison with Existing Models of ADC

The advantage of this work is that it has very little delay compared to digital ramp ADC and successive approximation ADC because it does not require any clock signal. The propagation delay is only present here. If we compare with conventional Flash ADC architecture, it has more delay because in conventional flash ADC all the comparators work parallel. But the area required will be more because in conventional flash ADC increase in resolution increases the number of comparators [10]. So, to reduce the complexity, power requirement, and area proposed model can be used. Its conversion time will be nearly equal to 2.5–1.5 us with a high-speed operational amplifier.

**Table 2.** Simulation results after error correction

| Analog input | Circuit response | | | | |
|---|---|---|---|---|---|
| | Digital output | Differential amplifier output due to $2^{nd}$ MSB ($V_2$) | | Differential amplifier output due to $1^{st}$ MSB ($V_1$) | |
| | | Desired voltage | Obtained voltage | Desired voltage | Obtained voltage |
| 0 | 000 | 0.00 | 0.16 | 0.00 | 0.18 |
| 0.80 | 001 | 0.80 | 0.80 | 0.80 | 0.80 |
| 1.40 | 010 | 1.40 | 1.40 | 0.15 | 0.15 |
| 2.15 | 011 | 2.15 | 2.15 | 0.9 | 0.99 |
| 2.75 | 100 | 0.25 | 0.25 | 0.25 | 0.25 |
| 3.00 | 100 | 0.50 | 0.50 | 0.50 | 0.50 |
| 3.55 | 101 | 1.05 | 1.05 | 1.05 | 1.05 |
| 4.00 | 110 | 1.50 | 1.50 | 0.25 | 0.25 |

In the complex and high-density circuits observing the correctness of the output is also difficult hence to detect the errors built-in-self-test (BIST) technique can be used in flash ADC [11]. Different optimization comparators, decoders can be used to improve the efficiency of Flash ADC [12]. Sample and hold circuit can be used to reduce power consumption instead of R-2R ladder [13]. The digital part in the Flash ADC can be configured using field programmable gate array [FPGA] and Verilog language can be used for programming [14]. A threshold inverter quantization (TIQ) based comparator can be introduced for high-speed low area flash ADC applications [15]. The Threshold Modified Comparator Circuit (TMCC) can be implemented to improve the speed and area of the Flash ADC [16]. This all optimizations can be done in Flash ADC to improve the performance but these methods increase the complexity of Flash ADC circuit. The work discussed in this paper makes the circuit simple without compromising the performance parameters [17, 18].

The proposed model requires $2N - 1$ operational amplifiers for N-Bit conversion. 15 operational amplifiers are required for 8-Bit conversion. This model eliminated the use of encoders and DAC, also less operational amplifiers are used which makes the circuit smaller requires less power, and relatively faster which makes it suitable to use inside any microcontroller.

Table 3 shows the comparison of various architectures of the flash ADC implemented by using different techniques for optimization in terms of speed, area and power.

**Table 3** Comparison of flash ADC's

| Specification | 10 | 11 | 12 | 13 | This work (unit) |
|---|---|---|---|---|---|
| Resolution (Bit) | 6 | 7 | 3 | 4 | 3 |
| Supply voltage (V) | 1 | 1 | 2.1 | 1.8 | 5 |
| Technology (nm) | 90 | 180 | 180 | 180 | 180 |
| Comparators required | 17 | 64 | 12 | 15 | 3 |
| Power (w) | 1.19 m | – | 0.34 μ | 2.88 m | 1.1 m |
| Frequency (Hz) | 10 M | – | 10 M | 3.20 M | – |
| Conversion time | 100 n | – | – | – | 2.5–1.5 |

## 6 Conclusion

The proposed model of 3-bit advanced Flash ADC has a simple circuit with a smaller number of components since encoder and DAC circuits are eliminated. It uses 5v of the reference voltage and 12 V of IC biasing voltage. The proposed model includes comparators and differential amplifiers to generate each bit. For N-Bit conversion, it requires a $2N - 1$ number of operational amplifiers. For example, the 8-Bit conversion 15 operational amplifiers are required. Thus, we can say that the proposed model requires less area and less power consumption and it is faster than successive approximation ADC and sigma-delta Adc. And thus, due to smaller size and improved performance, it can be included within any microcontroller. The Error correction is done efficiently by using the TLC2272 differential amplifier which gives 0.5 uA of input offset current and 300 uv of input offset voltage and the output range of 0.01–4.99 V. In the future, its speed can be improved with higher resolution by using high-speed operational amplifiers. Also, a low offset comparator can be designed to reduce the error occurrence.

## References

1. Yadav RVPS, Srivastava N (2018) Implementation of high speed flash ADC using multiplexer with reduced number of preamplifier and comparator count. IJSR 7(10)
2. Tocci RJ, Widmer NS, Moss GL Digital systems, principles and applications, 10th edn. [copyright date and publisher include]
3. Modi V et al. (2018) MUX based flash ADC for reduction in number of comparators. In: Proceeding IEEE international conference on intelligent circuits and systems. 978-1-5386-6483-4/18/$31.00
4. Texas Instruments (1999) ADC0820-N 8-bit high speed μp compatible A/D converter with track/hold function, ADC0820 datasheet. Accessed Mar 2013
5. Tan PBY et al (23004) Simplified half-flash CMOS analog-to-digital converter. In: NSTI-Nanotech 2004, vol 2. www.nsti.org. ISBN 0-9728422-8-4
6. Lee DY (2000) Half-flash type analog-to-digital converter capable of converting data at high speed. United States Patent 6,011,503, 3 January 2000

7. Mashrur Islam Md et al (2019) A successive method of flashing for ADC network. Asian J Sci Technol 01(01):17–20
8. Alexander CK, Sadiku MNO (2013–2014) Fundamentals of electric circuits, 4th edn. Mc-GRAW – Hill
9. Texus Instruments (1999) Application of Rail-to-Rail operational amplifier. Application Note December 1999
10. Kutre TJ, Patil SN, Kore SK (2020) Design of MUX based Flash ADC for reduction in number of comparators. IOP Conf Ser Mater Sci Eng 925:012063. https://doi.org/10.1088/1757-899X/925/1/012063
11. Sivakumar MS, Gurumekala T, Pulya SA (2019) Error detection of data conversion in flash ADC using code width based technique. Procedia Comput Sci 165:270–277
12. Pavan Kumar M et al (2020) Design and implementation of efficient flash ADC. Int J Eng Res Technol (IJERT) 9(05). ISSN 2278-0181
13. Patel NH (2021) Power efficient 4-bit flash ADC using Cadence Virtuoso. Int J Eng Res Technol (IJERT) 10(03). ISSN 2278-0181
14. Abualsaud AA, Qaisar SM, Ba-Abdullah SH, Al-Sheikh ZM, Akbar M Design and Implementation of a 5-Bit Flash ADC for Education. https://doi.org/10.1109/ICEDSA.2016.7818471
15. Senthil Sivakumar M, Sowmya Priya M (2019) Design and analysis of a comparator for flash ADC. Int J Recent Technol Eng (IJRTE) 7(5S4). ISSN 2277-3878
16. Srinu B, Nagendra Kumar M (2017) Design of 4 Bit FLASH analog to digital converter using TM comparator circuit and gray to base2 encoder using 0.13μm CMOS technology. IJECT 8(3)
17. Tan PBY, Suparjo BS, Wagiran R, Sidek R (2000) An efficient architecture of 8-bit CMOS analog-to-digital converter. In: ICSE conference, pp 178–186
18. Yin S, Shizu L, Ronghua Z, Shoujue W (n.d.) A new half-flash architecture for high-speed video ADC. In: 5th international conference on solid-state and integrated circuit technology. proceedings (Cat. No. 98EX105). https://doi.org/10.1109/icsict.1998.785900

# An Assessment of Criss-Cross Multilevel Inverter with Fault Tolerance for Electric Vehicle Applications


**B. Divyashree, S. Nagaraja Rao, and Veerabhadra**

**Abstract** In this paper Criss-Cross Multilevel Inverter (CC-MLI) configuration is introduced for Electric Vehicle (EV) applications. The CC-MLI uses less power switches, improving modularity with fault tolerant capability using symmetric sources for the nine-level. To produce target level for synthesizing the sinusoidal output, the CC-MLI structure consists of semi-half-bridge cells linked in series with criss-cross switches. Inverter switching is achieved by generating acceptable switching angles with optimization switching angle technique using Round Modulation Control (RMC) to improve CC-MLI power quality. The RMC switching technique is also compared with conventional switching angle methods, such as Equal-Phase Pulse (EPP) and Half-Equal-Phase Pulse (HEPP) methods. The CC-MLI requires only '(m + 11)/2' power switches, '(m – 1)/2' DC sources and bypass diodes for 'm' levels, whereas the conventional or classical MLIs require '(2m – 1)' power switches. In addition, the proposed CC-MLI has modular circuit layout and can be easily extended to higher levels. An assessment of proposed CC-MLI configuration is done in terms of Total Harmonic Distortion (THD) and fundamental voltage. The %THD of RMC method is obtained as 9.37%, whereas %THD of EPP and HEPP methods are obtained as 25.37% and 22.06% respectively for nine-level CC-MLI. The simulation for a single-phase nine-level CC-MLI with single-switch fault tolerance is validated using MATLAB/Simulink using EPP, HEPP and RMC switching angle control techniques.

**Keywords** CC-MLI · Control techniques · Fault tolerance and THD

## 1 Introduction

Multilevel inverters (MLIs) are gaining popularity these days due to their numerous benefits, including high power efficiency, lower harmonic elements, lower dv/dt, and lower switching losses. MLIs are categorized into three types, namely Diode clamped

B. Divyashree · S. N. Rao (✉) · Veerabhadra
M.S. Ramaiah University of Applied Sciences, Bangalore, Karnataka, India
e-mail: nagarajarao.ee.et@msruas.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_15

[1], Flying capacitor [2] and Cascaded H-bridge MLIs [3]. These configurations are called as classical or conventional MLIs. It needed a greater number of components such as a DC voltage source, power switches, capacitors, and clamping diodes if the number of levels exceeds five. However, due to its modularity and ease of control, the cascaded H-bride MLI has gained particular interest, and also various topologies have been proposed for MLIs [4, 5]. Due to discrete voltage sources, the CHBMLI is more reliable for high-voltage applications and hence free of voltage-balancing problems. If the degree grows, however, it necessitates a greater number of components. In this paper, an assessment and THD evaluation of CC-MLI along with single power switch fault tolerance operation for nine-levels have been performed using the RMC switching angle control technique and the compared with conventional EPP and HEPP switching angle control techniques [10].

The suggested fault tolerant based CC-MLI configuration can be utilized for the EV applications to drive the motor, since existing two-level inverters need a bulky inductor, costly and inefficient, so to overcome from these limitations a CC-MLI configuration can be used to drive the wheels with the integration of CC-MLI with an electric motor [11, 12]. Further, continuity of operation is critical, since numerous breakdowns might occur inside an EV's propulsion system [13]. As a result, fault-tolerance is identified as a necessity for the various EV's propulsion systems, including motor drive. The block diagram of CC-MLI fed motor drive for EV application is shown in Fig. 1. Most of the EVs requires two individual converters firstly bidirectional DC-DC boost converter and one more is suggested MLI i.e., CC-MLI to boost up the efficiency by improving the power quality with RMC based switching angle control technique [14]. The suggested CC-MLI is analyzed with fault tolerant operations using redundant states, as a result of the CC-MLIs ability to monitor which battery modules receive energy; the battery capacity is no longer constrained by the weakest battery modules or power switch failure. If one of the modules is defective or any power switch failure occurs, then the available energy is limited, the CC-MLI can be programmed to use the remaining energy in the battery pack, effectively bypassing the faulty module.

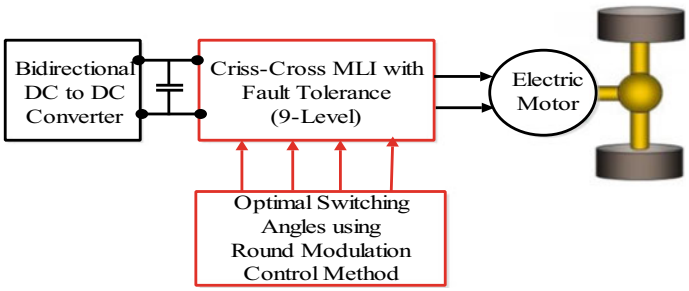



**Fig. 1** Block diagram for EV system

**Table 1** Comparison of classical and CC-MLI configurations

| Components | Classical MLIs | | | CC-MLI |
|---|---|---|---|---|
| | Diode clamped MLI | Capacitor clamped MLI | Cascade H-bridge MLI | |
| Main switches | 2(m − 1) | 2(m − 1) | 2(m − 1) | (m + 11)/2 |
| Main diode | 2(m − 1) | 2(m − 1) | 2(m − 1) | (m − 1)/2 |
| Clamped diode | 2(m − 3) | 0 | 0 | 0 |
| Balance capacitor | 0 | (m − 2) | 0 | 0 |
| DC-bus capacitor | (m − 1) | 0 | 0 | 0 |
| DC source | 1 | 1 | (m − 1)/2 | 2(m − 1) |

## 2 Related Works

The classical MLIs requires more power switches i.e., '2(m – 1)', therefore, the classical MLIs increases the hardware complexity, bulky and additional space is required for installation [5]. Comparatively diode clamped and flying capacitor MLIs have certain limitations [1, 6], the diode clamped MLI requires excess diodes whereas, the flying capacitor MLI requires excess capacitors in addition to power switches with increase in levels, in addition, the flying capacitor MLI creates voltage-balancing issues [7]. Therefore, to overcome these barriers, a CC-MLI is introduced with '(m + 11)/2' power switches, '(m – 1)/2' DC sources and bypass diodes for 'm' levels [8]. Further, a CC-MLI can have improved performance with fundamental switching frequency-based switching angle control techniques over high frequency modulation methods [9]. The requirement of power switches, DC sources and various components for classical MLIs and proposed CC-MLI is shown in Table 1 [16, 17].

The classical MLIs requires 16 power switches [1, 2] for nine-level, with half of them are in conduction at any given time to generate particular level, while in the CC-MLI configuration, the number of switches is always less at any level beyond five. It requires only 10 power switches [15], thus reduced switch count is the one of the main advantages of this configuration.

## 3 Proposed Method

The structure of the proposed CC—MLI configuration for nine-level generation is depicted in Fig. 2.

The CC-MLI consists of string of voltage sources (Va1, Va2) and (Vb1, Vb2) connected in a criss-cross fashion via switches Sc1 and Sc2. The bypass diodes (Da1, Da2) and (Db1, Db2) are used to bypass the voltage sources from the load, while the switches (Sa1, Sa2) and (Sb1, Sb2) connect the voltage sources in series. The power switches (S1, S2) complementary to the switches (S′1, S′2) are arranged in an

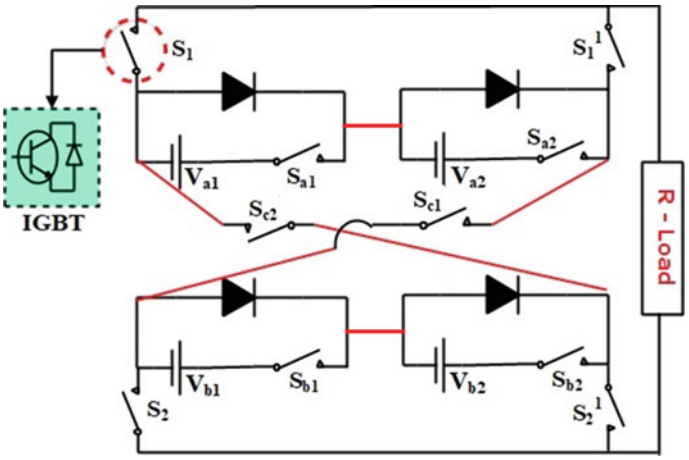



**Fig. 2** Configuration of nine-level CC-MLI

H-bridge fashion. The development CC-MLI topology allows the power switches to share voltage stress via a series connection of semi-half bridge modules intertwined with criss-cross power switches and symmetric DC sources [15]. Further, the valid switching states of CC-MLI configuration are listed in Table 2 for the generation of nine-levels.

The MLIs control techniques are critical since they are directly related to the system's overall performance. In this paper RMC switching angle control technique is used to trigger the power switches of CC-MLI and the performance of RMC has been compared with EPP and HEPP method control techniques [9, 10].

## *3.1 Equal-Phase Pulse (EPP) Method*

The EPP is the method which derived from the average distribution of switching angle in the range of 0 to 180°. Switching angle calculation for EPP method is given in Eq. (1).

$$\alpha_i = \mathrm{i}\left[\frac{180}{m}\right] \tag{1}$$

where, $\mathrm{i} = 1, 2, 3 \cdots \left[\frac{m-1}{2}\right]$ m = number of levels.

For the nine-level i.e., m = 9, the four main switching angles $\alpha_1$, $\alpha_2$, $\alpha_3$ and $\alpha_4$ using EPP method are $20^0$, $40^0$, $60^0$ and $80^0$ respectively.

**Table 2** Valid switching states for nine-level CC-MLI configurations

| Sl. no. | Switching states in healthy operation | Redundant switch states | Outputs voltage ($V_0$) |
|---|---|---|---|
| 1 | S1, Sc2, S2′,Sc1, S2 | S1, Sc2, S1′, Sc1, S2 | 0 |
| 2 | S2′, Sc2, Sa1, S1′ | S2′, Sc2′, Sa2, S1′<br>S2, Sb1, Sc2, S1<br>S2, Sb2, Sc2, S1<br>S2′, Sc2, Sa2, S1′ | +Vdc |
| 3 | S2′, Sc2, Sa1, Sa2, S1′ | S2, Sb1, Sc2, Sa1, S1′<br>S2′, Sc2, Sa1, Sa2, S1′<br>S2, Sb1, Sb2, Sc2, S1 | +2Vdc |
| 4 | S2, Sb1, Sc2, Sa1, Sa2, S1′ | S2, Sb1, Sb2, Sc2, Sa1, S1′<br>S2, Sb1, Sc2, Sa1, Sa2, S1′ | +3Vdc |
| 5 | S2, Sb1, Sb2, Sc2, Sa1, Sa2, S1′ | No redundant state | +4Vdc |
| 6 | S2, Sc1, Sa2, S1′ | S2′, Sb2, Sc1, S1′<br>S2, Sc1, Sa2, S1<br>S2′, Sb1, Sc1, S1′<br>S2, Sc1, Sa1, S1 | −Vdc |
| 7 | S2, Sc1, Sa2, Sa1, S1 | S2′, Sb2, Sb1, Sc1, S1′<br>S2′, Sc1, Sa2, Sa1, S1<br>S2, Sc1, Sa2, Sa1, S1 | −2Vdc |
| 8 | S2′, Sb1, Sc1, Sa1, Sa2, S1 | S2′, Sb2, Sc1, Sa1, Sa2, S1<br>S2′, Sb1, Sc1, Sa1, Sa2, S1 | −3Vdc |
| 9 | S2′, Sb2, Sb1, Sc1, Sa2, Sa1, S1 | No redundant state | −4Vdc |

## *3.2 Half-Equal Phase Pulse (HEPP) Method*

Since the EPP switching angle technique produces a multilevel waveform which is similar to a triangular wave, another method called the HEPP is used to coordinate the switching angles, resulting in a more extensive and stronger output waveform. Switching angle calculation for HEPP method is given in Eq. (2).

$$\alpha_i = \mathrm{i}\left[\frac{180}{m+1}\right] \tag{2}$$

where, $\mathrm{i} = 1, 2, 3 \cdots \left[\frac{m-1}{2}\right]$ m = number of levels.

For the nine level i.e., m = 9, the four main switching angles $\alpha_1$, $\alpha_2$, $\alpha_3$ and $\alpha_4$ using HEPP method are $18^0$, $36^0$, $54^0$ and $72^0$ respectively.

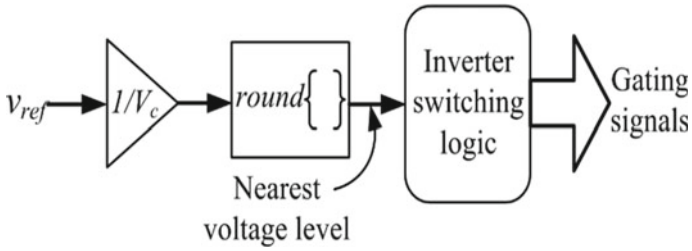



**Fig. 3** Switching pulses using RMC

### *3.3 Round Modulation Control (RMC)*

The RMC which is also known as nearest level control, the nearest staircase voltage is generated by RMC, which lowers the percent THD and improves power quality. The desired sinusoidal reference ($V_{ref}$) is compared to the corresponding inverter output voltage level to produce a nearest voltage level in this step. The '$V_{ref}$' magnitude depends on output voltage and number of levels. For a nine level CC-MLI, the magnitude of '$V_{ref}$' is chosen as 4 to generate the required levels. The switching pulse generation using RMC method is shown in Fig. 3.

### *3.4 THD Evaluation of CC—MLI for Nine—Level*

The general expression to evaluate THD for periodic output voltage waveform is given by (3).

$$\text{THD} = \sqrt{\left(\frac{V_{RMS}}{V_1}\right)^2 - 1} \tag{3}$$

where, '$V_{rms}$' and '$V_1$' are the RMS values of output phase voltage and the fundamental components respectively and their values are obtained using Eqs. (4) and (5).

$$V_{rms} = V_{dc}^{*}\sqrt{\frac{2}{\Pi}(\alpha_2 - \alpha_1) + 4(\alpha_3 - \alpha_2) + 9(\alpha_4 - \alpha_3) + 16(\alpha_4 - \alpha_3)} \tag{4}$$

$$V_1 = \frac{4V_{dc}}{(\pi\sqrt{2})}[(\cos\alpha_1 + \cos\alpha_2 + \cos\alpha_3 + \cos\alpha_4)] \tag{5}$$

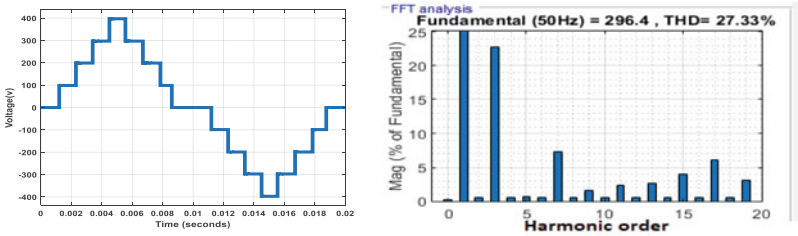
The evaluation of output phase voltage THD given in Eq. (6) for the proposed nine-level inverter can be obtain by substituting Eqs. (4) and (5) in (3).

$$\text{THD} = \sqrt{\frac{\Pi}{4}\frac{.(\alpha_2 - \alpha_1) + 4(\alpha_3 - \alpha_2) + 9(\alpha_4 - \alpha_3) + \left(\frac{\pi}{2} - \alpha_4\right)}{(cos\,(\alpha_1) + cos\,(\alpha_2) + cos\,(\alpha_3) + cos\,(\alpha_4))^2} - 1} \quad (6)$$

## 4 Simulation Results and Discussion

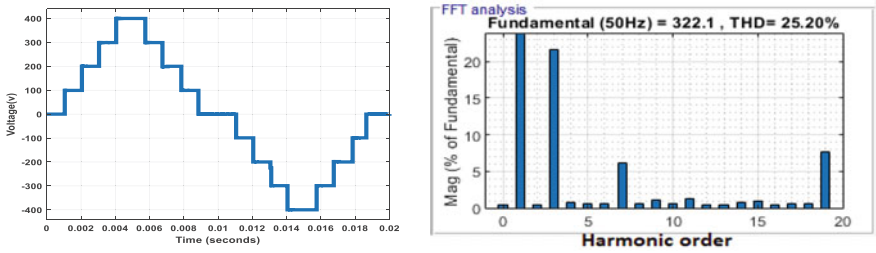
The simulation of proposed CC-MLI is carried out and analyzed for a nine-level operation using MATLAB/Simulink by applying the input of DC voltage ($V_{dc}$) of 100 V with a standard 50 Hz frequency for the EPP, HEPP and RMC methods and results are presented with respective THD bar graph obtained using FFT tool.

Figure 4 and it is observed that the Vrms is 221.08 and the THD of the output voltage is 27.33%. Figure 5 depicts the CC-MLI output voltage and THD analysis using EPP method for the nine-level the CC-MLI output voltage and THD analysis using HEPP method and it observed that the $V_{rms}$ is 244.94 and the THD of the output voltage is 25.2%. Further, the simulation results of CC-MLI using RMC method shown in Fig. 6. It can be noticed that the $V_{rms}$ is 270.89 and the THD of the output voltage is 9.38%. The theoretical values of the phase voltage THD for the



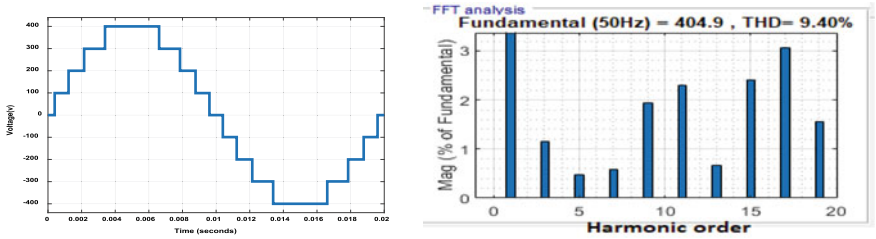

(a) (b)

**Fig. 4** CC-MLI results using EPP method **a** Output waveform **b** THD analysis




(a) (b)

**Fig. 5** CC-MLI results using HEPP method **a** Output waveform **b** THD analysis




(a) (b)

**Fig. 6** CC-MLI results using RMC method **a** Output waveform **b** THD analysis

**Table 3** Comparison of THD using EPP, HEPP and RMC methods

| Modulating technique | Switching angle (degree) | | | | %THD (Theoretical) | %THD (Simulation) | $V_{rms}$ (Volts) |
|---|---|---|---|---|---|---|---|
| | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ | $\alpha_4$ | | | |
| EP | 20 | 40 | 60 | 80 | 25.59 | 27.33 | 221.08 |
| HEP | 18 | 36 | 54 | 72 | 22.04 | 25.21 | 244.94 |
| RMC | 7.18 | 22.02 | 38.68 | 61.84 | 7.99 | 9.40 | 270.89 |

**Table 4** %THD of EPP, HEPP and RMC methods

| Sl. no | Level | EPP (%) | HEPP (%) | RMC (%) | Vishwanath et al. [21] (%) |
|---|---|---|---|---|---|
| 1 | 3 | 77.10 | 67.90 | 31.08 | 38.56 |
| 2 | 5 | 53.46 | 45.94 | 17.69 | 26.91 |
| 3 | 7 | 30.98 | 31.29 | 11.70 | 21.32 |
| 4 | 9 | 25.37 | 22.06 | 9.40 | 13.68 |

CC-MLI using EPP, HEPP and RMC methods along with the switching angles are listed in Table 3.

Comparisons of THD's for the EPP, HEPP and RMC methods are compared with sine modulation [21] control methods and the results are listed in Table 4 to find out best or optimal control method for the power quality improvement.

From Table 4, it can be noticed that the RMC method will give better THD when compared to EPP, HEPP and sine modulation control methods for any level of CC-MLI. Further, the fault tolerant analysis of CC-MLI is carried out with single power switch failure.

## *4.1 Fault Tolerance of Criss-Cross MLI Using RMC Method*

This section describes fault analysis of CC-MLI with RMC method for a single power switch failure. The simulation results of CC-MLI configuration are listed in

**Table 5** FFT analysis of CC-MLI for single faulty power switch

| Faulty switch | Missing voltage levels | %THD | | Voltage levels | Vishwanath et al. [21] (%) |
|---|---|---|---|---|---|
| | | During fault | After reconfigured | | |
| S1 | −Vdc, −2Vdc, −3Vdc, −4Vdc | 46.47 | 17.69 | 5 | 26.91 |
| S1′ | +Vdc, +2Vdc, +3Vdc, +Vdc | 46.46 | 17.01 | 5 | 26.91 |
| Sa1 | +4Vdc, −4Vdc | 21.59 | 15.71 | 7 | 21.32 |
| Sa2 | +4Vdc, −4Vdc | 16.52 | 15.06 | 7 | 21.32 |
| Sb1 | +4Vdc, −4Vdc | 15.71 | 15.04 | 7 | 21.32 |
| Sb2 | +4Vdc, −4Vdc | 14.78 | 13.78 | 7 | 21.32 |
| S2 | +3Vdc, +4Vdc, −Vdc | 86.26 | 17.67 | 5 | 26.91 |
| S2′ | +Vdc, +2Vdc, +3Vdc, −Vdc, −2Vdc, −4Vdc | 81.71 | 16.83 | 5 | 26.91 |

Table 5 by considering a faulty power switch $S_1$. It can be noticed that CC-MLI output has only positive levels and its %THD is 46.47% which is not acceptable due to poor power quality. Further, the power quality of CC-MLI has been improved using suitable redundant states of $-V_{dc}$, $-2V_{dc}$ and $-3V_{dc}$ voltage levels. Since there is no redundant state for $-4V_{dc}$ level, the output voltage will be asymmetric in nature and its %THD will be 64.53% which is even more when compared to during fault. Therefore, further by reconfiguring the switch states without considering $+3V_{dc}$, $+4V_{dc}$ and $-3V_{dc}$, $-4V_{dc}$ voltage levels, symmetric voltage will be obtained and the corresponding %THD of CC-MLI is listed in Table 5. It can be seen that the CC-MLI output consists of five-levels and its %THD is 17.69%.

The FFT analysis of CC-MLI for single faulty power switches i.e., $S_1$, $S_1^1$, $S_{a1}$, $S_{a2}$, $S_{b1}$, $S_{b2}$, $S_2$ and $S_2^1$ are listed in Table 5 with and without reconfiguration of CC-MLI redundant states. When compared to existing sine based modulation techniques [18–21], the proposed CC-MLI with RMC method has improved power quality. It is noticed that, %THD for a 5-level and 7-level inverters are 26.91 and 21.32% for the sine modulation techniques respectively for the modulation index of 0.8 [21], whereas the RMC method gives a maximum of 17.01 and 15.71% as represented in Table 5 for a 5-level and 7-level inverters.

## 5 Conclusion and Future Scope

The single-phase nine-level CC-MLI topology is assessed in terms of power switches, %THD and voltage levels using EPP, HEPP and RMC based switching angle control

techniques. From the Simulink results of CC-MLI, it can be observed that RMC method has better power quality in terms of %THD over EPP and HEPP methods for any number of levels. The %THD of RMC method is obtained as 9.37%, whereas %THD of EPP and HEPP methods are obtained as 25.37 and 22.06% respectively for nine-level CC-MLI. Further, the simulated %THD values of nine-level CC-MLI using EPP, HEPP and RMC methods are validated with the theoretical evaluation of THD. In addition, the CC-MLI using RMC method is tested under normal and single power switch failure conditions to confirm the CC-MLI as a fault tolerant topology. Therefore, due to modular circuit, reduced number of power switches and fault tolerant capability, the proposed CC-MLI configuration with RMC method is suggested for EV applications. The future scope of proposed CC-MLI topology can be integrated to AC motor drives to drive the EV.

**Acknowledgements** The authors would like to express their gratitude to the Hon'ble VC and RUAS management for providing all of the necessary facilities for this Research study.

## References

1. Siahbalaee J, Sanaie N (2021) Comparison of conventional and new cascaded multilevel inverter topologies based on novel indices. ISA Trans
2. Salem A, Van Khang H, Robbersmyr KG, Norambuena M, Rodriguez J (2020) Voltage source multilevel inverters with reduced device count: topological review and novel comparative factors. IEEE Trans Power Electron 36(3):2720–2747
3. Rao SN, Praveen Kumar V, Veerabhadra (2021) An enhanced boost Z-source inverter topology for electrical vehicle applications. In: Advances in energy technology, Lecture notes in electrical engineering, vol 766. Springer, Singapore. 978-981-16-1475-0, 500101_1_En (48)
4. Gupta KK, Ranjan A, Bhatnagar P, Sahu LK, Jain S (2015) Multilevel inverter topologies with reduced device count: a review. IEEE Trans Power Electron 31(1):135–151
5. Rao SN, Kumar DA, Babu CS (2013) New multilevel inverter topology with reduced number of switches using advanced modulation strategies. In: International conference on power, energy and control (ICPEC). IEEE, pp 693–699
6. Vijeh M, Rezanejad M, Samadaei E, Bertilsson K (2019) A general review of multilevel inverters based on main submodules: structural point of view. IEEE Trans Power Electron 34(10):9479–9502
7. Iqbal A, Siddique MD, Reddy BP, Maroti PK, Alammari R (2020) A new family of step-up hybrid switched-capacitor integrated multilevel inverter topologies with dual input voltage sources. IEEE Access
8. Arun N, Noel MM (2018) Crisscross switched multilevel inverter using cascaded semi-half-bridge cells. IET Power Electron 11(1):23–32
9. Rao SN, Kumar DV, Babu CS (2018) Implementation of multilevel boost DC-link cascade based reversing voltage inverter for low THD operation. J Electr Eng Technol 13(4):1528–1538
10. Hegde M, Rao SN, Indira MS (2021) Modified cascaded reversing voltage multilevel inverter using optimal switching angle technique for photovoltaic applications. Lecture notes in electrical engineering, vol 661. Springer, Singapore
11. Tolbert LM, Peng FZ, Habetler TG (1998) Multilevel inverters for electric vehicle applications. In: Power electronics in transportation (Cat. No. 98TH8349). IEEE, pp 79–84
12. Ronanki D, Williamson SS (2018) Modular multilevel converters for transportation electrification: challenges and opportunities. IEEE Trans Transp Electr 4(2):399–407

13. Poorfakhraei A, Narimani M, Emadi A (2021) A review of multilevel inverter topologies in electric vehicles: current status and future trends. IEEE Open J Power Electron 2:155–170
14. Vemuganti HP, Dharmavarapu S, Ganjikunta SK, Suryawanshi HM, Rub HA (2021) A survey on reduced switch count multilevel inverters. IEEE Open J Industr Electron Soc
15. Sulake NR, Pranupa S, Indira MS (2019) Cross connected source based reduced switch count multilevel inverter topology with fault tolerance. In: Global conference for advancement in technology (GCAT). IEEE, pp 1–7
16. Rao SN, Kumar DA, Babu CS (2018) Integration of reversing voltage multilevel inverter topology with high voltage gain boost converter for distributed generation. Int J Power Electron Drive Syst 9(1):210–219
17. Rao SN, Kumar P (2018) Comparative analysis of capacitor clamped and step-up switched capacitor multilevel inverters with self-voltage balancing. J Adv Res Dyn Control Syst 10(9):1841–1855
18. Manjunatha BM et al (2020) An optimized multilevel inverter topology with symmetrical and asymmetrical DC sources for sustainable energy applications. Eng Technol Appl Sci Res 10(3):5719–5723
19. Rao SN, Kumar BM, Pranupa S (2018) Three phase diode clamped multilevel DC link inverter with multi reference modulation techniques. J Adv Res Dyn Control Syst 10(9):1793–1805
20. Rao SN, Kumar DA, Babu CS (2018) Implementation of cascaded based reversing voltage multilevel inverter using multi carrier modulation strategies. Int J Power Electron Drive Syst 9(1):220–230
21. Vishwajith N, Rao SN, Sachin S (2020) Performance analysis of reduced switch ladder type multilevel inverter using various modulation control strategies. J Phys Conf Ser 1706(1):1–11

# An Energy-Efficient Load Balancing Approach for Fog Environment Using Scientific Workflow Applications


**Mandeep Kaur and Rajni Aron**

**Abstract** Fog computing seeks the attention of researchers by bringing a revolution in the Internet of Things (IoT). Fog computing emerged as a complement to cloud computing. It extends cloud services to the network edge and processes large and complex tasks near end users. Furthermore, fog computing can help process workflow tasks on its nodes only rather than sending them to the cloud, which helps to reduce the time consumed to request and process at the cloud layer. Scientific Workflow is used to represent data flow in scientific applications, which are very time-critical. This paper has proposed an energy-efficient load balancing approach for fog computing to reduce energy consumption in scientific workflow applications. The proposed algorithm works to reduce energy consumption in fog nodes by equal distribution of workload in fog resources. Genome and SIPHT workflow applications have been considered to evaluate in iFogSim.

**Keywords** Energy-efficient · Fog computing · Load balancing · Resource utilization · Scientific workflows

## 1 Introduction

Fog computing contains sensors, actuators, gateways, and other computing devices in its layered structure, which helps to store and process end-user requests at their end only. CISCO introduced fog computing to support the end-users facing obstacles while accessing the cloud data centers [1]. As its name indicates, fog is near the

M. Kaur (✉)
Lovely Professional University, Jalandhar, India
e-mail: k.mandeep@chitkara.edu.in

*Present Address:*
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

R. Aron
SVKM's Narsee Monjee Institute of Management Studies (NMIMS) University, Mumbai, Maharashtra, India
e-mail: rajni@nmims.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_16

end surface where all the Internet of things communicates and generates data. The amount of data increases daily, which needs proper storage and processing. Hence fog computing provides a layer near to end devices to cope with the high latency problem. Fog computing also helps to implement scientific workflow tasks. Workflow systems help manage different resources, which can be used in fog computing to manage scientific workflow tasks. Workflows are also defined as Direct Acyclic Graphs (DAG), which contain vertices and edges, where vertices denote different tasks to be executed, and edges show the relationship between these tasks [2].

Workflow applications like scientific tasks, face recognition, and sentiment analysis are complex tasks that increase complexity in the fog environment. Workflows contain dependent tasks, in which firstly available resources are found, and then tasks are assigned for execution. Due to the complexity of workflow tasks, there can be wastage of resources, resulting in more energy consumption [3]. Workflow scheduling is considered an NP-complete problem, which deems time and cost parameters while running the tasks [4]. With the distributed nature of fog computing, fog nodes are deployed near the end devices. Here are few examples of scientific workflows: Cybershake, LIGO, Sipht, Genome, Montage. Sipht and Genome workflows have been considered for evaluation of the proposed approach. Sipht is used to detect replicates of all bacteria in the national center. It helps to collect biological information [5, 6]. The genome can be any data related to microbiological resistance, pathogen's identity, genetic information [7, 8].

## *1.1 Load Balancing at Fog Layer*

In order to implement load balancing in workflow tasks also means conserving energy in fog resources. If the workflow tasks are unevenly distributed in the fog nodes, then the resource requirement may be more, or fewer resources could be utilized. So, to conserve energy consumed by fog resources, load balancing is a must in the fog computing layer. The task of the load balancer at the fog layer is to balance the distribution of workload in all the fog resources equally. Workload distribution can help for the efficient utilization of resources [3, 5, 9].

## *1.2 Our Contribution*

This article contributes the following:

1. It proposes fog computing architecture for maximum resource utilization in scientific workflow applications.
2. Proposed load balancing approach for a fog computing environment that aims to reduce energy consumption in fog resources while executing large and complex scientific workflow tasks.

The remaining article has been organized as follows. Section 2 reviews the existing literature in fog computing. Section 3 proposes Fog computing architecture implementing load balancing for maximum resource utilization. This section also contains the proposed EE-LB algorithm that has been evaluated to obtain simulation results shown in the next Sect. 4. The last Sect. 5, concludes the article and provides future scope.

## 2 Literature Review

This section contains the review of literature containing dynamic resource allocation and load balancing in workflows. There are many types of research works that provided different techniques for scheduling the workflows, but load balancing is still needed to explore. Literature review has been classified into two categories that are described below:

### *2.1 Resource Allocation in Fog Computing*

Li [3] proposed a load balancing based workflow scheduling model for resource allocation in cloud environment. Their proposed system model reduces response time and energy consumption in executing scientific workflow applications. The proposed workflow scheduling approach is based on the shortest path technique. The authors developed a social media application and performed a live video application. They created social media application and considered live video application of workflows to implement their proposed scenario. Naha [5] proposed a linear-regression method for energy-aware resource allocation. The authors minimize failure occurred due to energy constraints in the fog environment. Furthermore, the authors proposed an energy-efficient framework. Along with this, an energy-aware framework has been proposed to execute different applications in fog. The proposed approach has been compared with other existing techniques, reducing execution and processing time.

Rehman et al. [10] proposed a "Dynamic Energy Efficient Resource Allocation strategy (DEER)" for maximum resource utilization by implementing load balancing in fog environment. The proposed approach has been executed in simulation environment, and the obtained results are compared with other approach in terms of cost and energy consumption. It has been obtained that energy consumption has been improved by 8.67%, and computational cost by 16.77%. Xu et al. [11] proposed "Dynamic Resource Allocation Strategy (DRAM)" for fog computing to obtain maximum load balancing in fog environment. The major steps involved in DRAM are: fog operation partitioning, detecting available nodes, dynamic resource allocation to local and global users. Maximum resource utilization has been obtained by implementing load balancing, but energy consumption of nodes in fog environment is not given much attention.

### *2.2 Load Balancing in Fog Computing*

Rizvi [4] has reduced computational cost and execution time while executing workflow scheduling policy, i.e., fair budget policy. The proposed policy has been evaluated using different workflow applications and compared their results with the execution time and energy consumption in other approaches. The authors also perform ANOVA test on their proposed strategies. Kaur et al. [12] proposed equal distribution of workload-based load balancing approach for the fog computing environment. The authors considered the cloud analyst tool to evaluate their proposed system and compared its obtained results with existing round robin and the throttled load balancing approach. The main motive of the proposed algorithm is to enhance resource utilization and reduce the implementation cost.

Kaur et al. [16] proposed energy-aware approach for load balancing in fog computing environment. The authors considered scientific workflow applications (Genome, Cybershake) for evaluating their proposed approach, and simulation results are obtained using iFogSim simulation environment. They only considered few fog nodes to evaluate their proposed approach. More the number of fog nodes more will be the energy consumption in them so, the proposed approach will not be able to handle more workloads with lesser number of fog nodes. Table 1 shows the comparison of existing approaches.

## 3 Proposed Fog Computing Architecture for Maximum Resource-Utilization

Fog computing act as a middle layer between IoT and cloud layer. The traditional fog computing architecture brings the cloud services from core to network edge [2]. The fog architecture provided in this article contains load balancing in the middle layer. This section includes fog computing architecture for workload balancing in workflows. In distributed fog environment, fog nodes are deployed near to the end-users. In proposed architecture, load balancing has been applied in the fog computing layer so that workflow tasks should be evenly distributed to all fog resources. The following Fig. 1 shows fog computing architecture:

Figure 1 shows fog computing architecture containing three layers, i.e., end-users, fog layer and cloud layer. End user layer is connected to the nearby deployed fog nodes. Users submit their workflow tasks to the fog nodes. Fog nodes are having nano data centers, which store and process user requests locally. Each fog node is connected to the central controller, which controls all fog nodes. The central controller firstly schedules the tasks at the fog node's local queues. The tasks can be executed in any manner, i.e. First Come First Serve (FCFS), shortest path first etc. Then tasks are forwarded to the load balancer, which keeps track of all available and utilized fog nodes. The energy is assigned in the form of electric power consumed by various resources. The idle VMs also consume energy along with the overloaded

**Table 1** Comparison of various existing approaches

| Author | Year | Purpose of work | Type of network | Tool used | Application | Research gap |
|---|---|---|---|---|---|---|
| Naha et al. [5] | 2021 | Proposed energy aware approach based on multiple linear regression for load balancing in fog | Fog | CloudSim | Time-sensitive applications | Fog nodes can be clustered to enhance the performance of system |
| Mokni et al. [7] | 2021 | Propose a hybrid multi-agent approach for Cloud-Fog environment to schedule IoT tasks workflows | Cloud-Fog | CloudSim | IoT application | They do not consider energy consumption in cloud-Fog |
| Davami et al. [15] | 2021 | Proposed high-level architecture for scheduling of multiple work fows | Fog | Architecture tradeoff analysis method (ATAM) | Scientific workflow | Article does not consider load balancing to distribute equal workloads |
| Kaur et al. [16] | 2020 | Proposed Energy-aware load balancing approach | Fog-Cloud | iFogSim | Scientific workflows | Lesser number of fog nodes are considered |
| Hameed et al. [17] | 2021 | Proposed dynamic clustering approach for vehicular system | Vehicular ad-hoc network (Fog) | NS2 | Realistic vehicular network | Security of fog nodes can also consider |

VMs. Hence, to optimize energy consumption in fog nodes, proper load balancing is required, so that no VM remains underutilized, and no VM becomes overloaded. Load balancer equally distributes these tasks into the VMs. Workflow tasks are executed by fog nodes, and load balancing helps to improve utilization of all the resource in the fog layer. With the enhancement in resource utilization, the energy consumption of fog nodes can be reduced, which helps to reduce implementation cost. Fog layer is connected to cloud layer above it. Cloud layer have large data center having large storage and computing capacity that takes data from fog layer and store it for future use.

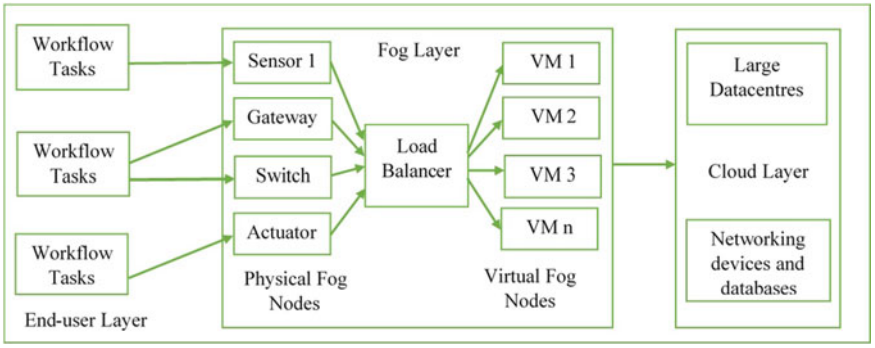



**Fig. 1** Fog computing architecture implementing load balancing for maximum resource-utilization

## *3.1 Energy-Efficient Load Balancing Approach (EE-LB)*

In fog computing environment, when large computational workflow applications are executed, then there is increase in demand of resources also. So, with the increased resources requirement there is more energy consumption. For efficient-energy consumption in fog environment there is a need for energy-efficient load balancing approach so that there can be no wastage of energy and resource utilization can be increased. In this section, proposed energy-efficient load balancing approach for fog computing environment has been provided. This section proposes a hybrid load balancing approach that is based on Simulated Annealing and water cycle optimization approaches. Optimization approaches used in this work are described as follows:

**Simulation Annealing Algorithm (SAA).** SAA has been used for intra-cluster mapping of tasks on fog nodes. The energy consumption in fog clusters has been analyzed using SAA. SAA has a large margin for error control so, it has been used to find global solution for scientific workflows.

**Water Cycle Optimization (WCO).** WCO works on the basis of natural water cycle process that has been used in this work when the optimized solution has not been found with SAA. WCO reduces the energy consumption and cost of intra-cluster resource mapping.

Both optimization techniques work in hybrid form to enhance the performance, and reduce the energy consumption and computational cost in fog nodes. Here is the pseudocode for proposed algorithm:

## 4 Result Analysis and Discussion

This section provides the results obtained by evaluating proposed approach in iFogSim environment. This section is divided into subsections i.e., parameter considered for comparison of obtained results, experimental requirements, and results are shown in graph form in the later subsection.

### *4.1 Parameters Considered*

The proposed approach has been evaluated based on two performance parameters i.e., computational cost, and energy consumption. These parameters are explained as follows:

**Computational Cost:** Computational cost can be calculated in terms of maintenance cost of fog nodes as well as cloud nodes. Sometimes only few nodes are utilized and others remains idle, but they also require maintenance. Hence maintenance cost can be calculated using following equation:

$$Cost = \sum C_r^{fog} + \sum C_r^{cloud} + R_{(c+f)} \tag{1}$$

Equation (1) used to calculate the computational cost in fog environment i.e., it considers as the sum of the total cost of resources of fog layer $\left(C_r^{fog}\right)$, total cost of resources of cloud layer $\left(C_r^{cloud}\right)$, and total available resources at fog and cloud layer $\left(R_{(c+f)}\right)$.

**Energy Consumption:** When tasks are assigned to resources for processing, the resources consume energy while executing these tasks. Sometimes in case of large computational tasks, few resources get more tasks to execute, and others remains idle. All resources consume energy whether they are in execution mode or idle [7]. Hence energy consumption can be calculated as follows:

$$EnergyConsumption(E_c) = E_{idle}^{fog,cloud} + E_{utilized}^{fog,cloud} + R_{max} \tag{2}$$

Equation (2) used to calculate energy consumption $(E_c)$ in fog environment that can be calculated as sum of energy consumption by all idle and utilized fog as well as cloud resources $\left(E_{idle}^{fog,cloud}, E_{utilized}^{fog,cloud}\right)$, and maximum number of available resources $(R_{max})$.

### 4.2 Experimental Requirement

Table 2 describes the experimental requirements used for executing proposed approach. Proposed EE-LB approach in has been evaluated by using iFogSim simulation tool and considered Genome and Sipht workflow applications. All the requirements have been explained as follows in Table 2.

### 4.3 Experimental Results

The experimental results obtained after evaluating proposed approach EE-LB are shown in the form of graphs by comparing with other existing approaches i.e. DEER [10], DRAM [11], EA-LB [16], SRFog [18] on the basis of few parameters that are described in the following Table 3.

Figure 2 shows the simulation results obtained by executing proposed Energy-efficient (EE-LB) and comparing it with the other existing approaches. It can be seen from the graphs that EE-LB approach outperforms the other existing approaches and reduces cost and energy consumption in evaluating considered workflows.

**Table 2** Experimental requirements

| Requirement | Value |
|---|---|
| Simulator | iFogSim |
| Operating system | Windows 10, 64 bit |
| Fog nodes | 20–140 |
| Workload | 100–1000 tasks |

**Table 3** Comparison of proposed approach (EE-LB) with other approaches

| Parameter | DEER | DRAM | EA-LB | SRFog | (EE-LB) |
|---|---|---|---|---|---|
| Simulation environment | CloudSim | CloudSim | iFogSim | Kubernetes | iFogSim |
| Number of Nodes | 500 to 2000 resources | 20–140 | 20 fog nodes | 2–7 nodes | 140 fog nodes |
| Number of workloads | N numbers | 500–1000 | 100–200 | 6 user requests | 100–1000 |
| Network type | Fog | Fog | Fog | Fog | Fog |
| Energy consumption (KJoules) | 1002.51397 | 756.345 | 935.321 | 1134.453 | 675.66 |
| Computational cost ($) | 7000–8000 | 8000–9000 | 5500–6000 | 4000–4500 | 2500–3000 |

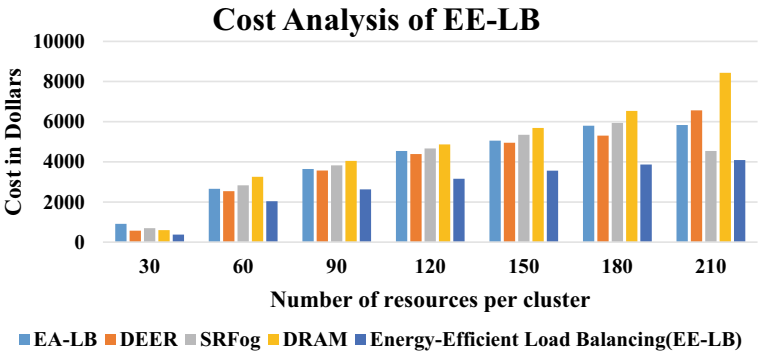
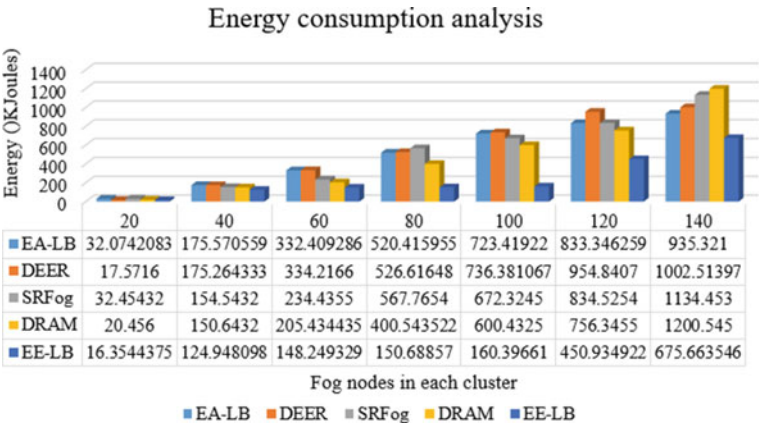



**Fig. 2** Cost analysis of EE-LB by comparing with other approaches




| | 20 | 40 | 60 | 80 | 100 | 120 | 140 |
|---|---|---|---|---|---|---|---|
| EA-LB | 32.0742083 | 175.570559 | 332.409286 | 520.415955 | 723.41922 | 833.346259 | 935.321 |
| DEER | 17.5716 | 175.264333 | 334.2166 | 526.61648 | 736.381067 | 954.8407 | 1002.51397 |
| SRFog | 32.45432 | 154.5432 | 234.4355 | 567.7654 | 672.3245 | 834.5254 | 1134.453 |
| DRAM | 20.456 | 150.6432 | 205.434435 | 400.543522 | 600.4325 | 756.3455 | 1200.545 |
| EE-LB | 16.3544375 | 124.948098 | 148.249329 | 150.68857 | 160.39661 | 450.934922 | 675.663546 |

**Fig. 3.** Energy Consumption analysis of EE-LB by comparing with other approaches

Figure 3 shows the simulation results obtained by evaluating considered workflows in proposed framework. It has been obtained from the graphs that with increase in number of fog nodes energy consumption also increase. Proposed approach EE-LB reduces energy consumption compared to the other approaches.

## 5 Conclusion and Future Scope

Load balancing in scientific workflows is necessary to fully utilize the resources at fog layer. This article provides architecture for fog computing implementing load balancing in scientific workflow. Furthermore, this article reviews the existing load balancing and scheduling techniques in workflows and provides a comparison between them. Different types of existing scientific workflows have been considered

and evaluated in iFogSim by applying proposed approach EE-LB and results are compared with EA-LB, DEER, SRFog, DRAM. It has been observed that EE-LB reduces computational cost by 28%, and energy consumption by 35% as compared to the other approaches. In future, QoS parameters in fog environment needs to be explored more.

## References

1. Bonomi F, Milito R, Zhu J, Addepalli S (2012) Fog computing and its role in the internet of things. In: Proceedings of the first edition of the MCC workshop on Mobile cloud computing, pp 13–16
2. Ding R, Li X, Liu X, Xu J (2018) A cost-effective time-constrained multi-workflow scheduling strategy in fog computing. In: International conference on service-oriented computing. Springer, Cham, pp 194–207
3. Li C, Tang J, Ma T, Yang X, Luo Y (2020) Load balance based workflow job scheduling algorithm in distributed cloud. J Netw Comput Appl 152:102518
4. Rizvi N, Ramesh D (2020) Fair budget constrained workflow scheduling approach for heterogeneous clouds. Clust Comput 23(4):3185–3201
5. Naha RK, Garg S, Battula SK, Amin MB, Georgakopoulos D (2021) Multiple linear regression-based energy-aware resource allocation in the fog computing environment. arXiv preprint arXiv:2103.06385
6. De Maio V, Kimovski D (2020) Multi-objective scheduling of extreme data scientific workflows in Fog. Future Gener Comput Syst 106:171–184
7. Mokni M et al (2021) Cooperative agents-based approach for workflow scheduling on fog-cloud computing. J Amb Intell Hum Comput:1–20
8. Ahmad Z et al (2021) Scientific workflows management and scheduling in cloud computing: taxonomy, prospects, and challenges. IEEE Access 9:53491–53508
9. Singh SP (2021) An energy efficient hybrid priority assigned laxity algorithm for load balancing in fog computing. Sustain Comput Inform Syst: 100566
10. Rehman AU et al (2020) Dynamic energy efficient resource allocation strategy for load balancing in fog environment. IEEE Access 8:199829–199839
11. Xu X et al (2018) Dynamic resource allocation for load balancing in fog environment. Wirel Commun Mob Comput 2018
12. Kaur M, Aron R (2020) Equal distribution based load balancing technique for fog-based cloud computing. In: International conference on artificial intelligence: advances and applications 2019. Springer, Singapore, pp 189–198
13. Shahid MH, Hameed AR, Islam S, Khattak HA, Din IU, Rodrigues JJPC (2020) Energy and delay efficient fog computing using caching mechanism. Comput Commun 154:534–541
14. Kaur A et al (2020) Deep-Q learning-based heterogeneous earliest finish time scheduling algorithm for scientific workflows in cloud. Softw Pract Exp
15. Davami F et al (2021) Fog-based architecture for scheduling multiple workflows with high availability requirement. Computing 1–40
16. Kaur M, Aron R (2020) Energy-aware load balancing in fog cloud computing. Mater Today Proc
17. Hameed AR et al (2021) (2021) Energy-and performance-aware load-balancing in vehicular fog computing. Sustain Comput Inform Syst 30:100454
18. dos Santos P, Pedro J et al (2021) SRFog: a flexible architecture for virtual reality content delivery through fog computing and segment routing. In: IM2021, the IFIP/IEEE symposium on integrated network and service management

# An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets


**B. Pranesh, T. Nitin, Shree Charan, D. P. Tejash, and K. Mahantesh**

**Abstract** An efficient image representation and extracting discriminative features in compressed domain is attracting researchers in computer vision and pattern recognition to develop efficient algorithms to classify and annotate images in large datasets. In this paper, an ensemble model combining DCT, Wavelet and HOG is developed to represent an image in compresses domain. DCT is useful to find low frequency coefficients of an image which express visual features, further subjected to multi-resolution analysis using wavelets with an advantage of developing robust and geometrically invariant structured object visual features through spectral analysis and finally PCA is used to map lower dimensional feature space with the transformational matrix in which a set of observed data is infused highlighting the edge orientation along with histogram of oriented gradients as its feature vectors. For classification purpose, different distance measure techniques and machine learning algorithm is used to obtain average classification rate. Proposed ensemble model is demonstrated on Caltech-101 and Caltech-256 datasets and compared the results with several benchmarking techniques in literature.

**Keywords** Discrete cosine transforms (DCT) · Distance metrics (DM) · Histogram of oriented gradients (HOG) · Image classification · Logistic regression (LR) · Wavelets

## 1 Introduction

The innovation and research and development sectors are growing rapidly due to the digital collection of data. The development of the internet and digital media techniques increased the demand for data. Hence, classification is an essential and required process for searching and efficient indexing of data. This leads to the existence of data classification and detection techniques and significant interest in the research community. The advancement of the internet and image acquisition

B. Pranesh · T. Nitin (✉) · S. Charan · D. P. Tejash · K. Mahantesh
Department of Electronics and Communication Engineering, SJB Institute of Technology, Bengaluru, India
e-mail: nitint1999@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_17

techniques has resulted in a huge increase in digital image collections produced by research, educational, medical, industrial, and other applications. The standard approach for retrieving images from large datasets is to describe the image using terms and then use the keywords to retrieve the images from databases, but manually annotating the images takes a long time. Furthermore, due to the natural inability of keywords to bridge the semantic divide. For massive image collections that are often specific to human experience, context-sensitive, and incomplete, manually annotating images for a wide variety of images is a time-consuming and costly process. In the early 1980s, content-based image retrieval (CBIR) was implemented as an important solution to text-based approaches because they struggled to accommodate a range of task-dependent queries.

CBIR indexes images based on their visual quality [1], such as texture, color and shape. Pattern recognition methods are used in image retrieval systems to explain visual information with partial semantics, while the process of extracting images has remained predominantly statistical in nature. While dealing with vast content image databases, some sophisticated algorithms are designed to interpret color, shape, and texture features which fails to exhibit image semantics and depict limitations [2]. In this paper, the distance measures and machine learning algorithms are used to build computer programs that produce solutions and evolve over time to solve problems that cannot be solved using enumerative or calculus-based approaches.

## 2 Related Works

To accomplish the tasks for the classification of images [3], approaches associated with the depiction of low image functionality are used, including shape, size, color, etc. Many papers have been written on image processing using various ensemble approaches. It has many advantages and uses in a wide range of areas and recommended HOG strategies for image classification [4]. In the histogram of oriented gradients (HOG), the effects of normalizing are investigated [5]. The normalizing function has been incorporated into the gradient creation stage. PCA was first used to remove substantial features from a function space by Turk and Pentland. Pankaj and Wilscy [6] suggested a system for comparing PCA [7] and features for face recognition. PCA is a computational technique for converting a huge amount of linked data into a smaller amount of congruent data [8]. The feature vectors are fed into PCA, which uses feature vectors derived from HOG to rewrite HOG features in terms of new variables. The image will be first preprocessed to improve its quality, such as by standardizing and smoothing it [9]. The gradient histogram (HoG) and wavelet properties are utilized to make predictions using machine learning. Extraction techniques examine images and remove the most significant attributes representing different kinds of objects [10]. Features are used to attribute items to the class where they use as inputs for classifiers. The extraction process allows to minimize the amount of data having through analyzing certain characteristics of images which contain pertinent information or functionality which distinguish one pattern

from the other. The several variations of features available including shape, color and texture-based features. Texture is amongst the most crucial attributes for distinguishing objects or areas of image enhancement is the process. These coefficients are decomposed into multi-resolution sub-bands by wavelets [11], allowing spectral analysis to create geometrically invariant structural object visual features. DCT eliminates low-frequency components that convey image visual features, and wavelets decompose these coefficients through multi-resolution sub-bands, causing spectral analysis to generate geometrically symmetric structural object visual features [12]. Using PCA, a dynamic model is used to map a collection of observed data into a lower-dimensional spatial domain [6]. In general, very less amount of significant work has been noticed in literature, where Caltech 101/256 datasets have been used extensively used for demonstration and benchmarking algorithms with standard train-test procedures as mentioned in [15–17]. Finally, various distance measuring techniques are used in order to derive an overall consistency score for object categorization. We demonstrated the proposed methodology in Sect. 3.4 on two very popular and challenging datasets and achieved the highest classification thresholds in addition to other benchmarking approaches discussed in the literature. Factors to consider when assessing the accuracy of models and the conclusions drawn from them objectively are discussed.

## 3 Methods

This section includes the frequency domain (DCT) technique and the methods for mixed time–frequency domains by isolating the borders and gathering the full energy of the image. This section goes into the details of combining DCT and wavelets and later HOG to obtain semantic features, followed by feature extraction and classification techniques.

### *3.1 Discrete Cosine Transforms*

Since PCA is found to have weaknesses in terms of lacking class separability and localization [13]. In the majority of images, the signal intensity at low frequency is high, and based on visual properties; DCT effectively divides image spectral bandwidth. The frequency domain of a signal is changed. The length of the f(x) sequence for x = 0, 1, 2, … , N is indicated where N is the length of the sequence.

The DCT of an image f(x, y) is computed using Eq. (1):

$$C(u,v) = \frac{2}{\sqrt{mn}}\alpha(u)\alpha(v)\sum_{x=1}^{m}\sum_{y=1}^{n} f(x,y)cos\left[\frac{(2x+1)u\pi}{2m}\right]cos\left[\frac{(2y+1)v\pi}{2n}\right] \tag{1}$$

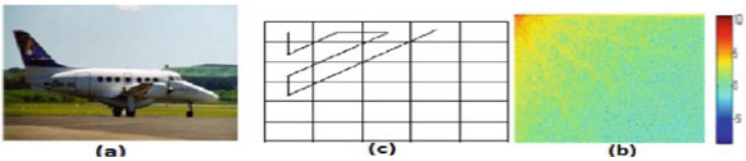



**Fig. 1** **a** original image; **b** color map of cipher degree of DCT; **c** scanning strategy of DCT coefficients

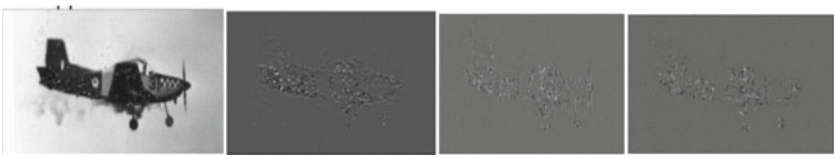


**Fig. 2** Approximation element, horizontal element, vertical element, diagonal element

where u = 1, 2, … , m and v = 1, 2, … , n are variables of scaling, as opposed to a 1D signal.

Figure 1(a) depicts a color aero plane image with a scale of 256 × 256 pixels, Fig. 1(b) shows the color map of the DCT's log magnitude, showing high energy compaction at the origin, and Fig. 1(c) shows the standard zigzag scanning scrutinize for DCT factors for an image.

### *3.2 Multi Resolution Analysis Using Wavelets*

dbN wavelets are the Daubechies' extreme phase wavelets. The Daubechies function extraction is discussed using scalar products with scaling signals and wavelets, similar to the Haar wavelet transform.

Daubechies combine the principles of multi-resolution analysis and pyramid coding techniques, then outlined their similarities and differences, and initially incorporated wavelet transform with the conventional definition of filter banks. when a source image is decomposed, four equal-sized output images are produced as shown in Fig. 2.

### *3.3 HOG*

HOG is a feature descriptor that is often used to extract features from image data [5]. The approximation element from the previous section is used as base image for feature extraction. Figure 3 gives HOG representation along with its original image data.

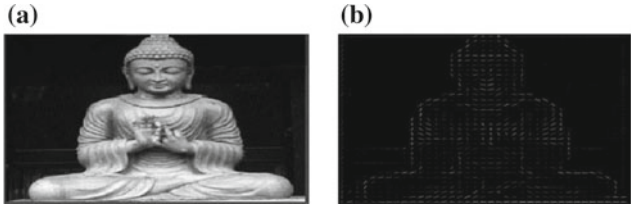



**Fig. 3** **a** Original image **b** HOG representation

The configuration of this work is somewhat different from that of Dalal and Triggs [14]. Linear-gradient voting into nine orientation bins ranging from 0 to 180°, 16 × 16 pixel blocks with four 8 × 8 pixel cells, and 16 × 16 pixel blocks with four 8 × 8 pixel cells are used in the selected configuration. The size of the descriptor can be computed using Eq. (2):

$$Descriptor\ size = b_s^2 n_b (v_{cells} - b_s + \sigma))(h_{cells} - b_s + \sigma) \tag{2}$$

The detector window has a resolution of 64 × 128 pixels. Converging the gradient masks [1 0 1] and $[1\ 0\ 1]^T$ with the raw image I yields the HOG descriptor by computing the 1-D p1oint derivatives Gx and Gy in the x and y directions. The Eq. (3) describes the gradient computation (being I the image and (i, j) the pixel coordinates):

$$G_x(i, j) = \frac{\partial I}{\partial x}(i, j) \quad G_y(i, j) = \frac{\partial I}{\partial y}(i, j) \tag{3}$$

The derivatives $G_x$ and $G_y$ are then applied to each pixel to calculate the gradient degree |G (x, y)| and path angle (x, y). The Eq. (4) indicates the magnitude of the gradient at a certain pixel:

$$M(i, j) = \sqrt{G_x^2(i, j) + G_y^2(i, j)} \tag{4}$$

The gradient degree of the direction histogram is simply used as a weighting factor. The gradient direction angle can be calculated using Eq. (5):

$$\theta(i, j) = arctan\left(\frac{G_y(i, j)}{G_x(i, j)}\right) \tag{5}$$

Depending on whether the angle is signed or unsigned, this histogram represents angles that are equally spaced between 0 and 180 or 0 and 360. As a result, the Eq. (6) is used for determining the kth bin of a histogram:

$$h_k = \sum_{i,j} M(i, j) 1[\Phi(i, j) = k] \tag{6}$$

In our case, k is a positive integer between 1 and 9. The vote will most likely take place between two bin centers at any given angle, if the previous angle division is contained in bins, so the voting is decided by bilinear interpolation between two bin centers mentioned in Eq. (7):

$$f(x|x_1, x_2) = f(x_1) + \frac{f(f(x_2) - f(x_1))}{x_2 - x_1}(x - x_1) \tag{7}$$

As the variations in gradient intensity show, some kind of illumination normalization is needed for a robust descriptor. Let the k-norm of v, with k ∈ 1, 2, and a small constant, as the k-norm of v, as a vector containing all the histogram for the block $||v||_k$. L1-norm, L1 squared-norm and L2-norm of a block $v$ descriptor vectors are mentioned in Eqs. (8), (9) and (10) respectively:

$$L1 - norm : v \rightarrow \frac{v}{||v||_2 + \in} \tag{8}$$

$$L1 - squared\ norm : v \rightarrow \sqrt{\frac{v}{||v||_2 + \in}} \tag{9}$$

$$L2 - norm : v \rightarrow \frac{v}{\sqrt{||v||_2^2 + \in^2}} \tag{10}$$

The values of v are constrained to 0.2 after an L2-norm, clipping, and re-normalization. This can be achieved by reducing, cutting and normalizing the product of the L2-norm. The tests revealed that L1-norm-squared, L2-norm, and L2-norm-Hys all behave equally and produce positive outcomes, but L1-norm degrades efficiency by 5%. If you do not normalize, the score is penalized by approximately 27%.

### *3.4 Classification*

The HOG features obtained from the previous section are used for classification purpose. The visual similarities are calculated using different distance measure techniques like Euclidean, Manhattan, KL Divergence and JS Divergence which can be computed using equations from Table 1. Further, logistic regression classifier is also used to compliment the results.

In the above Table 1, A and B are the two-dimensional points with coordinates (a1, a2, …), (b1, b2, …) respectively. Where A is considered as train data features and B is query data feature. Divergence methods calculate the difference between two probability distributions. The log of the event in P over the event likelihood in Q.

**Table 1** Mathematical representation of different distance measures

| Distance measures | Equations |
|---|---|
| Euclidean distance | $d(A, \mathrm{B}) = \lvert \mathrm{a} - \mathrm{b} \rvert = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_n - b_n)^2}$ |
| Manhattan distance | $M_{dist}(A, \mathrm{B}) = \lvert a_1 - b_1 \rvert + \lvert a_2 - b_2 \rvert + \ldots\ldots + \lvert a_n - b_n \rvert$ |
| Kullback–Leibler divergence | $\mathrm{KL(P} \,\lvert\rvert \mathrm{Q}) = \sum_x XP(x) * \log\left(\frac{P(x)}{Q(x)}\right)$ <br> $\mathrm{KL(P} \,\lvert\rvert\, \mathrm{Q) + KL(Q} \,\lvert\rvert\, \mathrm{P)}/_2$ |
| Jensen-Shannon divergence | $JS(P \lvert\rvert Q) = \mathrm{KL(P} \,\lvert\rvert \mathrm{M) + KL(Q} \lvert\rvert \mathrm{M)}/_2$ <br> $M = (P + Q)/_2$ |

KL divergence score is based on the idea that the odds of an occurrence from P vary greatly from those of the same event from Q. It may be used to differentiate between discrete and continuous probability distributions, which determine the integrality of outcomes rather than the number of discrete events that are expected to occur.

The JS divergence, a symmetrical normalized rating, is calculated using the KL divergence. It provides a smoothed and normalized variant of KL divergence, with scores ranging from 0 (identical) to 1 (extremely different), making it more functional as a metric by using the base-2 algorithm.

**Logistic Regression.** It is broadly utilized in ML and statistics. In both binary and multiclass grouping, it performs admirably. A LR model predicts P(Y = 1) as an X-function mathematically. LR employs the log odds ratio rather than probabilities to predict category, and an incremental multivariate technique rather than a least squares technique to build the final model. Assumption of logistic regression includes (1) For each instance, each independent variable should have a single value; (2) collinearity is expected to be minimal, although not necessarily totally independent; and (3) Irrelevant Independent Alternatives (IIA). The probabilities of preferring one class over another, despite the apparent presence of numerous unrelated and irrelevant choices are defined by IIA. In order to prevent multicollinearity in the model, the variables must be independent of one another. We need associated variables and a broad sample size for logistic regression. Using this approach, it is possible to predict the average classification rate.

## 4 Results and Performance Analysis

The results and analysis of the Caltech-101 and Caltech-256 datasets are covered in this section. Caltech—101 is a collection of 9,144 images from 101 distinct sorts of nature scenes (animals, butterflies, chandeliers, Garfield, cars, flowers, human faces, etc.) [15]. Most of the pictures are centered, occluded, influenced by corner

artifacts, and display high-intensity variability, making this dataset very problematic [16]. The Caltech-256 dataset, which includes image data from 256 item categories and a total of 30,607 pictures, is an augmentation of the Caltech-101 dataset [17]. It has higher intensity, clutter, item size, location, and posture variations, as well as more categories with at least 40 images per category and higher inter-class and intra-class variability. Table 2 gives classification rates of different distance metrics for 15 & 30 train per class/category respectively. Sample images of datasets are as shown in Fig. 4.

To evaluate the performance of the proposed methodology, the conventional experimental procedure outlined in [19] is considered, labeling the first 15 & 30 images in each category as training to generate feature vectors and the remaining as testing.

Using the image, a transformation strategy is performed to extract lower frequency coefficient information in the compressed DCT domain using 2D db wavelets. The approximated element is used as a base image for HOG feature extraction. Vectorization procedure (DCT Coefficients in zig-zag pattern) is initiated to generate the feature vectors. The trained feature dataset and the query feature vector use different distance measures and machine learning technique to obtain similarity. As a consequence, the performance of HOG descriptors in combination with transformation techniques are found to be efficient in comparison other descriptors, and the overall performance is increased.

Table 3 signifies that the proposed model's performance to a few state-of-the-art approaches mentioned in the relevant research. The histogram descriptor technique

**Table 2** Results of different distance metrics using HOG

| Distance metrics | Caltech 101 15 Train | Caltech 101 30 Train | Caltech 256 15 Train | Caltech 256 30 Train |
|---|---|---|---|---|
| | Classification rate in % | | | |
| Euclidean distance | 35.2246 | 38.1712 | 12.3379 | 14.5542 |
| Manhattan distance | 40.5264 | 44.2474 | 17.5962 | 20.5483 |
| KL method | 23.8639 | 25.5652 | – | – |
| JS method | 28.1980 | 28.1980 | – | – |

**Fig. 4** Sample images of Caltech 101 and 256 datasets

**Table 3** Analysis of different transformation techniques with logistic regression

| Methodology | Caltech 101 15 Train | Caltech 101 30 Train | Caltech 256 15 Train | Caltech 256 30 Train |
|---|---|---|---|---|
| | Classification rate in % | | | |
| Coslets [12] | 37 | 43.1 | 16.9 | 21 |
| Holub et al. [17] | 37 | 43 | – | – |
| Serre et al. [18] | 35 | 42 | – | – |
| HOG | 35.1316 | 43.3317 | 11.1489 | 15.1768 |
| DCT + HOG | 40.5596 | 46.9413 | 13.6233 | 15.2541 |
| Wavelet + HOG | 39.6711 | 46.1189 | 12.8623 | 15.9632 |
| DCT + Wavelet + HOG | 48.3920 | 53.0854 | 16.6244 | 19.0058 |
| PCA + Wavelet + HOG | 44.0699 | 48.0862 | 14.0952 | 17.2614 |

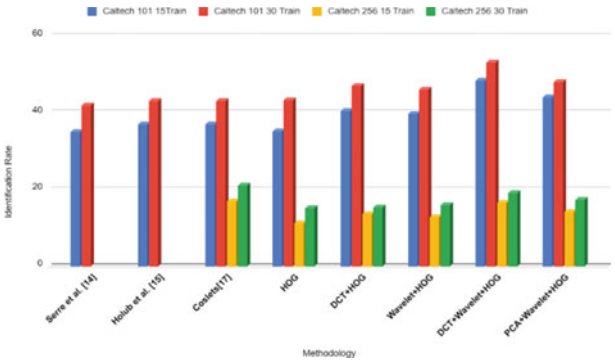



**Fig. 5** Comparative analysis of ensemble methodologies

with transformation techniques has turned out to be very interesting method for improving classification rate when compared to other works. On manifolds, Fig. 5 compares the performance of the proposed method to current dictionary learning techniques. In contrast to [12, 18, 20, 21], where grids of HOG features are proven to be very primitive and significant in outperforming the other strategies, the suggested method achieves a remarkable identification rate merely by employing a single visual descriptor. In comparison to conventional classifiers mentioned in [7, 19, 22], the proposed methodology with HOG has obtained leading classification rates of 48.4% & 53.08% for 15 & 30 images per category respectively, and was found consistent when compared with spatial pyramid feature technique [19] & method based on sparse localized features [23]. The proposed methodology using HOG outperformed [17] by 3.84% and produced very competitive results in comparison to the methodologies indicated in [6, 13, 24], which employed just 12,800 (i.e., no more than 50 images per category) images for testing.

## 5 Discussion and Conclusion

In recent years, most of the papers in literature have highlighted on combining transformation techniques for image representation but failed to preserve discriminative features. An ensemble model is proposed which combines DCT, Wavelet and HOG to preserve low frequency co-efficient, point singularities and edge orientations respectively from an image. Proposed ensemble method combined with different distance measure and logistic regression techniques are demonstrated on very large benchmark datasets for classifying images and has obtained competitive results by outperforming recent techniques such as coslet, bag of features, bag of words, spatial pyramid matching techniques. In future, the performance of our model can be examined for different distance measures, support vector machine, neural network architectures and deep learning techniques. The scope of this paper lies on real time object detection in cameras.

## References

1. Munjal M, Bhatia S (2019) A novel technique for effective image gallery search using content based image retrieval system, pp 25–29
2. Mahantesh K, Rao SA (2019) Content based image retrieval - inspired by computer vision & deep learning techniques. IEEE, pp 371–377
3. Liu B-D, Wang Y-X, Zhang Y-J, Zheng Y (2012) Discriminant sparse coding for image classification. In: 1988 international conference on acoustics, speech, and signal processing, 1988, ICASSP 1988, pp 2193–2196. https://doi.org/10.1109/ICASSP.2012.6288348
4. Deniz O, Bueno G, Salido J, De la Torre F (2011) Face recognition using histograms of oriented gradients. Pattern Recogn Lett 32(12):1598–1603
5. Zhang L, Zhou W, Li J, Li J, Lou X (2020) Histogram of oriented gradients feature extraction without normalization. In: 2020 IEEE Asia Pacific conference on circuits and systems (APCCAS), pp 252–255. https://doi.org/10.1109/APCCAS50809.2020.9301715
6. Pankaj DS, Wilscy M (2013) Comparison of PCA, LDA & gabor features for face recognition. Using Neural Netw 177:413–422
7. Shlens J (2003) A tutorial on PCA derivation, discussion and SVD
8. Rai P, Khanna P (2015) An illumination, expression, and noise invariant gender classifier using two-directional 2DPCA on real Gabor space. J Vis Lang Comput 26:15–28. SSN 1045-926X
9. Zhang H, Zhao L (2013) Integral channel features for particle filter based object tracking. In: Proceedings of the 2013, IHMSC 2013, vol 02, pp 190–193
10. Shyla NSJ, Emmanuel WRS (2021) Automated classification of glaucoma using DWT and HOG features with extreme learning machine. In: 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV), pp 725–730. https://doi.org/10.1109/ICICV50876.2021.9388376
11. Mahantesh K, Manjunath Aradhya VN (2014) An exploration of ridgelet transform to handle higher dimensional intermittency for object categorization in large image datasets, ICAICT, pp 515–521
12. Mahantesh K, Aradhya VNM, Niranjan SK (2015) Coslets: a novel approach to explore object taxonomy in compressed DCT domain for large image datasets. In: Advances in intelligent systems and computing, vol 320. Springer
13. Hemavathi N, Anusha TR, Mahantesh K, Manjunath Aradhya VN (2016) An investigation of Gabor PCA and different similarity measure techniques for image classification. In: Advances in intelligent systems and computing, vol 381. Springer, New Delhi

14. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR 2005), vol 1, pp 886–893. https://doi.org/10.1109/CVPR.2005.177
15. Yao B, Khosla A, Fei-Fei L (2011) Combining randomization and discrimination for fine-grained image categorization, CVPR 2011, pp 1577–1584. https://doi.org/10.1109/CVPR.2011.5995368
16. Liu B-D, Wang Y-X, ZhangY-J, Shen B (2013) Learning dictionary on manifolds for image classification. Pattern Recogn 46:1879–1890
17. Holub A, Welling M, Perona P (2005) Exploiting unlabelled data for hybrid object classification. In: NIPS workshop on inter-class transfer, Whistler, BC
18. Serre T, Wolf L, Poggio (2005) Object recognition with features inspired by visual cortex. In: CVPR, San Diego
19. Li F-F, Fergus R, Perona P (2004) Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories. In: 2004 conference on computer vision and pattern recognition workshop, pp 178–178
20. Sohn K, Jung DY, Lee H, Hero III AO (2011) Efficient learning of sparse, distributed, convolutional feature representation for object recognition, pp 215–223
21. Anusha TR, Hemavathi N, Mahantesh K, Chetana R (2014) An investigation of combining gradient descriptor and diverse classifiers to improve object taxonomy in very large image dataset, IC3I, pp 581–585. https://doi.org/10.1109/IC3I.2014.7019774
22. Griffin G, Holub A, Perona P (2007) Caltech-256 object category dataset. California Institute of Technology
23. Serre T, Wolf L, Poggio T (2005) Object recognition with features inspired by visual cortex. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR 2005), vol 2, pp 994–1000. https://doi.org/10.1109/CVPR.2005.254
24. Mahantesh K, Manjunath Aradhya V (2015) An exploration of neural networks & transformation techniques for image classification. Int J Adv Res Comput Sci Softw Eng 5(11)

# An Evaluation of Wireless Charging Technology for Electric Vehicle


**Vaishnavi Butale, Mohan Thakre, Vinayak Gaikwad, Yogesh Mahadik, and Tushar Jadhav**

**Abstract** The biggest hurdles for the globe throughout the coming decades will be climate change as well as the depletion of fossil fuels. One of the primary causes of climate change is an increase in the concentration of greenhouse gases. As a result, electric vehicles (EVs) have been identified as one of the possible solutions to reducing emissions as well as reliance on fossil fuels. Profitable commercialization and the faster adoption rate of EVs require faster, more cost-effective, and more reliable charging infrastructure. Conversely, several concerns, including such restricted driving range, a shortage with charging stations, battery destruction, and so forth, have been hampering the development of EVs. To alleviate EV range anxiety, it's also necessary to establish its highly developed charging infrastructure. To confront this same majority of these issues, a wireless charging system (WCS) could be one of the options for charging electric vehicles without a plug. The aim of this article should provide an overview of the main wireless power transfer techniques for charging a battery in an electric vehicle. Besides that, the obstacles and drawbacks of wireless charging technologies, and solutions to overcome these problems all are discussed extensively. Prospective ideas based on wireless electric vehicle charging systems, including "Vehicle-to-Grid (V2G)" as well as "In-wheel" WCS, also are addressed.

**Keywords** Conductive charging · Electric vehicle (EV) · Vehicle-to-grid (V2G) · Wireless electric vehicle charging system (WEVCS)

V. Butale · M. Thakre (✉)
K. K. Wagh Institute of Engineering Education and Research, Nashik, India
e-mail: mohanthakre@gmail.com

V. Gaikwad
MPSTM & E, NMIMS, University, Mumbai, India
e-mail: vinayak.gaikwad@nmims.edu

Y. Mahadik
Institute of Chemical Technology, Matunga, Mumbai, India

T. Jadhav
Government Polytechnic, Malvan, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_18

## 1 Introduction

Due to the rapid increase in oil prices as well as concern about the increasing emission of greenhouse gases, there is a growing interest in EVs around the world. EVs aid in the reduction of greenhouse gas emissions and the reduction of air pollution [1]. The battery, battery management system, power drive train, motor, controller, and other components comprise the majority of the EV. The battery is the EV's heart. As a result, battery protection is critical. During abnormal conditions, the insulation of the battery is more likely to be damaged. An HV battery protection system has been designed and explained in paper [2] to address the issue of insulation failure. Electric motors used throughout EVs include brushless DC motors (BLDC), DC series motors, three-phase induction motors, permanent magnet synchronous motors (PMSM), and switched reluctance motors (SRM). SRM is being investigated by industries and researchers due to its benefits such as rugged construction, high torque, effective speed control, and exceptional fault tolerance capability [3, 4]. EVs are becoming more popular around the world, but some concerns, including such limited driving range, a lack of charging stations within EV range, battery degradation, etc. existing roadblocks to EV advancement [5]. As a result, to address all of these issues, WEVCS may be a viable alternative technology for charging EVs without the need for a plug. In paper [6] the implementation of on-road charging is investigated to reduce the size of EV batteries as well as extend the driving range of EV.

There are two ways to charge an electric vehicle. 1) Charging by conductivity 2) Charging via wireless. Conductive charging requires physically connecting the EV to the electric power supply via a cable [7]. The publication [8] promotes the importance of a fast-charging system. The major drawbacks of a conductive charging system are its location within an EVs range, charging time, queuing time at the charging station, as well as potential traffic congestion. The wireless charging technique offers great ability to address EV charging infrastructure constraints. In these WEVCS emerging technologies, there is no direct relationship between both the EV as well as the power source. Power is wirelessly transferred to the EV [9, 10]. Wireless charging allows safe and convenient charging even if the vehicle is at rest or in motion. It also aids in the reduction of range anxiety. Wireless charging advancements include i) Inductive power transfer (IPT) ii) Resonant inductive power transfer (RIPT) iii) Capacitive power transfer (CPT) iv) Laser power transfer v) Microwave power transfer (MPT) vi) Permanent magnet power transfer (PMPT) [11–14]. Future upcoming technologies V2G [15] as well as IN Wheel wireless charging also discussed.

## 2 Schematic Review

Overview of conductive battery charger & international standards of charger related to power level has given by the author [7]. One of the conductive charging type "Fast charging" has been explained thoroughly in [8] including its various aspects such as

various technological development in this field, effect on a distributed system. Author [5] depicted that conductive charging requires a vehicle to get connected manually to the charging station hence there are chances of getting electric shocks. This can be overcome by WEVCS. Present WEVCS technologies were revised by the author [10, 13]. Basic principles, advantages, limitation were discussed in detail. Author [10], presented the proposed circuit topology for RIPT. Author [9], outlined a comprehensive overview of the conductive and inductive charging solution. Also stated that absence of wired connection between vehicle and charger in case of IPT make it safe technology for higher power charging application. Author [13] depicted promising technologies namely magnetic resonance power transfer and coupled magnetic gear including its key issue, challenges and the latest developments. Author [5] stated that the power transfer efficiency of the MPT system of wireless charging is not attractive as compared to other resonant WPT system. However, it can be improved by designing proper antennas. Author [14] stated the limitation of WPT regarding power transfer efficiency, installation cost etc. Various solution to improve power transfer efficiency such as by alignment of power pad, using different compensation networks, increasing operating frequency were explained in detail in paper [1]. Author [12] carried out a case study considering an IPT system with parallel-series compensation topology and simulation result obtained for induced voltage in the receiver due to magnetic flux variation by passing over the transmitter. Author [11], reviewed future upcoming concepts IN Wheel wireless charging system in which energy transfer efficiency is improved by reducing gap between transmitter and receiver coil. It has been discovered that with the rapid growth of EVs, there is a significant increase in power requirements from the distribution network. Vehicle-to-grid (V2G) was discovered to be a potential alternative technology for compensating for power demand. Author [15] emphasized various aspects of V2G such as benefits, limitation, challenges, economics of V2G etc. Conductive charging, methods of wireless charging technologies, limitation of WEVCS and its application were discussed in detail in upcoming section.

## 3 Conductive Charging

Throughout conductive charging, there seems to be a device connected in between the electrical supply and the EV. It comprises an AC to DC converter, a dc-dc converter, or a direct conversion from low-frequency AC to High-frequency AC (HFAC) with power factor circuitry (PFC). Boost converters are used to improve power factor (PFC). Based on their location, conductive chargers have been categorized as On-board or Off-board. Within the vehicle, the same onboard charger, rectifier, and battery current regulator circuit are located. It has two power stages: one for alternating current mains rectification and one for battery current regulation. In the off-board charger, rectifier, and battery current regulator are located outside of the vehicle, the external charging infrastructure includes a power converter. The power transfer levels for conductive chargers are given in Table 1.

**Table 1** Battery charger level according to SAEJ-1773[10]

| Level | Typical use | Power level | Utility interface |
|---|---|---|---|
| 1 | Home or office | 1.5 kW | 120 V,15A, 1φ supply |
| 2 | Personal /business entry point | 6.6 kW | 230 V,40A, 1φ supply |
| 3 | Recharging facility | 25–160 kW | 208-600 V, 3φ supply |

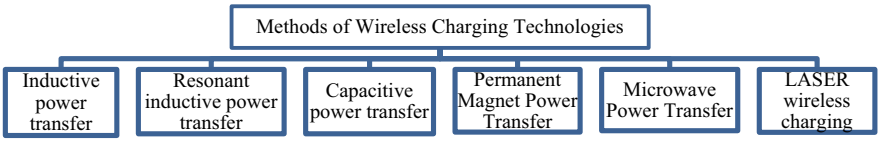



**Fig. 1** Classification of wireless charging technologies

Conductive charging limits may be specified as,

- There is a possibility of an electrical shock. There is therefore always a requirement for high operational safety.
- The system architectures for residential charging systems are complicated under SAE J1772 for EV conductive charging systems.
- The cost of electricity per mile exceeds the dynamic EV wireless charge.

Wireless charging has become a key field of research to reduce these limitations. A flexible, safe, and compact EV charge without direct human interaction is provided by wireless charging. The architectural design with WEVCS is indeed simple, but also less maintenance allows it for domestic and residential use.

## 4 Methods of Wireless Charging

In Fig. 1, classification of available wireless charging technologies has been provided. These are generally IPT, CPT, RIPT, PMPT, MPT, LASER wireless charging, etc.

### *4.1 Inductive Power Transfer (IPT)*

Inductive power transfer is indeed the process of transmitting power without a type of media that employs the mutual induction concept. Figure 2 shows a constructional diagram of IPT. Inductive power transfer is indeed the process of transmitting power without a type of media that employs the mutual induction concept. In this system, the transmitter coil has been embedded beneath the road surface, as well as the

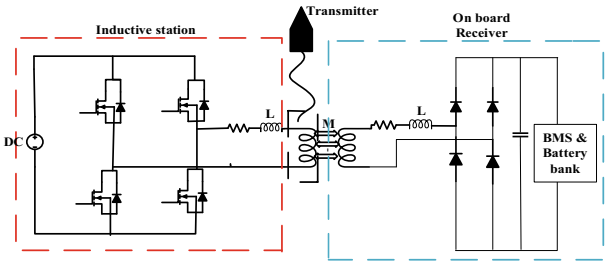



**Fig. 2** Circuit diagram of traditional inductive power transfer [11]

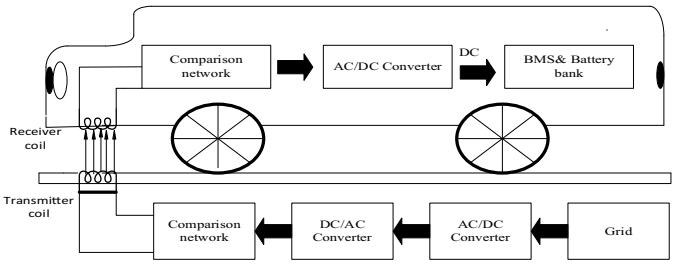



**Fig. 3** Block diagram of resonant inductive power transfer [11]

receiver coil has been embedded beneath the EV. AC from the main grid has been converted into high-frequency alternating current (HFAC) for power transfer between the transmitter and receiver coils using AC/DC and DC/AC converters. One such HFAC source is delivered to the transmitter coil. As a result, an electromagnetic field and therefore flux is developed. This flux comes into contact with the receiver coil, leading to voltage generation throughout the receiver coil.

This AC voltage would then be transformed into a steady direct current supply, which will then be supplied into battery packs to start charging those. In this system, magnetic coupling varies greatly. The magnetic coupling coefficient decreases as the air gap between two coils increases, and thus power transfer decreases. The large air gap between coils causes more leakage inductance and thus reduces magnetizing flux, necessitating the use of high magnetizing currents in a possible solution. As an outcome, as the power factor reduces, so does the overall system efficiency reduce [9].

## *4.2 Resonant Inductive Power Transfer (RIPT)*

Block diagram of RIPT is as shown in Fig. 3. RIPT is an advanced version of IPT and the most widely used wireless power transfer (WPT) technology. Its working principle is similar to that of IPT. But compared to conventional IPT, an additional

compensation infrastructure would be introduced to both the primary and secondary windings to establish resonant conditions & minimize losses incurred. The frequency at resonant condition given by Eq. (1)

$$Frps = \frac{1}{2\pi\sqrt{Lps * Lcs}} \tag{1}$$

whereas, Fr - resonant frequency of sec. and primary coil; Lcs & Lps -Resonant capacitor and self-inductance values of transmitter and receiver coil. If the resonant frequency of both coils is the same, efficiently maximum power transfer is possible.

### *4.3 Capacitive Power Transfer*

Figure 4 shows a block diagram of CPT power that can be transferred without incurring significant losses when passing through a barrier including such metal barriers. Through CPT, the power is transmitted from the main to the receiver winding through a coupling capacitor rather than coils or magnets as shown in Fig. 4. Power factor correction (PFC) integrated circuit applies the primary AC voltage to an H-bridge converter. The H-Bridge's HFAC is redirected via the coupling capacitor on the receiver end.

Whenever the circuit has been resonant, extra inductors have been included in series with such a coupling capacitor to decrease impedance between the transmitter and the receiver sides. Also, it facilitates the incorporation of soft-switching into electronic circuits [11]. With the help of filter circuitry and a rectifier, the AC voltage received at the receiver side is converted to DC for battery charging. Power transfer in CPT is affected by both the distance between two plates and the size of the coupling capacitor. Maximum power transfer is possible by increasing the frequency and decreasing the distance between two plates. CPT can transfer an electromagnetic field without an eddy current through a metal shield, resulting in low losses. Because the capacitor plates are loosely coupled and the coupling capacitance is low in the CPT system. The overall efficiency of the system will decrease. CPT is limited to

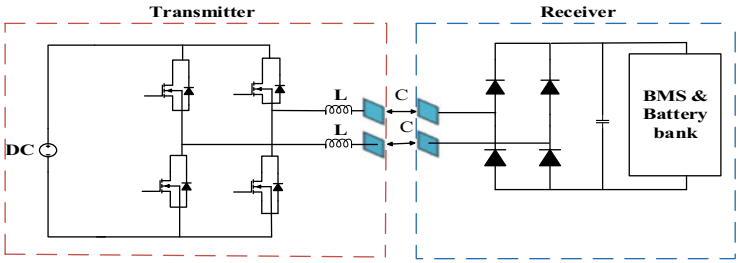



**Fig. 4** Circuit diagram of capacitive power transfer [11]

low-power applications due to the low density of energy stored in free space between metal plates.

### *4.4 Permanent Magnet Power Transfer (PMPT)*

Two synchronized permanent magnets have been positioned next to one another. The current is connected to the transmitter winding throughout order to create mechanical torque on the main permanent magnet. This same main permanent magnet begins to rotate as well as initiates torque just on supplementary permanent magnets such as through mechanical interaction. In this PMPT, the main magnet activities as a generator, and the supplementary magnet receive power, which would then be transported to the batteries through the power converter. One such technology resulting in low power transfer efficiency, complicated system design, as well as maintenance problems is unsuitable for EV wireless charging technologies [12].

### *4.5 Microwave Power Transfer (MPT)*

Receiving antenna (receiver) & transmitting antenna (transmitter) is mainly used in this technique of EV charging.

Figure 5 shows a block diagram of MPT. The transmitting antenna emits a microwave beam to transfer power from the transmitter to a receiver configured in the EV. Rectenna is yet another term for a receiving antenna. The Rectenna would then acquire the microwave beam released by the transmitting antenna. The collected beam would then be transformed into a direct current source, and this source has been supplied to the batteries to start charging.

Particularly in comparison to certain other resonant WPT systems, total power efficiency decreases. The efficiency of this system can be enhanced by formulating better antennas. The previous MPT architecture was using a frequency of 2.45 GHz [5]. It acknowledged that using high frequency enables the transmission of power over

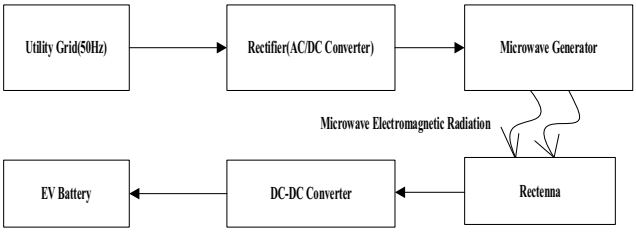



**Fig. 5** Block diagram of microwave power transfer

greater distances. Various automobiles can be charged that use the same MPT structure utilizing beam steering modern technologies. MPT can transfer power across such a wide air gap between the transmitter and the receiver. The issue with that same method would be that generating a high-frequency power signal in the MHz frequency range is indeed a considerable power electronic challenging issue. MPT is no longer as reliable or cost-effective as present technology. It was impossible to prevent such radiation without compromising functionality or range [14].

### *4.6 LASER Wireless Charging*

This technology will be useful for transmitting power over long distances but with limited efficiency. In this method of power transfer, the electric current is converted into a laser beam, which is then focused on a photovoltaic cell. This type of charging is also called power beaming as it refers to the transfer of power in the form of a laser beam. Laser power transfer is a complicated phenomenon. As a result, EV charging is not possible with this technology.

## 5 Application of Wireless Electrical Vehicle Charging System

### *5.1 Vehicle-to-Ground (V2G)*

As the number of electric vehicles increases, so does the electricity demand. There is a significant increase in power requirements from the distribution network. V2G technology is a system that allows for the controllable, bidirectional flow of electrical energy between an EV and the grid. And there is a flow of electric energy from the grid to the EV during the charging of an EV's batteries. During peak demand on the grid, this same grid requires energy; in this case, the circulation of electric power is from vehicle to grid.

In a plug-in, V2G electrical vehicle equipped via an onboard bidirectional charger as shown in Fig. 6(a), a device can access the V2G to produce power whenever the grid is experiencing peak load. EV has been charged with an ac signal plug socket during the off hours. This AC is converted to DC as well as supplied through a separated DC/DC converter to maintain a user's safety. BMS control and protective measures, as well as a bidirectional DC converter, are used to distribute the transformed DC to a battery. So if charging battery packs, this converter operates as a buck converter, and when discharging battery packs, it operates as just a boost converter. During the charging or discharging of such an EV, there has been manual ability to handle and also physical contact in this system, and there is a threat of electric shocks as well as accidents. The new smart V2G technique has been proposed to address

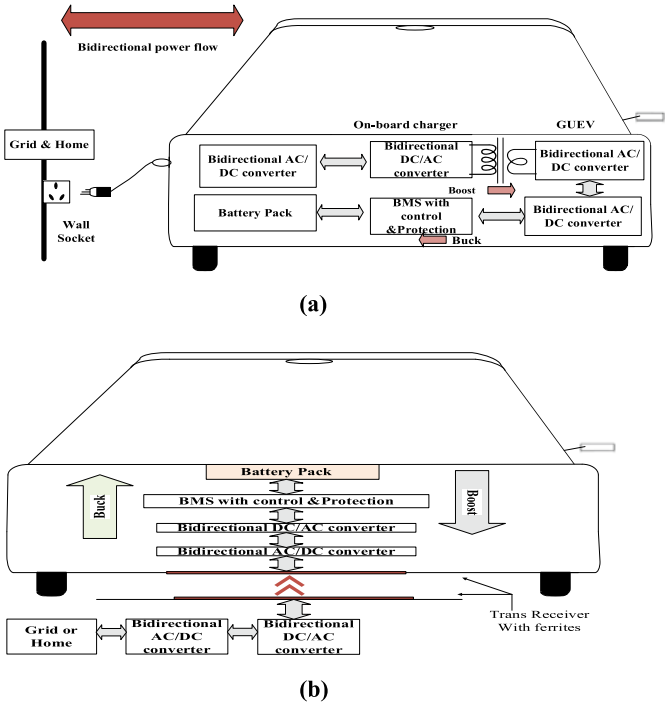



**Fig. 6** **a** Plug-in V2G [11]. **b** Wireless V2G [11]

every one of these limitations. This technique utilizes wireless primary transformers with bidirectional converter topologies. The primary of one such transformer was integrated into the parking surface, the secondary was installed beneath the vehicle, and the remaining bidirectional converters were also assembled in the vehicle body as per Fig. 6(b).

## *5.2 IN Wheel Wireless Charging System (IN-Wheel WCS)*

As observed in the previous topic, authors can deduce that the air gap, as well as coil arrangement, has the greatest influence on power transfer performance. Overall average gap distance for limited passenger vehicles between 150 and 300 mm, and might be more for heavier vehicles. The energy transfer performance deteriorates as the air gap rises, and as the air gap decreases, the energy transfers efficiency increases, allowing the air gap problems to be overcome. The IN-wheel WCS is intended for both dynamic as well as stationary applications. The primary difference between such a system and other WEVCS would be that the receiver has been configured inside the tire configuration instead of under the electric vehicles. So, if compared to

current static analysis and dynamic WEVCS, this same air gap here between source as well as receiver coils has been decreased.

Because of the smaller air gap, the coupling efficiency between the transmitter and the receiver is higher here. The higher efficiency of power transmission will give it a significant advantage over conventional quasi dynamic or dynamic wireless charging technology solutions.

Wireless transformer coils, a power source, as well as the internal structure of the tire are the main elements of this structure. To optimize static and dynamic IN-wheel WCS, such elements must be specially designed. Different receiver coils have been parallelly positioned inside of the tyre. The said arrangement has the advantage of activating the specific receiver coil that is in interaction with the transmitter. In the occurrence of horizontal misalignment, different receiver coils might be triggered at the same time, and all these coils would transfer power to the battery bank. A rectifier, a resonant capacitor, as well as a filter circuit, are all included in every other receiver coil.

## *5.3 Wireless Charging Technologies Difficulties*

Among the disadvantages of wireless charging technological advances seem to be,

- WPT has less power transfer efficiency, unlike plug-in charging.
- A wireless power transfer charging system is more expensive to install than a plug-in charging system.
- A huge issue with these technological advances is the potential for health hazards and the need for a fundamental limit on human health and the environment to radiofrequency radiation [14].

## *5.4 WPT Efficiency Enhancement Solutions*

To tackle above mentioned difficulties some solutions provided which are as follows,

**By Increasing Frequency.** Changing the operational frequency also allows us to make the scheme extra convenient and effective. The efficiency grows with frequency, but there is still a power level limit. Yvkoff suggested a 100–200 kHz frequency range system [1]. Yvkoff devised one expression, which is given as per Eq. (2),

$$TQ = wM/RO \tag{2}$$

whereas $TQ$ is transfer quality factor, $M$ is mutual inductance between transmitter and receiver, $\omega$ is resonant frequency & $R_0$ is equivalent resistance. To obtain

maximum efficiency $TQ$ should be large. It can be increased in 3 ways (i) Maximizing driving freq. (ii) reducing equivalent resistance (iii) increasing the mutual inductance.

**By Alignment of Power Pads.** To improve WPT efficiency, the transmitter and receiver coils must also be positioned. The effect of coil imbalance on efficiency has been presented in Sect. 4, IPT. The highest efficiency can be provided as per Eq. (3),

$$n_{max} = \frac{1}{1 + 2/k\sqrt{L_p * L_s}} \tag{3}$$

In the above Eq. (3), the coupling coefficient is given by $k = m/Lp * Ls$ & quality factor of the primary and secondary coil is given by Eq. (4),

$$Q_P = (\omega L_p/R_p) \& Q_S = (\omega L_S/R_S) \tag{4}$$

It's indeed straightforward from the above two solutions which enhance the quality factor, as well as coupling factor, improves the maximum achievable power transmission efficiency. A compensation infrastructure has been necessary to establish a resonant tank, compensate for inductance, as well as maximize power transfer. Compensation infrastructures have been introduced to both the transmitter and the receiver sides to make resonant characteristics of high-power transfer. Compensation infrastructure has been classified into four categories: (i) Series-Series (SS), (ii) Parallel-Parallel (PP), (iii) Series–Parallel (SP), and (iv) Parallel-Series (PS).

## 6 Conclusion

Several technologies, including conductive charging, IPT, CPT, PMPT, RIPT, and LASER charging, have been explained clearly. Among all of these techniques, RIPT is the most widely used technology today, as in this technique resonant condition is established with the help of compensation infrastructure and hence minimum losses incurred. CPT is common for low-power applications. The application of CPT is limited to low-power applications due to the low density of energy stored in free space between metal plates. Encouraging a Vehicle-to-Grid approach to developing dynamic wireless charging would then create a new generation of electrically powered vehicles with lower battery capacity and enhanced vehicle performance scope.

## References


1. Ahmad A, Alam MS, Chabaan R (2018) A comprehensive review of wireless charging technologies for electric vehicles. IEEE Trans Transp Electrif 4(1):38–63. https://doi.org/10.1109/TTE.2017.2771619
2. Thakre MP, Mahadik YV, Sardar N (2020) Architecture of an HV power battery protection devices for hybrid electric vehicles (HEV). In: IOP Conference Series: Materials Science and Engineering, vol. 1084, First International Conference on Circuits, Signals, Systems and Securities (ICCSSS 2020) 11th–12th, Tamil Nadu, India, December 2020
3. Chowdhary PK, Thakre MP (2020) MMC based SRM drives for hybrid EV with decentralized BESS. In: 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp 319–325, Coimbatore, India. https://doi.org/10.1109/ICECA49313.2020.9297508
4. Chowdhary PK, Thakre MP (2020) MMC-based SRM drives for hybrid-EV with decentralized BESS in battery driving mode. In: 2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), pp 1–6, Chennai, India. https://doi.org/10.1109/ICPECTS49113.2020.9337029
5. Liang X, Chowdhury MSA (2018) Emerging wireless charging systems for electric vehicles-achieving high power transfer efficiency: a review. In: IEEE Industry Applications Society Annual Meeting (IAS), pp 1–14, Portland, USA. https://doi.org/10.1109/IAS.2018.8544484
6. Patnaik L, Huynh PS, Vincent D, Williamson SS (2018) Wireless opportunity charging as an enabling technology for EV battery size reduction and range extension: analysis of an urban drive cycle scenario. In: 2018 IEEE PELS Workshop on Emerging Technologies: Wireless Power Transfer (Wow), pp 1–5, QC, Canada. https://doi.org/10.1109/WoW.2018.8450900.Montreal
7. Mude K, Dashora H, Bertoluzzo M, Buja G: From wired to in-moving charging of the electric vehicles. ISBN: 978-960-474-400-8
8. Thakre MP, Mahadik YV, Yeole DS, Chowdhary K (2020) Fast charging systems for the rapid growth of advanced electric vehicles (EVs). In: 2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), pp 1–6, Chennai, India. https://doi.org/10.1109/ICPECTS49113.2020.9336979
9. Khaligh A, Dusmez S (2012) Comprehensive topological analysis of conductive and inductive charging solutions for plug-in electric vehicles. IEEE Trans Veh Technol 61(8):3475–3489. https://doi.org/10.1109/TVT.2012.2213104
10. Sabki SA, Tan NML (2014) Wireless power transfer for electric vehicle. In: IEEE 8th International Power Engineering and Optimization Conference (PEOCO2014), pp 41–46
11. Panchal C, Stegen S, Lu J (2018) Review of the static and dynamic wireless electric vehicle charging system. Eng Sci Technol Int J 21:922–937. ISSN 2215-0986
12. Niculae D, Iordache M, Stanculescu M, Bobaru ML, Deleanu S (2019) A review of electric vehicles charging technologies stationary and dynamic. In: 2019, 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE), pp 1–4. Bucharest, Romania. https://doi.org/10.1109/ATEE.2019.8724943
13. Qiu C, Chau KT, Liu C, Chan CC (2013) Overview of wireless power transfer for electric vehicle charging. In: 2013 World Electric Vehicle Symposium and Exhibition (EVS27), pp 1–9. Barcelona, Spain. https://doi.org/10.1109/EVS.2013.6914731
14. Musavi F, Eberle W (2014) Overview of wireless power transfer technologies for electric vehicle battery charging. IET Power Electr 7:60–66
15. Thakre MP, Mahadik YV, Yeole DS (2020) Potentially affect of a vehicle to grid on the electricity system. In: IOP Conference Series: Materials Science and Engineering, vol 1084, First International Conference on Circuits, Signals, Systems and Securities (ICCSSS 2020) 11th–12th, Tamilnadu, India, December 2020

# Automated Dam Data Acquisition and Analysis in Real-Time


**Neelam Sanjeev Kumar, Gokul Chandrasekaran, and P. R. Karthikeyan**

**Abstract** The Mechanized Dam Information Procurement and Panic Reportage System was configuration to get, screen, and dissect, in real-time, basic wellbeing boundaries, for example, inflows, surges, door openings, pressure gages and level of the water kept up in the dam. This framework permits dams to be worked all the more securely and crisis intends to be all the more successfully organized and actualized with the crisis the executive's organizations. The model portrayed in this paper is the initial move toward our vision of improving wellbeing on dams by having a system of sensors sending defensive rigging and checking framework. The health care monitoring system also processed with the self-evaluation module which helps to identify the process and movement of the object. The real time received data also discussed in this research work. The wireless sensor plays an important role to locate the data and identify the drop in the transmission line. Overall, the dam data information is fetched and processed the IoT operation. The reports are analyzed and tabulated in this research work which identify the level of prediction and operation of the dam data. The energy consumption, delay, and percentage of reading performances are improved in the proposed method than the conventional works.

**Keywords** Circuit implementation · Data framework embedded · Mechanized Dam · Real time system · Wireless sensor network

N. S. Kumar (✉)
Department of Biomedical Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India
e-mail: neelamsanjeev1034@gmail.com

G. Chandrasekaran
Department of Electronics and Communication Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

P. R. Karthikeyan
Department of Electrical and Electronics Engineering, Velalar College of Engineering and Technology, Erode 638012, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_19

## 1 Introduction

An astute building site wellbeing framework would decrease wounds and fatalities by giving better security from mishaps, improving reaction times to mishaps, and giving increasingly careful information assortment that can be utilized to break down mishaps and close mishaps to forestall future events [1–3]. Automated Dam Data Acquisition and Alarm Reporting System (ADDAARS) gives data to chiefs continuously through a blend of radio and satellite telemetry, microwave, fiber optic, and web innovation to Improve the dispersion and utilization of labor, Provide constant information for safe dam and store activity, Reduce crisis reaction time and Provide an exact and predictable information assortment framework, Provide the devices to screen states of the dams and stores them progressively at the crisis activity focus and the regional operator center [4–6]. The microcontroller gathers, store and gets ready information for transmission through GPRS modem and Radio handsets. The GPRS modem gets information from the microcontroller and transmits it to the normal site [7].

Hydroelectric dams are intended to give a more naturally amicable type of producing power than consuming petroleum derivatives; be that as it may, the greater part of the downsides of hydroelectric dams is harming ecological reactions. As repositories are overwhelmed to make hydroelectric force they flood and crush here and their sections of land of important rural or natural life land. Now and again, jeopardized or compromised species are influenced [8].

What's more, one more thing we may utilize dam for shipping reason because of this there will be limit given in the event of vehicle weight, due to over-burdening dam might be harm, so with the assistance of burden cell, we can quantify heap of the vehicle and we can stop the harming of the dam [9, 10].

## 2 Related Works

Kumar et al. [11] proposed an integrated system for smart industrial monitoring system in the context of hazards based on the Internet of Things. In this work, the smart industrial devices were controlled by the IoT with effective operation. The process is performed with less execution time, but the power consumption of the entire design was high. Kumar et al. [12] presented a Novel Architecture of Smart Healthcare System on Integration of Cloud Computing and IoT. This work concentrated the smart health care system with biomedical application. The integration of the hardware processed with the appropriate managing modules which helped to process the data with less loss. However, the overall cost of the healthcare system was high. Manikandan et al. [13] proposed hash polynomial two factor decision tree using IoT for smart health care scheduling system. Hash function plays a vital role in the smart care system which helped to operate the confusion property. Due to the help of two factor healthcare system, the security of the healthcare system was improved.

But, the latency and accuracy were less which affected the system performances. Ellaji et al. [14] presented efficient health care systems using intelligent things using NB-IoT. Ammi et al. [15] proposed Customized blockchain-based architecture for secure smart home for lightweight IoT. Both the design was occupied more energy to monitor the health care system.

## 3 Proposed System

This framework permits dams to be worked all the more securely and crisis intends to be all the more adequately organized and actualized with the crisis the executives' offices ADDAARS gives data to heads continuously in form of a mix of wireless and satellite broadcasting, fiber optic and advent technology Internet innovation to Progress the dispersion and utilization of labor, Deliver the instruments to screen states of the dams and supplies progressively at the crisis activity focus and the provincial tasks habitats, the database programming additionally encourages the assessment of station conditions and information quality, just as gives a wide scope of reports to introduce both summed up and point by point data in printed structure. The transmitter and receiver sections are shown in Fig. 1 and 2.

### *3.1 Working Function*

Essential Physics of Pressure Sensing: Static Pressure. Weight, $P$, is characterized as power, $F$, per unit territory, $A$. The pressure sensor images is shown in Fig. 3 and

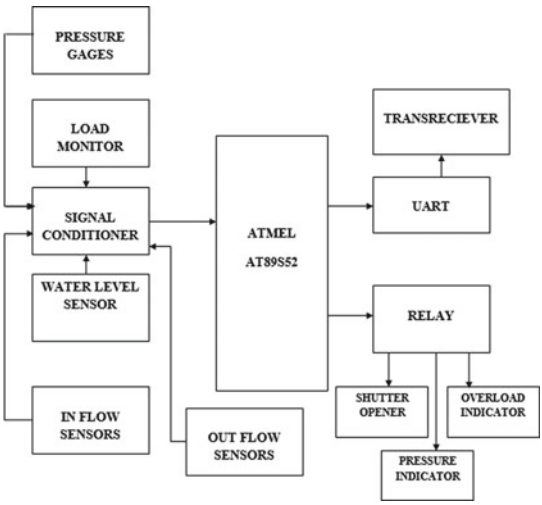



**Fig. 1** Transmitter section

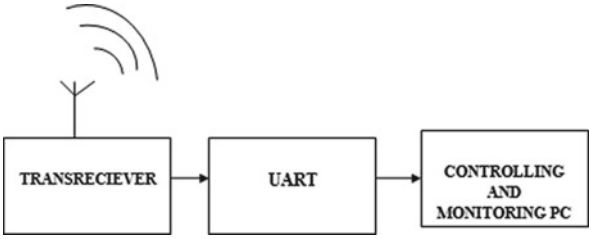



**Fig. 2** Receiving section

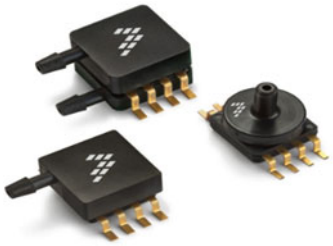


**Fig. 3** Images for pressure sensor

the overall circuit diagram is shown in Fig. 4. The power calculation is expressed in Eq. 1 and 2.

$$P = F/A \tag{1}$$

The estimation of weight is by and large connected with liquids, either fluids or gases. A compartment loaded up with a fluid (see Fig. 1) has a pressure (because of the heaviness of the fluid) at a point in the fluid of:

$$P = F/A = h * w \tag{2}$$

where:

$h =$ good ways from the surface forthright.

$w =$ weight of the fluid (most fluids are almost incompressible).

The pressure sensor helped to predict the pressure value which used in the dam acquisition process.

Standard of activity in Sensing components:

$$H = (1 - (P/Pref)0.190284) \times 145366.45\ ft \tag{3}$$

This condition is adapted to an altimeter, up to 36,090 feet (11,000 m). Outside that go, a goof will be introduced which can be resolved unmistakably for each remarkable weight sensor.

**Vibrating point:**
**Circuit Diagram:**

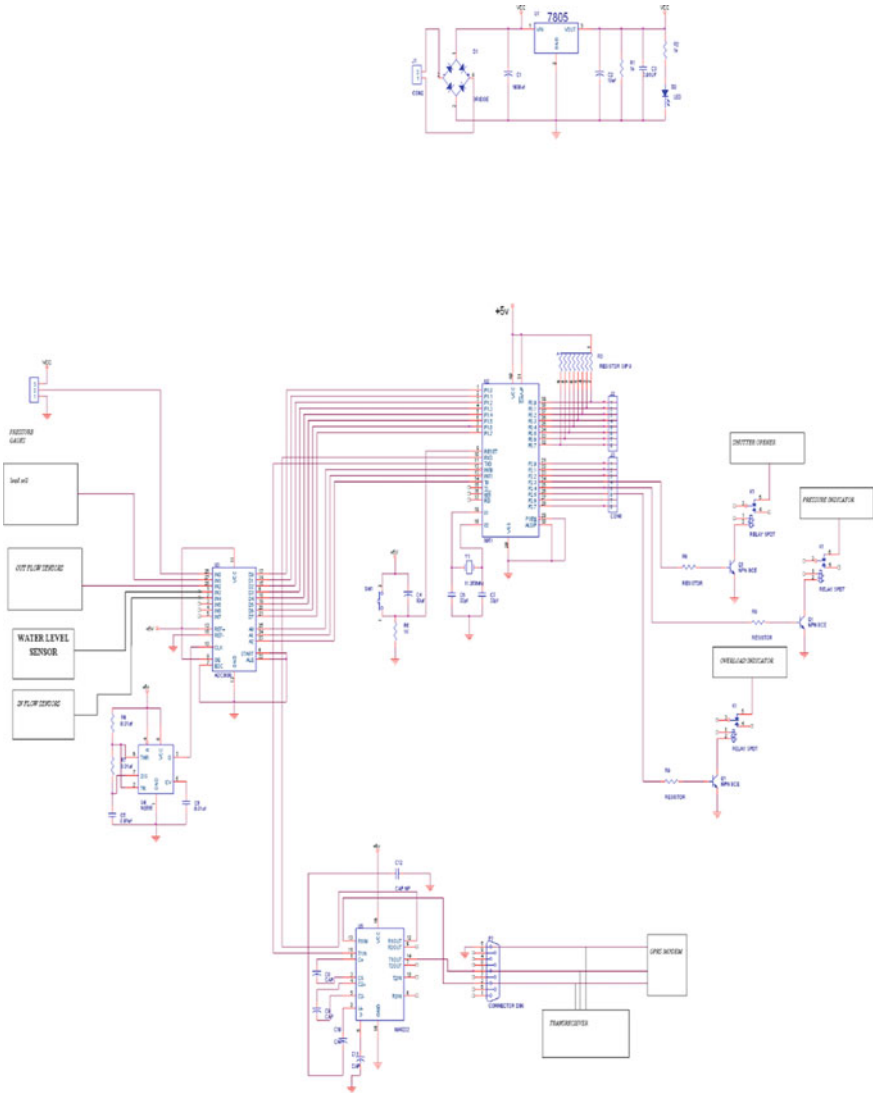



**Fig. 4** Schematic circuit diagram

## 4 Results and Discussion

The basic inspiration driving using the microcontroller in our undertaking is because of the prevalent CMOS 8-piece microcontroller with 8 K bytes of in-structure programmable Flash memory.

The ventures of the microcontroller have been written in Embedded C language and were totaled using KEIL, a compiler used for microcontroller programming. The correspondence between PC and the microcontroller was developed MAX 232 standard and those ventures were furthermore done in C language. The data of the proposed system is given in Table 1.

In this program, the distinctive special limit registers of the microcontroller are set so much that they can send and get data from the PC. This program uses the consecutive library to talk with the ports.

The analysis of Received Data in real time is shown in Fig. 5. The equipment and programming structure of an installed observing framework for ongoing applications is introduced in this paper. Vibration signals have been investigated to distinguish the mechanical flaws. The usage of the investigation method in time and recurrence space are given. The proposed framework lopsidedness discovery method is confirmed with various degrees of seriousness. Our exploration presents a successful method to execute a circulated virtual geographic condition framework dependent on the dispersed augmented experience and geographic data framework (GIS) innovation. It can incorporate current dispersed information and model assets, bolster the community-oriented hazard evaluation work, and utilize a perception domain to improve figuring proficiency. Therefore, our investigation can offer a worldview that will help advance the utilization of data advances to geographic data sciences. The advancement of DVGE frameworks is as yet continuous.

**Table 1** Data of the proposed system

| Sensor | Data logger | Total distance | Prototype fixed days | Frequency of data collection | No. of readings | No. of readings received | % of reading received |
|---|---|---|---|---|---|---|---|
| RS-1 | 1627 | 1.62 | Day 1 | 24 h | 92 | 92 | 100 |
| RS-2 | 1627 | 1.62 | Day 2 | 24 h | 92 | 92 | 100 |
| RS-3 | 1616 | 1.83 | Day 3 | 1 h | 3408 | 3090 | 90.67 |
| RS-4 | 1565 | 1.92 | Day 4 | 1 h | 3329 | 3256 | 97.81 |
| TS-1 | 1627 | 1.62 | Day 5 | 24 h | 92 | 92 | 100 |
| TS-2 | 1627 | 1.62 | Day 6 | 24 h | 92 | 92 | 100 |
| TS-3 | 1587 | 1.78 | Day 7 | 24 h | 92 | 86 | 93.48 |
| TS-4 | 1627 | 1.62 | Day 8 | 24 h | 92 | 92 | 100 |

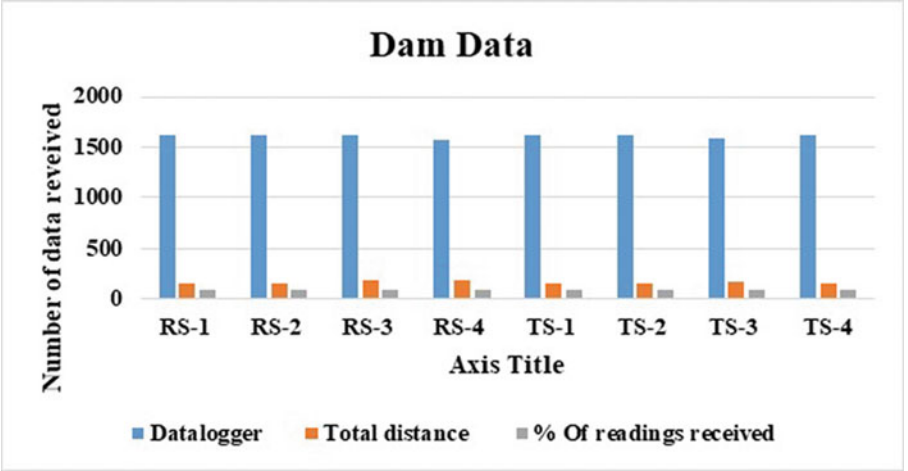



**Fig. 5** Analysis of received data in real time

**Table 2** Comparison of sensor type, energy consumption, percentage of reading, and delay

| Designs | Sensor type | Energy consumption (mW) | Percentage of reading | Delay (ps) |
|---|---|---|---|---|
| Kumar et al. [11] | TS-1 | 542.24 | 89.36 | 125 |
| Kumar et al. [12] | TS-2 | 457.31 | 75.12 | 145 |
| Manikandan et al. [13] | TS-3 | 124.31 | 91.36 | 198 |
| Proposed design | TS-4 | 98.41 | 100 | 114 |

### *4.1 Comparative Analysis*

The Comparison of sensor type, energy consumption, percentage of reading, and delay for different designs which are given in Table 2. In proposed design, sensor type, energy consumption, percentage of reading, and delay are improved than the conventional related works [11–13].

## 5 Conclusion

This paper centers around how to develop a disseminated virtual geographic condition framework, which can furnish geologically circulated clients with a mutual virtual space and a shared stage to actualize chance appraisal work. A five-layer framework engineering was right off the bat planned, Then, some key innovations

including dispersed virtual scene displaying, portable specialist registering administration and dam-break flood steering, were likewise talked about to offer a common and intuitive virtual synergistic workplace. At last, a model framework was executed to help chance appraisal and effect investigation of dam-break in Barrier Lake. In our model framework, topographically appropriated multi-clients can join the virtual geographic condition and execute chance appraisal work. The entire appraisal work can be accomplished in two hours. Contrasted and the conventional strategies, the training application has represented that it can lessen the outstanding task at hand, and the proficiency of the hazard appraisal has been discernibly improved. In the model framework, we don't consider the issue of framework security since it is utilized in the confided in end-clients. It will be settled in the next period of framework improvement.

## References

1. ASCE Task Committee on Instrumentation and Monitoring Dam Performance (2000): Guidelines for Instrumentation and Measurements for Monitoring Dam Performance, Chapter 7, pp 7.1–7.39: Automated data acquisition systems, Appendix B.3, pp. B.52-B.60: Lightning and transient protection, ASCE, Reston/VA
2. Clairmont J, Martin Dupuis M, Choquet P, Budin R (1999) Lightning protection and shielding for geotechnical instruments and data acquisition systems. Geotechnical News, September 1999
3. ICOLD Committee on Automated Monitoring of Dams and their Foundations (2004): Automated dam monitoring systems, guidelines and case histories, International Commission of Large Dams, Paris, France, Bulletin 118, p 253
4. IEEE Standard for Local and metropolitan area networks–Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs), IEEE, Piscataway, NJ (2011)
5. Klebba JM, Choquet P (2005) Eliminating lightning damage for dam instrumentation systems. In: Proceedings of the 25th Annual United States Society on Dams Conference, pp 545–556 (2005)
6. USSD Committee on Monitoring of Dams and their Foundations (2002): General guidelines for automated performance monitoring of dams. United States Society on Dams, Denver/CO, Technical report, p 70
7. Liu L, Wu Y, Zuo Z et al (2009) Monitoring and assessment of barrier lakes formed after the Wenchuan earthquake based on multitemporal remote sensing data. J Appl Remote Sens 3:1–12. https://doi.org/10.1117/1.3153915
8. Qiao QH, Zhang T (2009) 3D-GIS for barrier lake disaster reduction and risk management. In: International Conference on Geo-spatial Solutions for Emergency Management and the 50th Anniversary of the Chinese Academy of Surveying and Mapping (GSEM 2009). Beijing, China, pp 224–226, September 2009
9. Wang GQ et al (2008) Draining Tangjiashan Barrier Lake after Wenchuan Earthquake on numerical simulation. Sci China Press 53:3127–3133. China Beijing
10. Podder AK et al (2021) IoT based smart agrotech system for verification of Urban farming parameters. Microprocess Microsyst 82:104025
11. Kumar NS, Chandrasekaran G, Rajamanickam KP (2021) An integrated system for smart industrial monitoring system in the context of hazards based on the internet of things. Int J Saf Secur Eng 11(1):123–127. https://doi.org/10.18280/ijsse.110114

12. Kumar NS, Nirmalkumar P (2019) A novel architecture of smart healthcare system on integration of cloud computing and IoT. In: 2019 International Conference on Communication and Signal Processing (ICCSP), pp 0940–0944. https://doi.org/10.1109/ICCSP.2019.8698048
13. Manikandan R, Patan R, Gandomi AH, Sivanesan P, Kalyanaraman H (2020) Hash polynomial two factor decision tree using IoT for smart health care scheduling. Expert Syst Appl 141:112924
14. Ellaji C, Sreehitha G, Devi BL (2020) Efficient health care systems using intelligent things using NB-IoT. Materials Today: Proceedings
15. Ammi M, Alarabi S, Benkhelifa E (2021) Customized blockchain-based architecture for secure smart home for lightweight IoT. Inf Process Manag 58(3):1024

# Chaotic System Based Modified Hill Cipher Algorithm for Image Encryption Using HLS


**Anvit Negi, Devansh Saxena, Kunal, and Kriti Suneja**

**Abstract** Lately, information has risen as an imperative asset, and henceforth information safety is of essential concern. A novel and secure image cryptographic system based on the hill cipher algorithm and incorporating a Nose Hoover Chaotic generator has been presented. In the proposed system, the transmission unit mixes an input image with chaotic noise produced by a Chaotic generator and then a bitwise XOR operation is performed to complete the encryption with higher degree of robustness which has been proved by using various measures like entropy and correlation. Moreover, a comparison has been made in the paper. The proposed algorithm includes a key generation system on board which provides an enhanced robustness as well as better computation speed. This makes the algorithm unique and novel since the traditional Hill Cipher has been modified to accommodate pixel values. In addition to the proposed algorithm, we have offered a way to implement the design on a SoC using Xilinx Vivado Design suite.

**Keywords** Chaotic system · FPGA · SOC

## 1 Introduction

With development in networking and multimedia coding technologies, multimedia such as images, text, audio etc. are widely stored and exchanged over the internet. This makes them defenseless against utilization for malignant purposes. Hence, image protection and encryption have now become a widely sought-after field to protect anonymity and prevent unapproved access to web information. Symmetric

A. Negi (✉) · D. Saxena · Kunal · K. Suneja
Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India
e-mail: anvitnegi24@gmail.com

D. Saxena
e-mail: sdevansh2@gmail.com

Kunal
e-mail: kunal99sehrawat@gmail.com

K. Suneja
e-mail: kritisuneja@dtu.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_20

ciphers, for example, Advanced Encryption Standard (AES) are planned with acceptable disarray and dispersion properties [1]. But these advanced encryption standard algorithms seem unacceptable and inappropriate for data items such as images [2]. Since 1990, chaotic generators have gained lot of importance in the field of cryptography due to their attractive properties, for example, high affectability to beginning conditions, ergodicity and pseudo arbitrariness.

This paper presents an advanced algorithm for image encryption which constitutes the use of triple Hill Cipher algorithm and chaotic sequence generator. We have leveraged the property of XORing and have successfully devised a scheme for image encryption. The use of pseudo random number generator as stated above has improved the shortcomings of hill cipher algorithm that it encrypts identical input blocks to identical ciphertext blocks and cannot encrypt images that contain large areas of a single color. In addition to this we have we have proposed and implemented an advanced hill cipher algorithm which utilizes an advanced involuntary fundamental key matrix for the encryption. The objective of using this key generator in the paper is to overcome the pitfall of using a random key matrix in the hill cipher algorithm for encryption, where during decryption we might not retrieve the input image if the chosen key matrix is not invertible. In addition to this, the computational complexity is minimized by avoiding the calculation of inverse of the matrix at the time of decryption, as we use involuntary key matrix for encryption [7].

## 2 Related Works

Chaotic systems have likewise exhibited incredible potential for data security particularly image encryption [3]. Unlike the regular cryptographic calculations which are fundamentally founded on discrete arithmetic, chaos-based cryptographic systems depends on complicated dynamics of nonlinear framework or system which are deterministic yet very straightforward [4]. Along these lines, it can give a quick and secure method for information protection, which is urgent for sight and sound information transmission over quick correspondence channels, for example, the broadband web correspondence [5, 6] (Table 1).

**Table 1** FPGA based recent image encryption systems

| | |
|---|---|
| [8] | FPGA implementation of two-phase image encryption based on chaotic maps is proposed in this paper. Chaotic Behavior of Arnold Cat and Logistic Map are used respectively |
| [9] | RGB image encryption based on 3D chaotic system has been proposed. The system has been implemented on Altera field programmable gate array (FPGA) DE2-115 |
| [10] | This paper showcases an improved chaotic image encryption system which also scrambles pixel position besides changing pixel value and its FPGA implementation |
| [11] | Image encryption algorithm based on four-dimensional chaotic system is proposed. Pipeline and parallel computation features of FPGA has been utilized |

# 3 Methods

## *3.1 Nose Hoover Chaotic Generator*

The subsequent section gives information about the Nose Hoover chaotic generator's digital modelling process with no equilibrium points. Differential Eqs. 1, 2 and 3 express the generator, and they are resolved using the Euler method:

$$x_{n+1} = x_n + y_n * h \tag{1}$$

$$y_{n+1} = y_n + (x_n + y_n * z_n) * h \tag{2}$$

$$z_{n+1} = z_n + \left(x^2 - a * y^2 + b\right) * h \tag{3}$$

Here x, y and z are designated system's state variables. The variable h in the given equation is the step size and the parameters are a and b. Figure 1 depicts Nose Hoover's chaotic attractors incorporating h = $2^{-6}$ at a = 9 and two distinct values of b [12].

The Fig. 2 provides the hardware architecture of the generator, that includes a 32-bit 2's complements fixed-point with 8- bits designated for integer part and 24-bits for the fractional part. In compendium, combinational circuits have been used to compute the different values of state variables. Moreover, adders are incorporated to store these variables in registers. Shift registers are included to achieve multiplication, comprising of a step size variable which is numerically equal to $2^{-6}$.

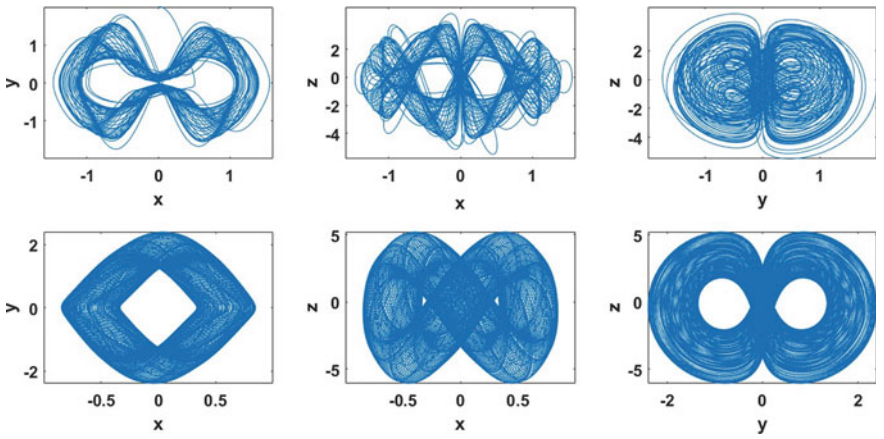



**Fig. 1** Nose Hoover chaotic attractors' projections while applying h = 2 6 at a = 9 and (a) b = 2 and (b) b = 6

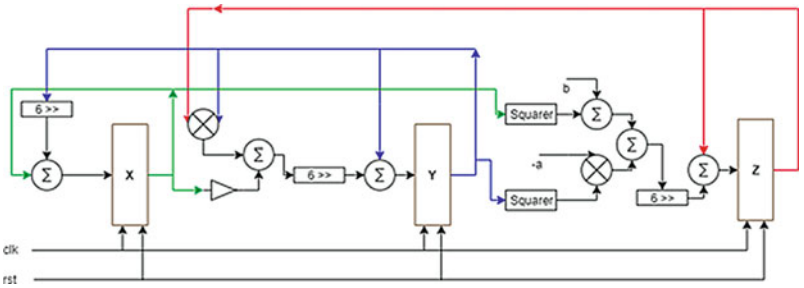



**Fig. 2** Hardware architecture [12]

## *3.2 The Hill Cipher and Its Transformation*

It is well known polygraphic substitution cipher which is formulated on the concepts of linear algebra. The algorithm traditionally includes three matrices namely the plain text matrix, the key matrix (K), and finally the ciphertext matrix (C). Equations 4 and 5 depict the working of hill cipher encryption and decryption. This cryptographic method was created to prevent the message from being decrypted by frequency analysis techniques. It encrypts and decrypts data using a n x n matrix as the key.

$$C = (K * P) mod(256) \tag{4}$$

$$P = \left(K^{-1} * C\right) mod(256) \tag{5}$$

*where $K^{-1}$ is the inverse of the matrix.*

Since the original cipher algorithm was made for text subsequent alteration has made possible the use this algorithm for image encryption system. Not only the cipher algorithm is modified but the corresponding key generator has also been modified to process the image [13]. The classic Hill Cipher has been modified to accommodate the 256 range of pixel values of our input image data. For colour image, the image matrix is broken into its RGB (Red Blue Green) components and is then passed onto the hill cipher block.

## *3.3 Hill Cipher Key Generator*

The most significant flaw in Hill Cipher is the difficulty in determining the proper encryption key matrix. The generation of key matrix that is used in decryption process is unattainable if the encryption key is not chosen properly. The reason being that inverse of encryption key is possible only if the matrix is invertible. This paper uses playfair cipher to generate the key matrix. Plaintext and key are the required inputs in this algorithm, with the plaintext used not being the original plaintext to be encrypted

as the ciphertext, but rather the text used solely to generate the key. The output key matrix is developed using this ciphertext as our input by randomly iterating through it. In Hill Cipher, a matrix containing number is used for encryption and this is labelled as key matrix. The decryption matrix here is the modular arithmetic inverse of the encryption key matrix. Two critical criterias for the key matrix is: it needs to be invertible, and once invertible, the gcd(det(K), 26) must be 1, otherwise only decryption matrix is possible. To inculcate the 256 range of pixel values for our image encryption, the second criteria is modified to gcd(det(K),256) to be equal to 1 [7].

The process to create the key matrix is as follows: Firstly, the given plaintext is encrypted using Playfair Cipher to generate the ciphertext. Then, each character of the ciphertext in converted to the numeric domain for the generation of our key matrix. Thus, we have a bunch of random numbers equivalent to the number of characters in the ciphertext. We then randomly iterate through these numbers and make a square matrix (n by n, where n is the size of key matrix) out of them. This randomly generated square matrix is then checked for the two conditions stated above which are required for hill cipher key matrix. If any one of the conditions is not satisfied, we then reiterate through the numbers and regenerate the square matrix [7]. This process is repeated until both the conditions are met and hence, the correct key matrix is obtained.

## 4 The Proposed Algorithm Incorporating a Chaotic Generator

The proposed algorithm has been represented in Fig. 3. Initially the algorithm generates the encrypted key using the novel key generator block from the perspective of hardware designer and then the generated key is passed onto the triple hill cipher block along with our input image which encrypts as well as decrypts the image thrice based. Everything is designed in order to work on a Soc board environment. Moreover, the algorithm includes a stir up operation which fuses chaos with the output of encrypter by using the results generated from the chaotic generator. This novel

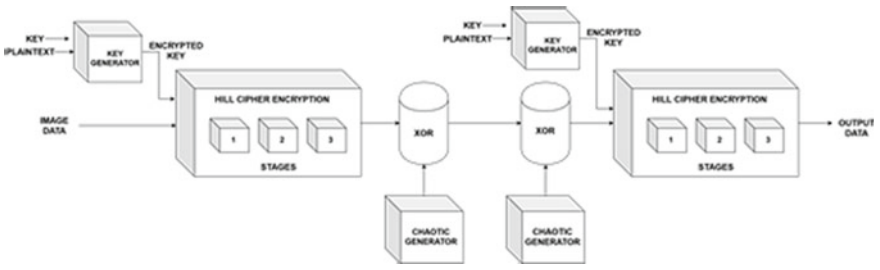



**Fig. 3** The proposed cipher algorithm

feature which the algorithm encompasses makes the input data secure even at times of cyber-attacks. Since the output of the image is not necessarily between the permissible pixel values of 0 to 255 it makes it almost impossible to decrypt the image. The statistical analysis has been done in the following section.

Ordinary Hill cipher algorithm is flawed and if by chance a cryptographer obtains a part of original image input and an encrypted copy simultaneously it was very well possible to solve the cipher. The modification by incorporating the chaos-based system creates an enigma which is very well secured from attacks of these kinds. Even if a part of original image and encrypted is obtained the stir operation disables the solving of cipher in this manner. The whole concept of chaos-based devices on a platform like FPGA is unorthodox and the motivation to include a chaotic generator entirely lies on the fact that a versatile yet simple XOR operation can make the system secure many fold times.

Moreover, the triple hill cipher not only increases the security but with the advance computation capabilities of FPGA it has been made possible that the triple hill cipher is achieved in limited trade off. Hence there is a remarkable increase in robustness and safety of the system which even makes the proposed system immune against frequency analysis techniques which are applicable on traditional cipher algorithms. Pointing out the advantage of chaos-based system, the algorithm's nose hover chaotic generator provides a continuous stream of random numbers and the selection of number from the stream is randomly determined by the time taken by the triple hill cipher which varies for each image. Hence theoretically it is impractical to find the possible numbers used from the pool of random numbers. The encryption scheme mentioned in the paper was modified for a text-based system, here in this image encryption algorithm the classic hill cipher is modified to accommodate all pixel values to encrypt the colour image on a FPGA board. With involvement of a key generator in the same encryption environment we have defeated the problem of non-invertible key on the input user end. The major flaw of the encryption scheme mentioned in was the use of invertible key matrix which has been subdued in this proposed algorithm [14].

This algorithm specified in this paper is unique. The current algorithm encrypts an image while the earlier algorithm in [14] was used for text encryption. Moreover, as already specified earlier the current algorithm overcomes the problem of key generation and improves the usability from a user perspective. The algorithm mentioned in the [14] was not capable of encrypting images since only a 29-character hill cipher was used but here we have made the hill cipher accommodate 256-pixel value. The entropy value shows the effectiveness of the new and modified hill cipher-based image encryption system. Moreover, the work is unique since it is not a software-oriented algorithm. This scheme uses hardware capabilities of FPGA board which makes the whole algorithm novel and unique.

The traditional hill cipher has been modified by incorporating a key generator and XOR operation. The encryption scheme has been different from the scheme mentioned in [14] since the hill cipher is modified for image by making the cipher including 255 pixels. This makes the encryption scheme very unique and robust the entropy analysis has been shown which proves this fact. It also includes a key

generator which has been synthesized on hardware which makes the scheme novel. No hardware synthesis of this kind of encryption scheme has been done till date which makes this very novel and robust. In addition to this the novel hardware based key generator safeguards the issue of key and provides an extra layer of security. It has also been observed that another advantage of incorporating a random stir value which is based on the delay of the first stage of triple encryption is that even if a leak happens this quality of the algorithm makes is immune by including the five 32 keys x0, y0, z0, a, b used in stir up stage which accounts to 2160 combinations in addition to the cipher algorithm and this has been labelled computationally non practical to solve as advocated by the Advanced Encryption Standard (AES). Hence this gives an edge to the proposed scheme over the existing one [15].

## 5 Results and Observations

### *5.1 Correlation and Histogram Analysis*

In the literature, statistical work showing the advantage has been carried out in order to prove the robustness of the pro- posed encryption system. The characteristics of uncertainty and diffusion lead to high resistance to statistical and analyst attacks. This is achieved by examining the distribution of cypher pixels, the analysis of the correlation in the cypher image between the neighboring pixels is presented in the Table 2 (Fig. 4).

The proposed scheme has been implemented in Zynq board xc7z035fbg676-3 and the complete hardware utilization has been provided in Table 3. The implementation results were obtained using the Vivado design suite. The Computational capabilities of the Soc used boasts the robustness of the proposed scheme. The Key sensitivity

**Table 2** Correlation between original & encrypted image

| Type of image component | Encrypted red | Encrypted green | Encrypted blue |
|---|---|---|---|
| Lena Red | 0.00125 | 0.00004 | 0.00076 |
| Lena Green | 0.00174 | –0.00081 | 0.00160 |
| Lena Blue | 0.00177 | –0.00101 | 0.00205 |

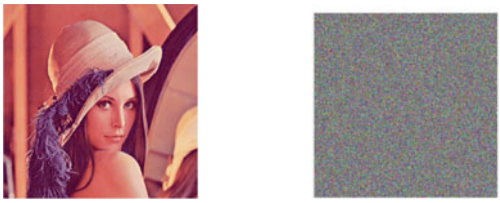


**Fig. 4** Plain image ‘Lena’ and its resultant cipher image

**Table 3** Hardware utilization

| | BRAM | DSP | FF | LUT |
|---|---|---|---|---|
| Hardware used | 9 | 91 | 8944 | 53,440 |
| Available | 1000 | 900 | 343,800 | 171,900 |
| Utilization % | ˜0 | 10 | 2 | 31 |

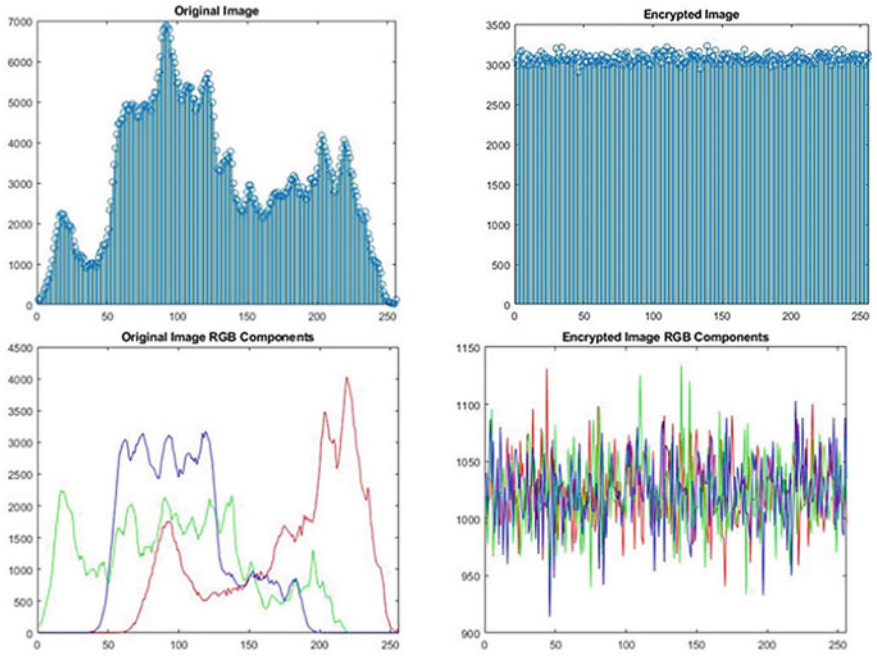



**Fig. 5** Histograms corresponding to RGB components of plain image 'Lena' and its resultant cipher image

is a significant encryption index. The main feature of the sensitivity analysis is the vulnerability analysis in the encryption and decryption process using the incorrect key. The system in this article has a strong sensitivity from a theoretical level due to the sensitivity of the initial value of the chaotic system. Before and after the estimate, the similarity or the correlation of the two images is presented in the Table 2, which clearly indicates that there is absolutely no correspondence between the two.

Numerous histograms with varying of several cypher pictures and their respective plain images were examined. An example of histogram analysis for the well-known picture "Lena" is presented. The histogram of red, blue and green components of cypher images are shown in the Fig. 5. The method of encryption results in bruising photographs by comparing histograms. Cipher picture histograms, approximated by uniform distribution, are very distinct from the image itself and contain no specific

**Table 4** Information entropy values comparison

| Image | Information entropy |
|---|---|
| Original image | 7.7569 |
| Encrypted image (Our method) | 7.9998 |
| [16] | 7.9994 |
| [17] | 7.9978 |
| [18] | 7.9979 |

resemblance. The proposed system including the nose hover chaotic generator resists any attack from analysts.

### *5.2 Information Entropy Analysis*

In thermodynamics, entropy is a physical number that quantifies the degree of unnecessariness in a physical system. Information entropy can be used to display the distribution of each grey value in a digital image. The formula shown in Eq. 6 is used to calculate information entropy:

$$H(m) = -\sum_{i=1}^{M} p(m_i) \log_2 p(m_i) \tag{6}$$

If the encrypted image information has a high entropy, it shows that the image's grey value distribution is uniform. As a result, attackers will have a difficult time getting image information by analyzing this data. Table 4 shows the information entropy of the original image and the encrypted image. We have also compared the entropy values with different research papers as shown in the Table 4.

## 6 Conclusion

As Communications and computer network technologies evolve quickly, more and more attention has been paid to the problems of secure transmission of information, and encryption is an important way of guaranteeing safe information transmission. The conventional encryption approach is not appropriate for image encryption because of the high volume of images, heavy correlation and high redundability of the image itself, so it's important to find a new solution. The invention and birth of the theory of chaos has given hope to picture encryption science. The strong vulnerability to initial circumstances, history of each state and pseudorandom are typical characteristics of chaos that correlate with essential cryptography requirements, namely complexity and diffusion. As a result, chaotic image encryption technologies flourished ever since chaos theory was introduced to image encryption in the 1990's.

The proposed scheme incorporates the nose hover chaotic generator which has not been used widely for an image encryption. This unconventional system combined with the computation capabilities of an FPGA board makes the encryption scheme reliable and secure. Moreover, due to the size of the key space, the sensitivity of the key-mountable methods, the pixel-relation between the encrypted image and original image as well as entropy proves that the method has definite advantage on the safety and reliability of the Lena image and subsequently the real-life images.

## References


1. Chen JX, Zhu ZL, Fu C, Yu H, Zhang LB (2015) A fast chaos-based image encryption scheme with a dynamic state variables selection mechanism. Commun Nonlinear Sci Numer Simul. 20(3):846–860. https://doi.org/10.1016/j.cnsns.2014.06.032
2. Lian S (2009) A block cipher based on chaotic neural networks. Neurocomputing. 32:1296–1301
3. Avasare MG, Kelkar VV (2015) Image encryption using chaos theory. In: Proceedings of the–2015 International Conference on Communication, Information & Computing Technology (ICCICT) 2015
4. Fridrich J (1998) Symmetric ciphers based on two-dimensional chaotic maps. Int J Bifurcat Chaos. 08(06):1259–1284. https://doi.org/10.1142/S021812749800098X
5. Chen G, Mao Y, Chui CK (2004) A symmetric image encryption scheme based on 3D chaotic cat maps. Chaos Solitons Fractals 21(3):749–761. https://doi.org/10.1016/j.chaos.2003.12.022
6. Kwok HS, Tang WKS (2007) A fast image encryption system based on chaotic maps with finite precision representation. Chaos, Solitons Fractals 32:1518–1529. https://doi.org/10.1016/j.chaos.2005.11.090
7. Mahendran R, Mani K (2017) Generation of key matrix for hill cipher encryption using classical cipher. In: Proceedings of the 2017 World Congress on Computing and Communication Technologies WCCCT 2017, pp 51–4
8. Baruah B, Saikia M (2016) An FPGA implementation of chaos based image encryption and its performance analysis. IJCSN Int J Comput Sci Netw [Internet] 5:712–20. www.IJCSN.org
9. Yang CH (2021) Secure color image encryption algorithm based on chaotic signals and its FPGA realization, pp 1–18
10. Zhu J, Du B (2019) Image encryption algorithm based on chaos and its implementation on FPGA. J Inf Hiding Multimed Signal Process 10:278–288
11. Yang CH, Chien YS (2020) FPGA implementation and design of a hybrid chaos-AES color image encryption algorithm. Symmetry (Basel) 12:1–17
12. Tolba MF, Sayed WS, Radwan AG, Abd-El-Hafiz SK, Soliman AM (2018) Hardware speech encryption using a chaotic generator, dynamic shift and bit permutation. In: Proceedings of the 2018 30th International Conference on Microelectronics (ICM), pp 100–3, December 2018
13. Swain G, Lenka SK (2012) A technique for secret communication using a new block cipher with dynamic steganography. Int J Secur Appl 6:13–24
14. Negi A, Saxena D, Suneja K (2020) High level synthesis of chaos based text encryption using modified hill cipher algorithm, pp 3–7
15. Lian S (2009) Multimedia Content Encryption: Techniques and Applications, pp 1–430
16. Liu L, Miao S (2016) A new image encryption algorithm based on logistic chaotic map with varying parameter. Springerplus 5:1–12. Springer International Publishing. https://doi.org/10.1186/s40064-016-1959-1
17. Liu L, Hao S, Lin J, Wang Z, Hu X, Miao S (2018) Image block encryption algorithm based on chaotic maps. IET Signal Process 12:22–30
18. Li R, Liu Q, Liu L (2019) Novel image encryption algorithm based on improved logistic map. IET Image Process 13:125–134

# Chronological-Squirrel Earth Worm Optimization for Power Minimization Using Topology Management in MANET


**B. Devika and P. N. Sudha**

**Abstract** This paper developed a Chronological-Squirrel Earth Worm optimization (C-SEWO) algorithm in MANET for clustering using the topology management. The clustering is performed based on the developed C-SEWO approach and objective functions. The proposed C-SEWO algorithm is designed by combining the Chronological-Earth Worm Optimization algorithm (C-EWO) and Squirrel Search Optimization Algorithm (SSA). Further, the objective functions are computed in terms of the factors, like power, mobility, connectivity, distance, and link lifetime. After the cluster selection, every node in the cluster generates the Gabriel graph for equivalent cluster, then every node updates the file of the neighbor and also preserves graph connectivity and regulates transmission power using the connectivity. The performance of the developed C-SEWO method obtains better performance regarding connectivity, delay, throughput, energy consumption, average residual energy and average routing distance of 95.36%, 0.297 s, 490.66 kbps, 0.024 J, 2.48 J and 147.77 m respectively.

**Keywords** Distance · Gabriel graph · MANET · Power · Topology

## 1 Introduction

MANET has a group of mobile nodes, which creates a dynamic independent network through completely movable infrastructure [1–3]. The nodes communicate with each other without the involvement of centralized base stations or the access points [4]. However, the MANET normally contains wireless nodes, which works without any fixed infrastructure [5, 6]. The rapid changes in the collision, higher error rates, power control, network partitions, connectivity, interference creates the issues in the design of higher level protocols, like routing and implementing the quality of service [7, 8].

B. Devika (✉)
Department of TCE, KSIT, Bengaluru, India
e-mail: devikab@ksit.edu.in

P. N. Sudha
Department of ECE, KSIT, Bengaluru, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_21

Generally, the MANET has various applications in the field of military operations, mobile networks and disaster relief process [9, 10]. The alternation in the radio frequency topology permits the network services to support the specific quality of services and abilities [11]. In MANET, there are mainly two techniques for topology management, namely clustering and power control. Power control method regulates the power on the basis of per node, thus one hop neighbor connectivity is balanced, and to guarantee network connectivity is [4, 12]. The wireless ad hoc network and the highly dynamic topology develop the difficulty in the management of group membership because the connections are temporary [13].

The topology control mechanism is used to manage some metrics, like node mobility, power management and the frequency management [11]. In MANET, the main goal of power control approach is to attain the performance necessity, like network connectivity. The power control method not only enhances the network capability, but also improves the battery capacity of nodes [14]. In the node, the power is consumed at the time of transmission, routing, receiving and processing packets of data [15]. Therefore, the power control method is the significant concern for the MANET [14]. In addition, various routing protocols are utilized for the power management, such as Power Aware Routing Optimization (PARO) [15]. Additionally, other routing methods for MANET can be classified into four types, such as reactive routing, flooding, dynamic cluster-based routing, and the proactive routing [5, 16, 17]. The early congestion detection and adaptive routing method is developed in MANET for enhancing the performances by reducing the routing overhead and delay, and also increases the packet delivery ratio [9, 18]. The multi-hop method executes the transports control protocol congestion control method for the applications, such as data transmission in MANET. [9, 19]. The agent-based congestion control technique is introduced for MANET in which the information about the network congestion is gathered and distributed by the mobile agents [9].

## 2 Literature Survey

This segment illustrates literature survey of various methods used for a topology management in MANET approaches, and the challenges of existing works are described. Rahmani et al. [20] presented a learning automata-based topology control mechanism for controlling network conditions. This method introduced the cognition with the whole network protocol for attaining the network and stack wide performance goals. Chaudhry and Tapaswi [21] introduced a clustering based scalable topology control technique for topology management in MANET. Sharifi and Babamir [22] modelled a clustering technique for effective energy management in MANET. Here, the evolutionary algorithm, named Imperialist Competitive Algorithm (ICA) was introduced by numerical coding for finding the cluster heads. Safa et al. [23] developed a power aware routing procedure for the MANET which developed for maintaining the heterogeneous routing in the network.

## 3 System Model of MANET for C-SEWO

Let us assume a graph $Z = \{X, Y\}$ in the MANET, where $X$ indicates the normal node set, and is represented as, $X = \{x_1, x_2, \ldots, x_a, \ldots, x_b\}$; $1 \leq a \leq b$ in which $b$ represents the total amount of normal nodes and $Y$ indicates the link set, which is denoted as, $Y = \{y_1, y_2, \ldots y_c\}$ in which $c$ represents the total number of links, which are utilized for linking two nodes $x_a$ and $x_b$. Multi objective functions, like connectivity, mobility, power, distance and link lifetime are utilized to design the MANET. Next, let us assume source node, $S$ that transfers the data packets to target node $T$. The set of cluster head is denoted as $C = \{w_1, w_2, \ldots, w_u, \ldots, w_v\}$; $1 \leq u \leq v$ where, $v$ represents the total cluster head.

### *3.1 Proposed Chronological-Squirrel Earthworm Optimization Algorithm for Power Optimization*

This section illustrates the proposed C-SEWO method for topology management in MANET. Figure 1 portrays the block diagram of developed C-SEWO method for power management using topology management in MANET. The developed optimization method is newly designed by combining the, Squirrel Search Algorithm (SSA) [21], and C-EWA [20]. The factors, like life time, connectivity, power, distance and the mobility are computed for clustering. Once the clusters are selected, then, every nodes belongs to the cluster generates the Gabriel graph for the corresponding cluster.

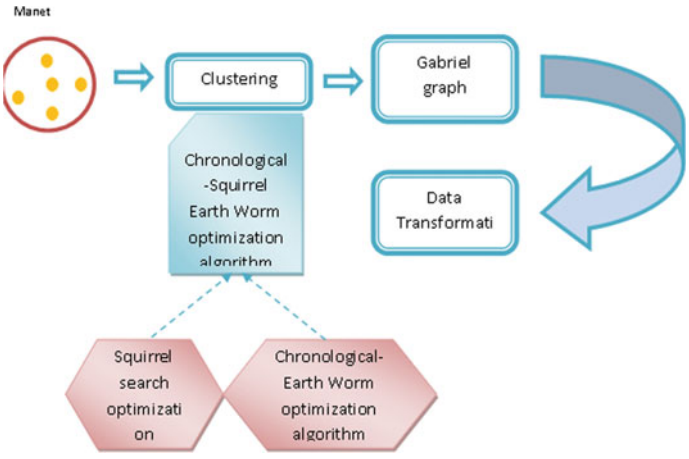



**Fig. 1** Block diagram of developed C-SEWO method for Power optimization in MANET


**Fig. 2** Solution encoding for identifying optimal cluster head

**Initialization.** Initially, the graph is initiated in which every node $w_u$ ; $1 \leq u \leq v$ is initiated along with the value for various parameters, such as mobility, power, connectivity, link lifetime and distance for choosing the optimal cluster head at time $t$.

In this section, the developed C-SEWO method is explained for the clustering process. The developed C-SEWO method is the combination of C-EWA [20] and the SSA [21]. The C-EWA approach is developed for clustering, but the delay time is not much improved. So, the SSA is integrated with the C-EWA approach.

*Solution Encoding.* The optimal cluster head is verified by the solution encoding and is illustrated in Fig. 2. Let us assume the MANET with $v$ number of cluster heads and the proposed C-SEWO along with the fitness function identifies the $o$ optimal clusters. Consequently, the solution is represented as a vector of dimension $o$.

As $W = \{w_1, w_2, \ldots, w_e, \ldots, w_o\}$ ; $1 < e \leq o$. Therefore, the optimal cluster head is selected by considering the maximum link lifetime, increased connectivity, less distance, large power and minimum mobility.

*Formation of Fitness Function.* For identifying the optimal solution, the fitness function is estimated from a solution set. Here, the fitness function is estimated based on the parameters, such as mobility, connectivity, power, distance and link lifetime. The fitness function of developed C-SEWO method is given by Eq. 1,

$$F = \sum_{u}^{v} \frac{1}{5}\left[\frac{P_u}{N_O} + Q_u + L_u + \left(1 - \frac{M_u}{N_O}\right) + \left(1 - \frac{D_u}{N_O}\right)\right] \tag{1}$$

where, $P_u$ denotes the power at the $u^{th}$ node, $N_O$ represents the normalization factor, $Q_u$ indicates the connectivity, $L_u$ expresses the link lifetime, $M_u$ represents the mobility and the distance is portrayed as $D_u$.

Connectivity of the nodes: Based on the bi-directional links the connectivity [23] is estimated, and it is expressed as Eq. 2,

$$Q_u = \frac{1}{z}\left[\sum_{f=1}^{z} \frac{Q_f}{q}\right] \tag{2}$$

where, $Q_f$ represents the connectivity of $f^{th}$ node and $q$ is the total connections.

Power of the Node: Power [23] is considered as one of the significant parameters to select the optimal cluster head. The minimum power to transmit the packets from the node $u$ for obtaining the minimum received power is expressed as Eq. 3,

$$P_u = \sum_{f=1}^{z} P_{\max} \cdot \frac{P_{\min}}{P_f} \tag{3}$$

where, $f$ represents the normal node, $z$ denotes the total normal nodes, $P_{\min}$ illustrates the received power, the maximum power is represented as, $P_{\max}$ and the received power at the $f^{th}$ node is depicted as $P_f$.

Mobility of the Node: The mobility [23] of the nodes is identified based on their responsibilities in MANET. Therefore, the mobility factor is represented as Eq. 4,

$$M_u = \frac{1}{|R_u|} \sum_{f=R_u} M_f \tag{4}$$

where, the relative mobility is represented as, $M_f$ and $|R_u|$ denotes the set of neighbours at the node $u$.

Squared Distance: The squared distance is computed by the distance between $u^{th}$ cluster head and $f^{th}$ node is defined as Eq. 5,

$$D_u = \sum_{\substack{f=1 \\ f \in u}} \left(K_f, R_u\right) \tag{5}$$

Link Lifetime of the Node: The link lifetime [12] is utilized for connecting the two nodes to transfer the information, and it is estimated based on the energy model. The link lifetime is illustrated as Eq. 6,

$$L_u = \frac{1}{z} \sum_{f=1}^{z} E_f \tag{6}$$

where, $E_f$ represent the energy dissipation of $f^{th}$ node, and the updated energy for the normal node after the data transmission is defined as Eq. 7,

$$E_{i+1}\left(K_f\right) = E_i\left(K_f\right) - E\left(K_f\right) \tag{7}$$

where, the energy of current time in normal node is denoted by $E_i\left(K_f\right)$ and $E\left(k_g\right)$ denotes the dissipated energy of normal node.

Likewise, the updated energy for the cluster head after the data transmission is defined as Eq. 8,

$$E_{i+1}(R_u) = E_i(R_u) - E(R_u) \tag{8}$$

where, $E_i(R_u)$ represents the energy of current time in cluster head node and $E(R_u)$ illustrates the dissipated energy of cluster head.

### 3.2 Algorithmic Procedure of the Proposed C-SEWO Algorithm

The algorithmic procedure of the developed C-SEWO approach is elaborated in this section.

**Initialization.** At first, the developed C-SEWO method initializes the solution space to determine the optimal solution, and it is represented as Eq. 9,

$$G = \left\{G_1, G_2, ...G_n...G_p\right\} \tag{9}$$

where, $G_p$ represents the $p^{th}$ population of the proposed C-SEWO.

**Estimation of Fitness Function.** The fitness function is estimated for every solution based on the multi objective fitness parameters, like squared distance, power, mobility, connectivity and link lifetime. Furthermore, the fitness function is computed using Eq. 1.

**Update the Solution Using the Proposed C-SEWO.** The C-EWA is used for clustering the graph, but the method was not solved the meta-heuristic problems. Here, the C-EWA and the SSA is integrated to overcome the meta-heuristic problems. The C-EWA is updated with the SSA for effective clustering. The update solution of the C-EWA equation is given by Eq. 10,

$$G_{m,n}(s+1) = \frac{G_{m,n}(s) + G_{m,n}(s-1) + 3(V_r * U)}{2} \tag{10}$$

where, $V_r$ denotes the weight vector and $U$ represents the random number generated from the Cauchy distribution. In the C-EWA equation, the updation is executed by using the SSA, and it is given by Eq. 11,

$$G_{m,n}(s) = \frac{G_{m,n}(s+1) - g_l H_i G_{best}(s)}{1 - g_l H_i} \tag{11}$$

Substitute Eq. (11) in (10) and the Eq. 12 become,

$$G_{m,n}(s+1) = \left[\frac{(1 - g_l H_i)\big(G_{m,n}(s-1) + 3(V_r * U) - g_l H_i G_{best}(s)\big)}{2(1 - g_l H_i) - 1}\right] \tag{12}$$

where, $g_l$ represents the random gliding distance, $U$ is the random number from the Cauchy distribution with $f = 1$, $H_i = 1.9$ and $V_r = \frac{\sum_{v=1}^{k} G_{m,n}}{k}$.

**Check the Feasibility of Solutions.** Once the C-EWA is updated, the solution goes through the estimation of fitness function. In this case, the optimal cluster head is

recognized by finding the position with the least fitness. Then, the optimal cluster head is replaced by the pervious solution.

**Termination.** The above processes are continued until detecting the optimal cluster heads. The algorithm is completed until it reaches the maximum iteration.

### *3.3 Gabriel Graph*

The Gabriel graph [4] is measured as a vertices set in the plane, where the edge among two nodes exist if other nodes do not overlap the edges among two nodes. Every node creates the graph in cluster to mitigate the transmission power. The graph is assumed as $h$ spanner, and $h$ represents the spanning ratio, and is expressed as Eq. 13,

$$h = \max_{K_f, K_u} \frac{A_{k_f, K_u}}{B_{K_f, K_u}} \tag{13}$$

where, $A_{K_f, K_u}$ represents the shortest path length between the two nodes $K_f$ and $K_u$ and $B_{k_f, K_u}$ denotes the Euclidean distance.

## 4 Results and Discussion of C-SEWO Algorithm

The experimental setup, performance metrics, dataset description, competing techniques and the comparative analysis are given below.

### *4.1 Experimental Setup*

The implementation of the proposed model is performed using NS2 software and the simulation is done in Ubuntu OS, 4 GB RAM, with Intel core i-3 processor.

## 5 Experimental Results

### *5.1 Analysis for 100 Nodes*

The analysis in terms of connectivity, power, delay, throughput, average residual energy, energy consumption and the average routing distance for 100 nodes is represented in Figs. 3, 4, 5 and 6 below. The analysis based on connectivity by varying

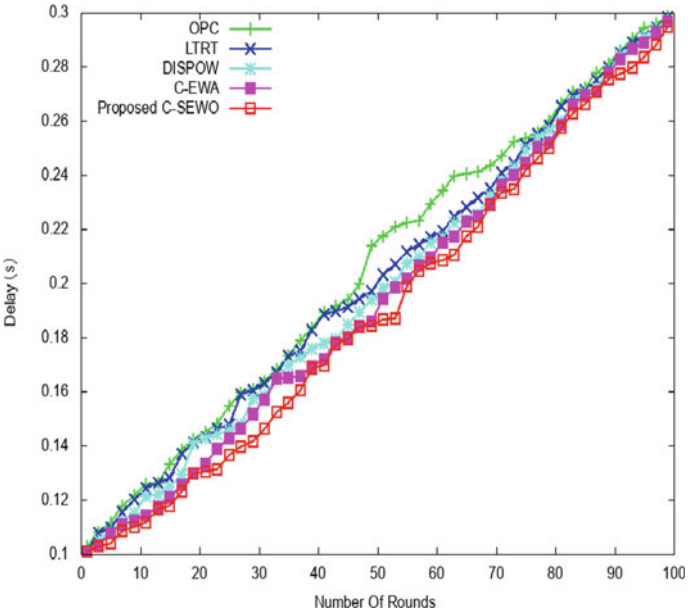



**Fig. 3** Analysis of connectivity

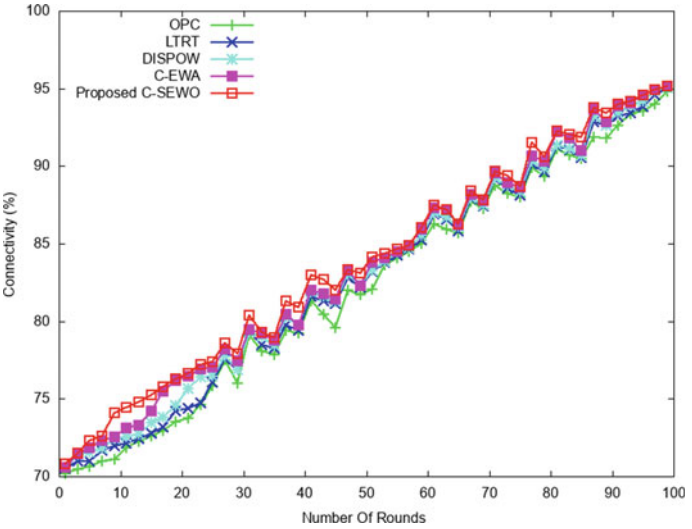



**Fig. 4** Analysis of delay

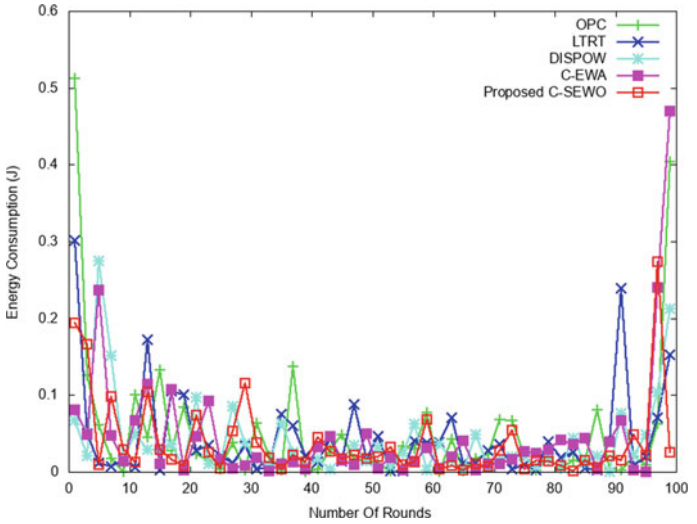



**Fig. 5** Analysis of energy consumption

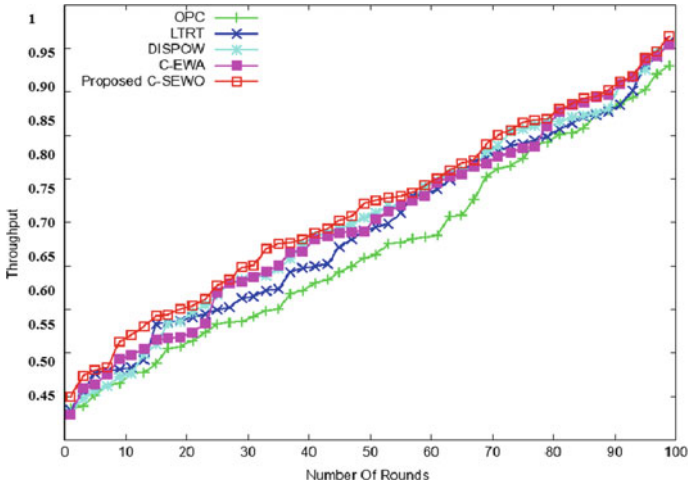



**Fig. 6** Analysis of throughput

rounds from 0 to 100 is depicted in Fig. 3. At 40th round, the values of connectivity measured by the existing OPC, LTRT, DISPOW, C-EWA and the proposed C-SEWO are 79.20, 79.35, 79.35, 79.68 and 80.55% respectively. The analysis of delay parameter is illustrated in Fig. 4. For round 60, the delay obtained by existing OPC, LTRT, DISPOW, C-EWA and the proposed C-SEWO are 0.237, 0.222, 0.220, 0.217 and 0.210 s. Similarly, the analysis of energy consumption is illustrated Fig. 5. For the 10th round, the energy consumption values computed by the existing OPC, LTRT,

**Table 1** Comparative analysis

| Metrics | OPC | LTRT | DISPOW | C-EWA | Proposed C-SEWO | EEDREAM | MAC Layer Energy Consumption and Routing Protocol Algorithm Optimization |
|---|---|---|---|---|---|---|---|
| Connectivity (%) | 95.01 | 95.24 | 95.35 | 95.35 | **95.36** | 95.12 | 95.10 |
| Delay (sec) | 0.299 | 0.299 | 0.298 | 0.298 | **0.297** | 0.298 | 0.299 |
| Energy consumption (J) | 0.40 | 0.15 | 0.21 | 0.46 | **0.024** | 0.43 | 0.04 |
| Throughput | 0.92 | 0.94 | 0.94 | 0.94 | **0.96** | 0.95 | 0.94 |

DISPOW, C-EWA and the proposed C-SEWO are 0.032, 0.057, 0.097, 0.0036 and 0.071 J. The evaluation based on throughput is presented in Fig. 6.

Table 1 illustrated the analysis of various techniques based on connectivity, delay, energy consumption, and throughput. Therefore, the proposed C-SEWO method has maximum connectivity, throughput and obtains minimum energy consumption, and delay. The recent Protocols EEDREAM, MAC Layer Energy Consumption and Routing Protocol Algorithm Optimization have also been considered for the Analysis and shown in Table 1.

## 6 Conclusion

Initially, the MANET nodes are fed to the clustering process and optimal cluster heads are selected. After the cluster head selection, the Gabriel graph is generated to reduce the transmission power of nodes, which maintains the energy and the power in MANET. The nodes, which offers the maximum value for the parameters are considered as the optimal cluster heads.

## References

1. Peng S, Wang Y, Xiao H, Lin B (2020) Implementation of improved AODV routing protocol for maritime ad hoc networks. In: International Congress on Image and Signal Processing, Biomedical Engineering and Informatics (CISP-BMEI), 17–19th October 2020
2. Kanellopoulos D, Cuomo F (2021) Recent developments on mobile ad-hoc networks and vehicular ad-hoc networks. Electronics 10:364. https://doi.org/10.3390/electronics10040364
3. Ding X, Guo J, Wang Y, Li D, Wu W (2021) Task driven charger placement and power allocation for wireless sensor networks. Ad hoc Netw 119:102556

4. Nikaein N, Bonnet C (2004) Topology management for improving routing and network performances in mobile ad hoc networks. Mob Netw Appl 9(6):583–594
5. Basu SS, Chaudhuri A (2003) Self-adaptive topology management for mobile ad-hoc network. J Inst Eng India Part Electron Telecommun Eng Div 84:7–13
6. Katiyar S, Kumar S (2018) An Efficient topology management algorithm in MANETs. In: Proceedings of International Conference on Advances in Computing, Communication Control and Networking IEEE, pp 267–272, October 2018
7. Samanta S, Ray SS, Gupta SS, Naskar MK (2006) A novel algorithm for managing network configuration of a mobile ad-hoc network. In: Proceedings of the 4th Asian International Conference on Mobile Computing (AMOC 2006), pp 51–58, January 2006
8. Pramanik A, Choudhury B, Choudhury TS, Arif W, Mehedi J (2015) Decentralized topology management on mobile ad hoc networks. In: Proceedings of the Global Conference on Communication Technologies (GCCT) IEEE, pp 117–120, April 2015
9. Devarajan K, Padmathilagam V (2015) An effective and efficient congestion control scheme for MANET using COCO method. Int J Emerg Technol Adv Eng 5(8):185–189
10. Sasikala K, Reka R, Prasath MA (2014) Performance evaluation of improving network capacity with optimized cooperative (COCO) topology control scheme in MANETS. Int J Res Emerg Sci Technol 1(6):58–64
11. Carvalho M, Granados A, Subramanian S, Perez C (2010) A cross-layer approach to mixed-control topology management for MANETs. In: Proceedings of the MILCOM 2010 Military Communications Conference, pp 411–416
12. Tseng YC, Ni SY, Chen YS, Sheu JP (2002) The broadcast storm problem in a mobile ad hoc network. Wirel Netw 8(2–3):153–167
13. Liu J, Sacchetti D, Sailhan F, Issarny V (2005) Group management for mobile ad hoc networks: design, implementation and experiment. In: Proceedings of the 6th International Conference on Mobile Data Management, pp 192–199, May 2005
14. Gautam DR, Sharma S, Sahu S (2011) Enhanced transmission power control mechanism based on RSSI for MANET. Int J Comput Appl 28(1):25–30
15. Gomez J, Campbell AT, Naghshineh M, Bisdikian C (2003) PARO: supporting dynamic power controlled routing in wireless ad hoc networks. Wirel Netw 9(5):443–460
16. Mc Donald AB, Znati TF (1999) A mobility-based framework for adaptive clustering in wireless ad hoc networks. IEEE J Sel Areas Commun 17(8):1466–1487
17. Gerla M, Pei G, Lee SJ (1999) Wireless, mobile ad-hoc network routing. In: Proceedings of IEEE/ACM FOCUS 1999, pp 1112–24, May 1999
18. Kumaran TS, Sankaranarayanan V (2011) Early congestion detection and adaptive routing in MANET. Egypt Inform J 12(3):165–175
19. Sharma VK, Bhadauria SS (2012) Mobile agent based congestion control using aodv routing protocol technique for mobile Adhoc-network. Int J Wirel Mob Netw (IJWMN) 4:229–314
20. Devika B, Sudha PN (2019) Power optimization in MANET using topology management. Eng Sci Technol Int J 23(3):565–575. ISSN 22150986, ELSEVIERBV
21. Jain M, Singh V, Rani A (2019) A novel nature-inspired algorithm for optimization: squirrel search algorithm. Swarm Evol Comput 44:148–175
22. Rahmani P, Javadi HH, Bakhshi H, Hosseinzadeh M (2018) TCLAB: a new topology control protocol in cognitive MANETs based on learning automata. J Netw Syst Manag 26(2):426–462
23. Sharifi SA, Babamir SM (2020) The clustering algorithm for efficient energy management in mobile ad-hoc networks. Comput Netw 166(15):106983

# Classification of Neurological Disorders with Facial Emotions and EEG


**T. G. Geethesh, S. Surya Prasad, K. Harshak Krishnaa, S. Karthick Saran, and O. K. Sikha**

**Abstract** Neurological disorders are diseases that affect the brain and the central nervous systems, which, when left untreated, can lead to severe repercussions in later stages. The symptoms for these are very subtle, and the afflicted are often not capable of expressing their discomfort. Furthermore, these disorders often go unnoticed by technical experts at times. The detection of these maladies often involves the Electroencephalogram (EEG), and recent studies proved that the facial emotions of the afflicted are a major factor as well. In order to boost the public awareness and increase the means to identify such illnesses, a technological solution is required. This paper analyses the effect of facial emotions and EEG data in the detection of neurological disorders, and their potential in serving as standalone parameters for detection and classification of the same.

**Keywords** Disease classification · EEG · Facial emotions · Neurology

## 1 Introduction

Neurological disorders are diseases that affect the brain and the central and autonomic nervous systems. The World Health Organization reports that millions of people around the world are affected by various types of neurological disorders every year. Right now, there is a need for a technological solution for detecting neurological disorders. The signs pertaining to these diseases are highly either emotion-based or physical in nature and they lead to severe repercussions if left untreated. Identification of the symptoms often require technical expertise, and the patient is often not aware that the conditions that he experiences are symptoms. The symptoms of neurological disorders are heterogeneous, with variations in the type of disorder. The patient could also experience physical or emotional symptoms. The emotional symptoms are harder to identify, and the most prevalent method today, is by word of mouth, which is prone to error if not properly expressed. As neurological disorders are

T. G. Geethesh · S. Surya Prasad · K. Harshak Krishnaa · S. Karthick Saran · O. K. Sikha (✉)
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: ok_sikha@cb.amrita.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_22

degenerative in nature, they can cause instability in cognitive functions, and can worsen over time. Thus, an effective way of identifying such disorders is desired, and seek assistance as soon as possible. The primary objective of this paper is to analyze the relation between facial emotion and map them to classify neurological disorders. Rest of this paper is organized as follows: the related works are detailed in Sect. 2; Sect. 3 describes the proposed model in detail; the results obtained are detailed in Sect. 4. State-of-the-art comparative analysis has been done as part of Sect. 5. Finally, the paper concludes with closing remarks in Sect. 6.

## 2 Related Works

Facial Emotion Recognition (FER) models try to recognize emotion of a person from images/videos by identifying the facial contours. Mellouk et al. [1] have proposed their review and insights about Facial Emotion Recognition using Deep Learning models. Sesha et al. [2] mentions that certain parts of the face are more useful to determine a certain set of emotions. Keshari [3] suggests that different focused strategies can be used to determine emotions in different people which allows them to gain more information about the facial emotions. Grabowski et al. [4] states that subtle changes in FER can be an indicator of the onset of a Neurodegenerative disorder or can act as a catalyst for other deficits that the disorder causes. The work done by Marcó-García et al. [5] indicates that small changes in the detection of negative emotions like anger, sadness and fear occur in case of mild cognitive deficits, but the positive emotions remain stable. This study provides several cases where negative emotions like sadness deteriorates to Alzheimer's and Dementia. Several FER changes mark this change. Detection of Huntington's disease involves recognizing negative emotions, particularly disgust. Parkinson's disease marks a distinct lack of ability to express negative emotions. In case of Sclerosis, there is a notable impairment in the perception of anger, surprise and disgust. The authors describe the importance of psychological responses to determine what an individual is actively thinking about, and it could be used to determine the cognitive wellbeing of an individual. Ravi et al. [6] have made a comparison between different methods of Facial Emotion Recognition. Yagis et al. [7] have formulated a positive approach to find Multimodal Emotion Recognition using the fusion of Electroencephalograph (EEG) and Facial Expressions. There is a significant link between facial expression and the corresponding EEG produced by an appropriate stimulus. The work done by Li et al. [8] supports this claim, by taking in a dual input of EEG and images of Facial data and proving that Facial expressions can be predicted based on variations in EEG data. The research done by Fahd Alturki et al. [9] have discussed the efficiency of EEG signals to diagnose neurological disorders. Zhu et al. [10] has designed a Cost-Efficient Classification for Neurological Disease Detection from EEG data using a machine learning model. Al-Amyn et al. [11] have done a systematic review on different Deep Learning models used for Neurological disorders. The research done by Srinidhi et al. [12] focuses on predicting neural seizures using EEG signals. Their

research also states that EEG data can be analyzed by passing in the raw format, rather than subjecting it to DWT and reducing dimension. Patel et al. [13] have experimented with CNNs for analyzing the higher-dimensional data. The work done by Yolcu et al. [14] explores the effectiveness of facial expression as a feature for the identification of neuro-disorders. Gautham et al. [15] has done a meta-analysis on the prevalence and diagnosis of neurological disorders. Facial emotion recognition is an integral part of social cognition that enables a healthy social life. The identification and treatment of the FER deficits has, however, not been implemented despite its effectiveness.

## 3 Methods

Facial Emotions are generally regarded as a cornerstone for cognitive and social functioning. A deficit in expressing the same is a clear indication of reduction of mental ability and could be a precedent to neuro-disorders. Hence, the research is undertaken in this area to ascertain neuro-disorders based on facial emotions. Normally, Electroencephalogram (EEG) is the most common test used to diagnose neurological disorders. This primary objective of this work is to correlate the emotions and EEG readings of a person in order to identify various neurological disorders. An ANN powered approach is proposed to combine the EEG readings and the facial emotions to predict the prospect of getting affected by these diseases. The proposed solution uses the facial images and EEG readings of a person as the ground truth with the types of disorders being the target value.

### *3.1 Dataset Used*

**DEAP.** The DEAP Dataset contains a video feed of 22 Participants, with varying neuro-diseases, and the preprocessed psychological readings for each of them. The data was down sampled to 128 Hz. A band pass frequency filter from 4.0–45.0 Hz was applied. The data was averaged to the common reference. The data was then broken up into Valence, Arousal, Dominance, Liking, and subjected to normalization and scaling accordingly, before being split into training and testing data.

**CK+.** The Cohn-Kanade Plus (CK+) contains facial images with annotated expressions. There are 7 different emotions captured, namely happiness, sadness, anger, contempt, disgust, fear and surprise. Preprocessing involves converting the images to Binary/Grayscale, Pixel brightness transformation and Geometric transformation. The preprocessed images are then separated and labelled for testing purposes.

### 3.2 *Proposed Architecture*

The architecture used here relies on parallel processing between the FER and EEG data. Hence, the processing needs to be light weighted and efficient to preserve resources. Furthermore, faster convergence needs to be prioritized, hence a trapezoidal architecture is adopted for the EEG Data. Whereas for FER, convolutional layers are used, and the appropriate filter size is applied, with max-pooling. For further refining regularization and dropout, techniques like batch normalization are used to reduce overfitting. Finally, both the outputs of the models are then combined to produce a probable list of neuro-diseases. This is done through a custom function that maps the intensity of the emotion to the possible neuro-disease, as demonstrated by Fig. 2. The architecture flow is shown in Fig. 1.

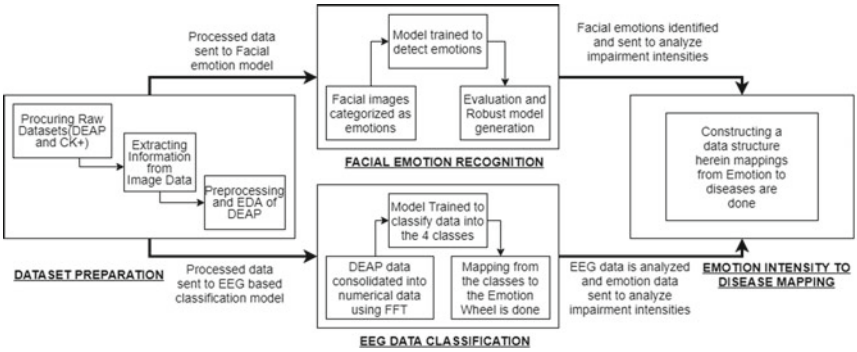



**Fig. 1** Architecture diagram


**Fig. 2** Emotion mapping to neurological diseases

### 3.3 Facial Emotion Recognition in Neurological Disorders

Emotions are the most basic physiological reactions in any living beings and are studied and classified into six basic categories. Neurological disorders can severely affect people's facial emotions. Facial Emotion Recognition (FER) identifies the type of disorders caused by neural degeneration, and a deficit would serve as a clear indicator. In order to differentiate between the types of disorders, the intensities of FER impairments are identified along with the person's age, gender and other reliable factors. Six categories of neurological diseases were taken for the study: Mild Cognitive impairment, Alzheimer's disease, Parkinson's disease, Huntington's chorea, Sclerosis and Epilepsy. FER impairment intensities are then mapped to the neurological diseases under study as shown in Fig. 2. This mapping has been generated from the work of [5] as part of their study on facial emotion recognition in neurological disorders.

### 3.4 Mathematical Model

For classifying the neurological disorders, change in the percentages of FER are taken into consideration for the most significant emotions. The change in FER values is represented as $C_A$ (Anger), $C_S$ (Sadness), $C_F$ (Fear), $C_D$ (Disgust) and $C_P$ (Positive Emotions). The expected value of FER is calculated first which is then used to find the change in FER for any emotion using the percentage of change ($C_E$) as in Eq. (1). This Change in FER of all the emotions are considered to classify the Neurological disorders.

$$\begin{gathered} C_E\% = ExpectedFER\% - MeasuredFER\% \\ Changein\,FER = \begin{cases} Drastic & for\, C_E\% > 10 \\ Mild & for\, 5 < C_E\% < 10 \\ NoChange & for\, C_E\% < 5 \end{cases} \end{gathered} \tag{1}$$

### 3.5 Proposed System

A live video feed is input to the FER model, along with the corresponding stimuli. The EEG data is fed in parallel to the EEG model in the form of frequencies. The FER model detects the emotion intensities of the sample, while the EEG model outputs the classes that are used as the ground truth. The results of both these models are used in conjunction with each other, and the deficit in FER is calculated, which is to determine the neuro-disease. This approach is taken so that eventually, the FER system would be sufficient to determine the neuro-diseases without the EEG readings.

## 4 Results

This section analyses the effectiveness of the proposed classification model. The FER model used in the current study is trained on CK+ and tested on input of the facial data as frames, spliced from a real-time video feed. Each frame is then subject to processing of emotions, and the mean of emotion intensities are taken and stored, within a fixed timeframe in which the stimulus is induced. Figure 3 shows an example of a person's emotions being detected along with the intensities at two different timeframes. These average emotion intensities are used for comparison with those of a normal person. The difference in these values is then recorded and are then mapped against a custom mapping as shown in Fig. 2 to determine the possible neuro-diseases at that point in time. Figure 4 shows the accuracy and loss plot of the FER deficit model. The final performance of the model to predict the FER deficits gives an accuracy of 90.11%.

A test is conducted where the subjects react to a set of stimuli. Their facial expressions to the stimulus are captured and recorded in order to identify the corresponding natural facial emotions, stored along with timestamps. The reason behind this is that all the mapping elements of the disease classification are represented as change and are classified as per their severity. The abstract quantity of change by comparing the results obtained from the patient with that of a normal person is calculated. The amount of deviation from the normal state is then classified as Drastic/Mild/No

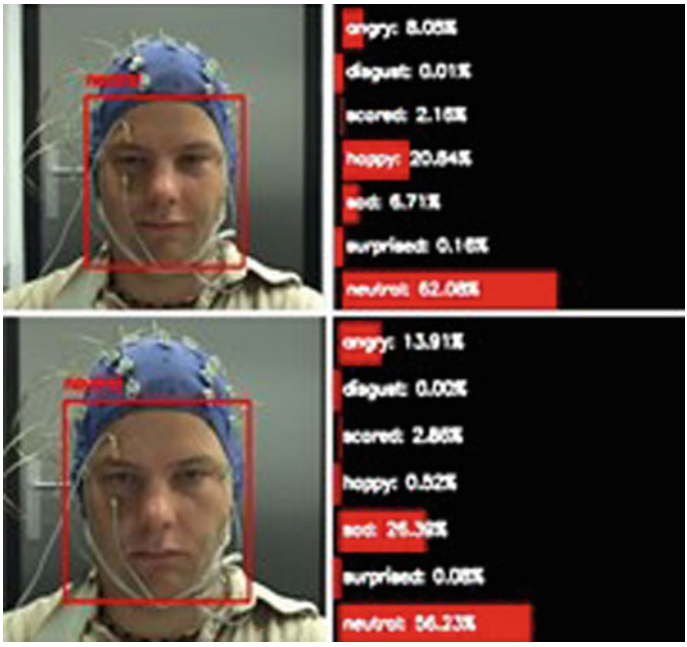



**Fig. 3** FERResultswithintensities

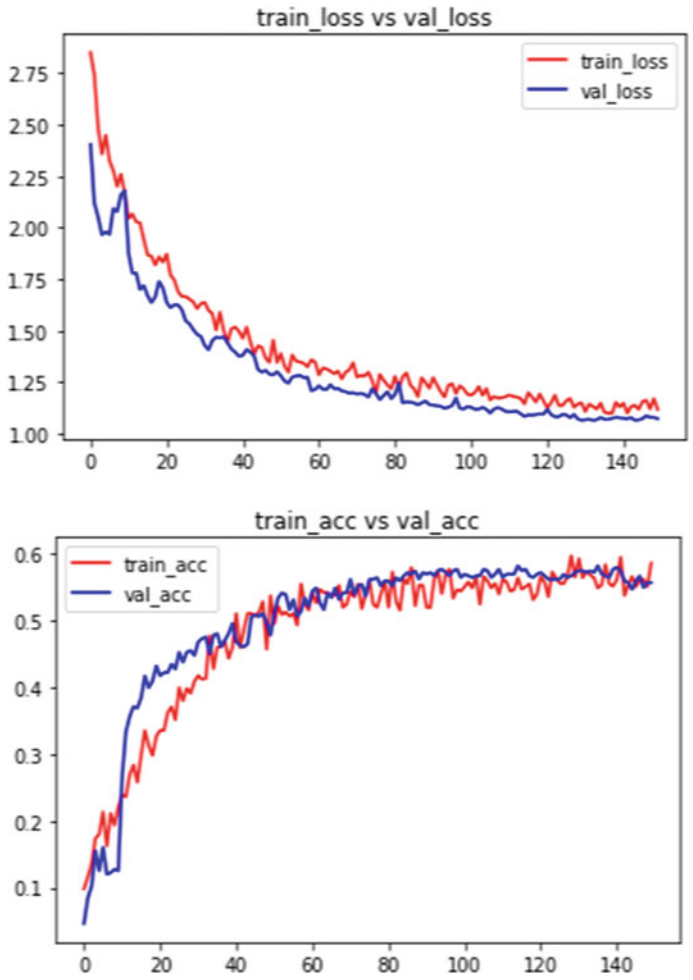



**Fig. 4** FER-accuracy and loss

Change, to identify the predominant emotion. A custom mapping is made, and the output of the FER model are used in a final mapping which returns a list of probable diseases as shown in Fig. 5. To improve the accuracy of the model, the features of the EEG, i.e., Valence, Arousal, Dominance and Liking could be taken and used further to ascertain the major emotion for certainty.

| Subject | Emotion | Change | Disease | Subject | Emotion | Change | Disease |
|---|---|---|---|---|---|---|---|
|  | Anger | Drastic | Mild Cognitive Impairment |  | Anger | No | Alzheimer's |
| | Sadness | Mild | | | Sadness | Drastic | |
| | Fear | Drastic | | | Disgust | Mild | |
| | Happiness | No | | | Surprise | Mild | |
|  | Anger | Mild | Epilepsy |  | Anger | Mild | Huntington's Chorea |
| | Sadness | Mild | | | Sadness | Mild | |
| | Fear | Drastic | | | Disgust | Drastic | |
| | Happiness | No | | | Happiness | No | |
|  | Anger | Drastic | Parkinson's |  | Anger | Drastic | Sclerosis |
| | Sadness | Drastic | | | Sadness | No | |
| | Fear | Drastic | | | Disgust | Drastic | |
| | Happiness | Mild | | | Surprise | Drastic | |

**Fig. 5** Final disease classification results (The subjects used in this demonstration are not actual patients but participants of the DEAP dataset. Their videos were used in checking the validity of the model.)

## 5 Discussion

The effectiveness of the proposed FER-EEG model is compared against state-of-the-art diagnostic models of various neuro-diseases. Models proposed by Ravi et al. [6], Yagis et al. [7], Li et al. [8] and Yolcu et al. [14] were considered for the comparison. Fahd A. Alturki et al. [9] demonstrated the effectiveness of EEG as a gold-standard diagnostic tool against various neuro-diseases. The work done by Marcó-García et al. [5] suggests that variations in FER has a standalone effect on prediction of neuro-diseases. There are several attempts made by Li et al. [8] to combine models that read both EEG and Facial images to create a more efficient model for Emotion Recognition. Results obtained in terms of classification accuracy is tabulated in Table 1. The table shows that there is a difference in accuracy between the proposed system and existing systems (as shown by [6–8, 14]), which is due to the latter directly computing the FER values. The proposed model, on the other hand, computes the difference in FER (as shown in Eq. (1)) to classify the change infer as drastic, mild, or no change and takes the test set from the live video feed, rather than CK+. As a result of the higher data variation and the change in FER being relative to the timeframe, there is a noticeable gap in accuracy. Hence, the comparison with the models can only be indirect, and any conclusions drawn should be regarded as proportional rather than absolute. Despite the decrease in accuracy when compared to existing systems, the change in FER as an emotion parameter is significant. The accuracy of the EEG model is comparable to the existing systems, and is bound to improve over time, when deployed.

**Table 1** State of the art comparison

| Methodology | Database | Architecture | Performance (Accuracy %) |
|---|---|---|---|
| Proposed FER model | CK+ | CNN (FER) | 90.11 (FER Deficit) |
| Proposed EEG model | DEAP | ANN (EEG) | 61.5 (EEG) |
| Ravi et al. | CK+ | CNN (FER) | 97.2 |
| Yagis et al. | PMMI & OASIS | ResNet50 (EEG) | 67.3, 67.1 (Subject Based) |
| Li et al. | Custom | Facial Landmark Localization (FER) | 86.94 |
| Yolcu et al. | RaFD | MUG & CNN (FER) | 94.44 |

## 6 Conclusion

The purpose of this research is to identify a viable means of detecting symptoms of neuro-diseases, due their degenerative nature if left untreated. Existing methods of detection are mostly symptomatic in nature and if the user is unable to express their symptoms properly, they stand the risk of a false negative, and the disease progressing further. The symptoms are not very easy to detect, and they need either the patient's statement about their discomforts, or an EEG reading, which is not readily available. Due to the medical nature of the application, there is danger of false positives as well. Appropriate sets of negative and positive cases are to be checked accordingly.

## References

1. Mellouk W, Handouzi W (2020) Facial emotion recognition using deep learning: review and insights. Procedia Comput Sci 175:689–694. The 17th International Conference on Mobile Systems and Pervasive Computing (MobiSPC), The 15th International Conference on Future Networks and Communications (FNC), The International Conference on Sustainable Energy Information Technology
2. Maddula NVSS, Nair LR, Addepalli H, Palaniswamy S (2021) Emotion recognition from facial expressions using Siamese network. In: Thampi SM, Piramuthu S, Li, KC, Berretti S, Wozniak M, Singh, D (eds.) Machine Learning and Metaheuristics Algorithms, and Applications. SoMMA 2020. Communications in Computer and Information Science, vol 1366, pp 63–72. Springer, Singapore. https://doi.org/10.1007/978-981-16-0419-5_6
3. Keshari T, Palaniswamy S (2019) Emotion recognition using feature-level fusion of facial expressions and body gestures, pp 1184–1189
4. Grabowski K, et al (2019) Emotional expression in psychiatric conditions: new technology for clinicians. Psychiatry Clin Neurosci 73(2):50–62
5. Marcó-García S, Ferrer-Quintero M, Usall J, Ochoa S, Del Cacho N, Huerta-Ramos E (2019) Facial emotion recognition in neurological disorders: a narrative review. Rev. Neurol 69(5):207–219
6. Ravi R, Yadhukrishna SV (March 2020) A face expression recognition using CNN & LBP. In: 2020 fourth international conference on computing methodologies and communication (ICCMC), pp 684–689. IEEE

7. Yagis E, De Herrera AGS, Citi L (November 2019 ) Generalization performance of deep learning models in neurodegenerative disease classification. In: 2019 IEEE international conference on bioinformatics and biomedicine (BIBM), pp 1692–1698. IEEE
8. Li D, Wang Z, Gao Q, Song Y, Yu X, Wang C (2019) Facial expression recognition based on electroencephalogram and facial landmark localization. Technol Health Care 27:1–15
9. Alturki FA, AlSharabi K, Abdurraqeeb AM, Aljalal M (2020) EEG signal analysis for diagnosing neurological disorders using discrete wavelet transform and intelligent techniques. Sensors (Basel Switzerland) 20:2505. 32354161[pmid]
10. Zhu B, Taghavi M, Shoaran M (October 2019 ) Cost-efficient classification for neurological disease detection. In: 2019 IEEE biomedical circuits and systems conference (BioCAS), pp. 1–4. IEEE
11. Valliani AA-A, Ranti D, Oermann EK (2019) Deep learning and neurology: a systematic review. Neurol. Ther. 8:351–365
12. Bulusu S, Sai Surya Siva Prasad R, Telluri P, Neelima N (2021) Methods for epileptic seizure prediction using EEG signals: a survey. In: Hemanth D, Vadivu G, Sangeetha M, Balas V (eds.) Artificial Intelligence Techniques for Advanced Computing Applications. LNNS, vol 130, pp 101–115. Springer, Singapore. https://doi.org/10.1007/978-981-15-5329-5_10
13. Patel UK et al (2021) Artificial intelligence as an emerging technology in the current care of neurological disorders. J Neurol 268:1623–1642
14. Yolcu G, Oztel I, Kazan S, Oz C, Palaniappan K, Lever TE, Bunyak F (2019) Facial expression recognition for monitoring neurological disorders based on convolutional neural network. Multimed Tools Appl 78:31581–31603
15. Gautam R, Sharma M (2020) Prevalence and diagnosis of neurological disorders using different deep learning techniques: a meta-analysis. J Med Syst 44:49

# Comparative Analysis of Machine Learning Approaches for the Early Diagnosis of Keratoconus


**P. Subramanian, G. P. Ramesh, and B. D. Parameshachari**

**Abstract** Keratoconus refers to the condition of the eye, where the corneal thinning occurs accompanied by a change in the corneal shape. A cone shaped protrusion occurs in the cornea and hence the name keratoconus. This manifests in the forms of vision problems. Treatment of the Keratoconus ranges from contact lenses to corneal transplants. Detection of Keratoconus at earlier stages is important to prevent visual repair or vision loss or costly treatments. Detection of keratoconus is being done by clinical signs including Fleischer's rings, Munson's sign, Rizutti's sign, Vogt's Striae and Hydrops. Advanced techniques like Keratoscope and Videokeratoscope are being used for diagnosis of Keratoconus. With the advent of machine learning, new techniques including multilayer perceptron, Radial Basis Function Neural Network (RBFNN), Decision Trees and Convolutional Neural Network (CNN) are being developed in the automatic detection and screening of various disorders. This survey paper reviews the machine learning approaches in diagnosis of Keratoconus.

**Keywords** Classification · Convolutional neural networks · Diagnosis · Keratoconus · Machine learning

## 1 Introduction

Keratoconus is a medical condition where the cornea changes in shape from dome shaped to conical shape due to the reduction in corneal thickness. The cornea of human eyes is oval shaped with diameters in the range of 10.6 to 11.7 mm along the axes. The thickness of the cornea ranges from 0.5 mm at the center and increases to 1 mm on the peripheral. The cornea has a spherical shape in the optic zone and flattens towards the periphery. The thinning of the corneal stroma leads to Keratoconus, called

P. Subramanian (✉) · G. P. Ramesh
St Peter's Institute of Higher Education and Research, Chennai, India
e-mail: psurya602@gmail.com

B. D. Parameshachari
GSSS Institute of Engineering and Technology for Women, Mysuru, India
e-mail: paramesh@gsss.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_23

because of the conical cornea. Keratoconus can be initially unilateral and can become bilateral affecting both eyes. Symptoms depend on the stage of the keratoconus with the early stages getting unnoticed. The advanced stages, though, can be identified by vision loss and protrusion. Diagnosis and classification of Keratoconus has gained significance with the advent of imaging technologies and machine learning. In this article we review the signs of Keratoconus, diagnosis of Keratoconus and the machine learning techniques for improved and early diagnosis.

## 2 Signs and Diagnosis of Keratoconus

In Greek Kerato means cornea and Konos means cornea. Hence the name Keratoconus has been coined for the condition of cone shaped cornea. Keratoconus can occur in both eyes of the same person; meaning it can be bilateral or unilateral. It is also an asymmetric disease meaning that the degree of affection varies between the eyes. Keratoconus is a corneal irregularity that presents itself in the form of progressive protrusion accompanied by thinning, leading to irregular astigmatism and visual impairment. Keratoconus is caused by reduction in thickness of inferior temporal and central cornea. The thinning of cornea causes reduction of intraocular pressure, which results in the deformation of the cornea into a conical shape as in Fig. 1 and hence the name keratoconus. Though Keratoconus normally starts in the age group of 11 to 19, earlier and later onset have also been reported. Once after the onset the disease progresses till the 40 s and gets stabilized. Though keratoconus is bilateral the onset of the condition varies from one eye to another. In some cases, it might even take years, as much as 16 years, for the other eye to show marked effect from the time the first eye shows symptoms for the aberration. Sawaguchi et al. [1], have discussed on Keratoconus manifesting itself as ruptures and discontinuities of the Bowman's layer. Takahashi et al. [2] have studied the Transmission Electron Microscopy of diseased tissue to observe that the count of collagen lamellae to be less than that in the normal tissue. Shetty et al. [3] have observed folds and ruptures of Descemet's membrane in keratoconus eyes. In the beginning or incipient or forme fruste stage the symptoms are barely noticed with the chances of practitioners missing them altogether. Vision accuracy loss which cannot be corrected by wearing lens spectacles sets in with the progression of the disease from the incipient stages. Charleux oil drop, Fleicher's rings, Vogt's Striae, corneal nerve visibility, corneal opacity, corneal scarring, Munson's sign, Rizutti's sign and Hydrops are some of the indications of the onset of keratoconus. Fleicher's ring is formed due to the accumulation of iron deposits on the cornea due to changes in corneal curvature which are induced by the keratoconus. In Vogt's Striae the compression of Descemet's membrane results in the appearance of vertical lines in the corneal stroma, deep and fine. When the patient looks downward a V shaped deformation occurs in the lower eye-lid and is called the Munson's Sign. Rizutti's Sign refers to the condition where a light focused on the temporal area of the limbus results in a reflection that is bright and from the limbus nasal area. Hydrops is sudden vision loss caused by breaks in Descemet's

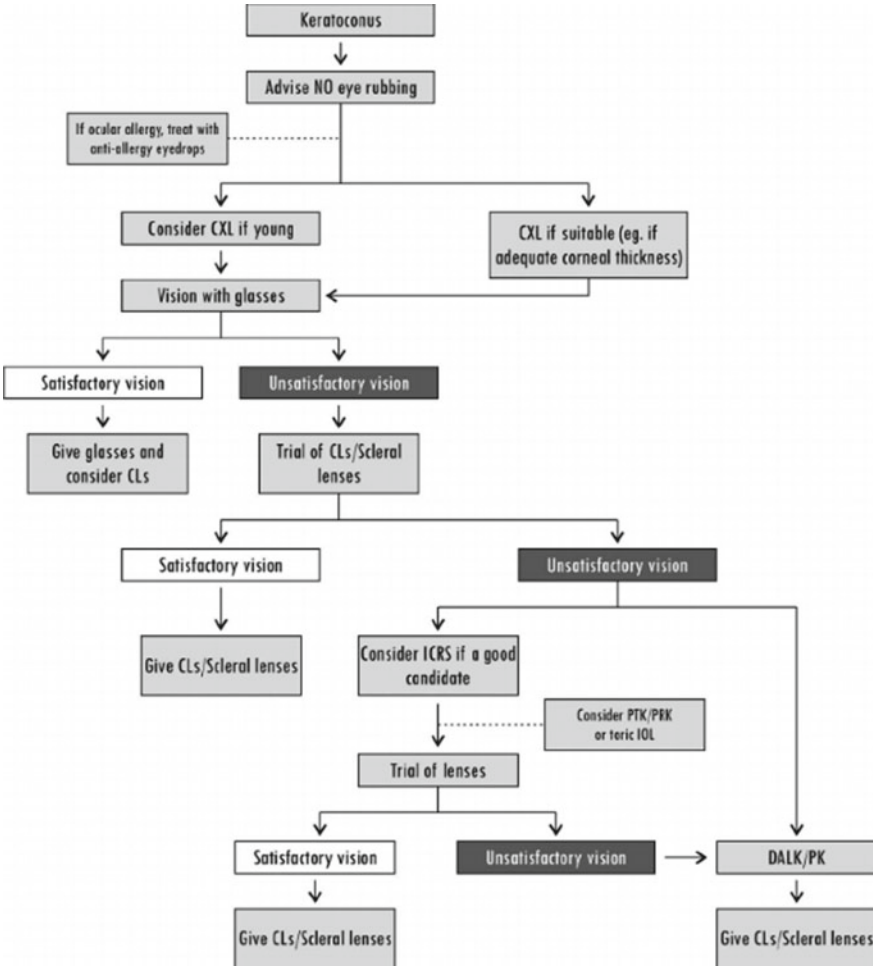



**Fig. 1** Taxonomy diagram Keratoconus detection

membrane. Mazen Sinjab [4] has written on the external signs including Munson's sign and Rizutti's sign as the lower eye lid margin forming a V shape when the patient looks downward and the nasal sclera getting lighted due to temporal illumination of the cornea. Focal thinning at the cone apex is another sign that can be used to confirm the Keratoconus. The focal thinning can be observed either in slit-lamp or OCT.

## 3 Machine Learning for Diagnosis of Keratoconus

Automatic diagnosis or the application of machine learning techniques help in the faster and accurate diagnosis more so as early screening methods. Placido based Keratoscope has been used in earlier period for detection of Keratoconus. The placido rings are thin and squeezed together in flat regions whereas they are dispersed in flat curvature regions. In normal cornea the placido rings are circular whereas oval rings suggest astigmatism. The rings are distorted and grouped in cone regions. Videokeratoscope in general study the specular reflection rather than directly measuring the corneal elevation. Sagittal Curvature is the radius of arc of a circle which is centered on the optical axis of Videokeratoscope and aligned to the corneal vertex. The tangent from the cornea to the point under consideration is the same as that of the Sagittal Curvature. Since accurate morphological description of the cornea is not possible with the Sagittal Curvature alone it became necessary to have tangential curvature also to describe the corneal steepening. Topographical maps representing elevation data have been used for more clear understanding of the corneal elevation data. The elevation maps need reference surfaces which are mostly best fit surfaces like spherical, aspherical or aspherotoric. The choice of the best fit affects the measurement of the corneal height. Simultaneous calculation of the curvatures and elevations in arc step has led to measurement of all morphological variations of the corneal shape though they face the disadvantage of skew ray error approximation. Optical scanning devices, like the Scheimpflug camera have the advantage of high transverse coverage and the measurement of whole of anterior chamber. Optical tomography coherence (OCT) is based on the measurement of interference of two beams of broadband radiation from a reference arm and a sample arm. Different classes of OCTs including Time Domain OCT (TD-OCT), Fourier Domain OCT (FD-OCT), Spectral Domain OCT (SD- OCT) and Swept Source OCT (SS- OCT) are used in the diagnosis of Keratoconus. With the advent of multimodal data and large size of the data machine learning plays an important role in the diagnosis of diseases. Machine learning helps in identifying new patterns and thereby new classes. Feature selection and extraction plays an important role in the efficiency of the machine learning algorithms. Machine learning algorithms have been used in diagnosis of various diseases. Of particular interest is the application of machine learning for the diagnosis of Keratoconus. We review here the different machine learning techniques that have been employed for the diagnosis and grading or classification of Keratoconus eyes. The taxonomy diagram is graphically stated in Fig. 1.

## 4 Related Works

Accardo et al. [5] have used a three layered Neural Network with Back propagation training for classification as Normal, Keratoconus and Others. Different architectures were considered by varying the number of input, hidden and output layers. Maps from single eye and both eyes of the same patient have been considered for classification. Different combinations of layer sizes and learning rate parameters have also been considered resulting in varying global sensitivities. The authors have achieved global sensitivity ranging from 81.7 to 94.1 and global accuracy of 92.9 to 96.4. The results vary with consideration of the single or both eyes. With single eye the accuracy level decreases.

Souza et al. [6] have considered only one eye of each patient. Eleven attributes are shortlisted from Orbscan II and are applied as input parameters for different algorithms namely Radial Basis Function Neural Network (RBFNN), Multilayer Perceptron (MLP) and Support Vector Machine (SVM). The parameters used included Anterior and Posterior best fit spheres, Maximum and Minimum Simulated Keratometry, Maximum Anterior and Posterior Elevation, 5 mm irregularity, thinnest point and central corneal power. The system provides detection of keratoconus but does not provide classification of stages of keratoconus. Different combinations of attributes need to be considered apart from the eleven attributes considered and data from both eyes of the patient can be considered. MLP, RBFNN, SVM and Decision Tree (DT) are employed by Toutounchian et al. [7] with direct features from Pentacam and features derived from topographical images. The attributes obtained directly obtained from Pentacam are Corneal Thickness, Anterior and Posterior best-fit spheres, Progression index, Sagittal Curvature, Tangential Curvature, Relative Pachymetry, Corneal Thickness Spatial Profile and Age. The second group of features consists of Randleman Score System, Ambrosia Score System and Symmetry in Sagittal Map. Though better accuracy for distinguishing between Normal and Keratoconus eyes has been obtained, the accuracy level decreases when suspect- to- Keratoconus is involved. Arbelaez et al. [8] have classified as Abnormal, Keratoconus, Subclinical keratoconus and Normal using support vector machine (SVM). Four groups of eyes, based on clinical diagnosis, namely, clinically diagnosed keratoconus, clinically diagnosed sub keratoconus, eyes undergone refractive surgery and normal eyes. The precision of classification increases in the distinction between normal and subclinical cases, when posterior corneal surface data are included. The accuracy levels are the least with respect to the subclinical Keratoconus in both cases. Hidalgo et al. [9] have used support vector machine (SVM) with linear kernel that employs dimensionality reduction. The eyes are fist classified by specialists into Normal, Astigmatism, Post Refractive surgery (PRS), Forme Fruste (FF) and Keratoconus. Pentacam provides around 1000 parameters which include correlated variables raising the need for reduction to reduce computation. Correlation based hierarchical clustering is employed resulting in a dendrogram followed by selection of one variable from each branch.

This resulted in the selection of 22 parameters. The classification performance indicates that the performance is less for forme fruste Keratoconus classification, with the sensitivity as low as 37.3%, compared to others.

Lavric et al. [10] utilized a convolutional neural network (CNN) for extracting and learning the features of a Keratoconus eye. Topographic maps with color scales are used for analysis. Here warm colors are used to represent curved/steep areas and mild colors are sued for flat areas. SyntEyes KTC model has been used to generate input data to overcome the difficulty in obtaining real data from patients. The accuracy obtained by Lavric is in the range of 99.33% and this can be used as a screening mechanism by ophthalmologist. The performance of the system depends on the image steps used; with smaller steps increasing the sensitivity and larger steps resulting in missing the mild cases of Keratoconus. Santos et al. [11] have employed a customized fully Convolutional Neural Network, CorneaNet, for early detection of Keratoconus. The network is used for segmentation of OCT images. The thickness maps of full cornea, Bowman layer, epithelial layer and Stroma are used and the segmentation speed considerably improved with the usage of the CorneaNet. In cornea with Keratoconus the thin zone could be detected and used for early detection of Keratoconus with 99.5% accuracy. Kamiya et al. [12] have obtained improved diagnostic accuracy by using deep learning on six color coded maps, namely anterior elevation map, anterior curvature map, posterior elevation map, posterior curvature map, refractive power map and pachymetry map. Six neural networks were trained for classification into 4 grades by taking into the average of the six outputs. Comparison of accuracy of the classification between normal and Keratoconus eyes shoes that the accuracy improves when the six maps were used when compared with the usage of the individual maps. Lavric et al. [13] have integrated the data generated in the SyntEyes algorithm with anterior and posterior Keratometry along with pachymetry. The anterior Keratometry is added with the input CNN network. The network provides the classification as normal and Keratoconus eyes. Shi et al. [14] have used neural network based machine learning for classification of eyes into Normal, Keratoconus and sub-clinical Keratoconus eyes. Combined features from Scheimpflug based camera and Ultra High Resolution Optical Coherence Tomography (UHR-OCT) provided better discrimination between the three classes than the features obtained from a single instrument. The UHR-OCT provides better discrimination between the sub-clinical and Clinical Keratoconus. Kuo et al. [15] have used three well known CNN models, namely VGG -16, Inception –V3 and ResNet-152, for detecting the pattern differences between normal and keratoconic eyes. The networks provided accuracy of 0.931 to 0.958. It has been observed that the results depend on the patterns and not on the color scales used and hence can be used across different platforms. Also the classification of sub-clinical cases was less satisfactory compared to the keratoconus cases.

Lavric et al. [16] have compared the performances of 25 different models in machine learning with an accuracy ranging from 62 to 94.0%. The best performance of 94% has been observed in a Support Vector Machine (SVM) that used eight corneal parameters which were selected using feature selection. Thulasidas et al. [17] have observed that the usage of Pachymetric Progression index and the BAD-D value resulted in improved detection of sub-clinical Keratoconus when compared to the

use of other Pentacam parameters as inputs. It is also observed that a combination of data results in improved detection of sub-clinical Keratoconus rather than that of a single parameter. Mustafa Koc et al. [18] have observed that though the topo metric and tomographic parameters are useful in earlier detection of keratoconus, for better performance individual parameters are not sufficient enough and combinations of parameters are needed for better performance. Gayathri et al. [19] have used 3-dimensional modeling of cornea for the diagnosis of Keratoconus. The correlation between the posterior and anterior corneal surfaces has been used to improve the performance of the modeling. Hallet et al. [20] employed a multilayer perceptron and variational auto encoder for classification of Keratoconus and Normal eyes. The unsupervised variational auto encoder gives better results with 80% accuracy when compared with the supervised multilayer perceptron with 73% accuracy.

Mahmoud et al. [21] have presented the 3-dimensional construction of corneal images from the 2-dimensional front and lateral images of the eye for detection of Keratoconus and its stages. The Convolutional Neural Network used for feature extraction uses the constructed 3-dimensional image. The accuracy of diagnosis obtained is 97.8% with a sensitivity of 98.45% and specificity of 96.0%. Feng et al. [22] have proposed a customized CNN, KerNet, for detecting Keratoconus and sub-clinical Keratoconus, from the Pentacam data. The data selected from Pentacam included the curvature of the front and back surfaces, elevation of the front and back surfaces and the pachymetry data in the form of numerical matrices. The network proposed, KerNet, uses multilevel fusion architecture. The proposed network improves the accuracy of the detection of Keratoconus by 1% and sub-clinical Keratoconus by 4%. Lavric et al. [23] have employed support vector machine to classify the different severity levels of the Keratoconus apart from the regular classification of Normal, Sub-clinical and Keratoconus. Starting with 2-class classification of healthy and keratoconus eyes with the highest Area under Curve (AUC) of 0.99, the work has been extended to a 3-class Classification of Normal, Forme Fruste and Keratoconus eyes with an AUC of 0.96. Further extension to a five class classification as Normal, Forme Fruste, Stage II, Stage III and Stage IV yielded an AUC of 0.88. Zaki et al. [24] used lateral segment photographed images (LSPI) which were extracted from videos of patients taken sideways with a smartphone camera. VGGNet 16 model is used for the classification achieving 95.75% accuracy, 92.25% sensitivity and 99.25% specificity. The comparative analysis of the conventional methods is explained in Table 1. In this table, latest machine learning approaches have been explained with limitations.

**Table 1** Comparative analysis

| Author | Year | Methodology | Advantage | Limitation | Performance analysis |
|---|---|---|---|---|---|
| Hallett et al. [20] | 2020 | Deep learning | Machine learning helped to analyses data from 124KC patients | Unsupervised algorithm selected only 29 variables only | True positive rate = 1.0 Accuracy = 85% |
| Mahmoud et al. [21] | 2020 | 3D Corneal Images Reconstruction | The image construction worked effectively with less lose | There may be a no obvious change in the earliest Keratoconus stages | True positive rate = 0.8 Accuracy = 80% |
| Feng et al. [22] | 2021 | Deep learning approach using raw data | The detection of the issue can be predicted earlier | There is an insufficient image availability for each disease stages, due to increased posterior evaluation | True positive rate = 0.7 Accuracy = 83% |
| Lavric et al. [23] | 2021 | Machine learning approach using topography | All type of raw data can be measured by this method | lower thinnest corneal thickness | True positive rate = 0.6 Accuracy = 70% |
| Zaki et al. [24] | 2021 | Lateral Segment Photographed Images | The segmentation helped to identify the problem easily | Measures only anterior corneal surface | True positive rate = 0.75 Accuracy = 88% |

## 5 Conclusion

Machine learning techniques particularly Support Vector Machines and Convolutional Neural Networks have been employed for efficient diagnosis of Keratoconus and its stages. The performance of the algorithms depends on the parameters chosen as input, unilateral or bilateral eyes, anterior or posterior, single parameter or multiple parameters and also on fusion of different modalities. The performance also depends on if the sub-clinical or forme fruste Keratoconus cases are involved. The detection of sub-clinical Keratoconus is important for early diagnosis of Keratoconus and for avoiding vision loss. The choice of parameters also decides the accuracy levels of the classification and diagnosis.

Algorithms including Back Propagation Neural Networks (BPN), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Radial Basis Function Network (RBFN), Decision Tree (DT) and Convolutional Neural Network (CNN) have

been used for detection and classification of Keratoconus based on the corneal attributes/indices. Though hundreds of corneal indices are available from scans, only a few have been used for detection of keratoconus. The output and accuracy depends on the choice of the attributes/parameters. No fixed rule is available on the parameters to be used. The accuracies of the algorithms in the literature are better for classification between Normal and Keratoconus cornea. But when there is a need to classify/ distinguish between Normal & Sub-clinical Keratoconus or Sub-clinical & Keratoconus cornea the accuracy decreases. Only few efforts have been taken in the detection and classification of Keratoconus using topography elevation images. Convolutional Neural Networks have been employed for detection of Keratoconus alone and not for classification, particularly not for the detection of sub-clinical Keratoconus.

## References


1. Sawaguchi S, Fukuchi T, Abe H, Kaiya T, Sugar J, Yue BY (1998) Three-dimensional scanning electron microscopic study of keratoconus corneas. Arch Ophthalmol 116(1):62–68
2. Takahashi A, Nakayasu K, Okisaka S, Kanai A (1990) Quantitative analysis of collagen fiber in keratoconus. Nippon Ganka Gakkai Zasshi 94(11):1068–1073
3. Shetty R, Sathyanarayanamoorthy A, Ramachandra RA, Arora V, Ghosh A, Srivatsa PR (2015) Attenuation of lysyl oxidase and collagen gene expression in keratoconus patient corneal epithelium corresponds to disease severity. Mol Vis 21:12–25
4. Sinjab MM (2012) Quick Guide to the Management of Keratoconus: A Systematic Step-by-Step Approach. Springer-Verlag, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-21840-8
5. Accardo PA, Pensiero S (2002) Neural network-based system for early keratoconus detection from corneal topography. J Biomed Inform 35(3):151–159. https://doi.org/10.1016/s1532-0464(02)00513-0
6. Souza MB, Medeiros FW, Souza DB, Garcia R, Alves MR (2010) Evaluation of machine learning classifiers in keratoconus detection from orbscan II examinations. Clinics (São Paulo, Brazil) 65:1223–1228. https://doi.org/10.1590/S1807-59322010001200002.
7. Toutounchian F, Shanbehzadeh J, Khanlari M (2012) Detection of keratoconus and suspect keratoconus by machine vision. In: Proceedings of the International Multi-Conference of Engineers and Computer Scientists, Hong Kong, China, March 2012
8. Arbelaez MC, Versaci F, Vestri G, Barboni P, Savini G (2012) Use of a support vector machine for keratoconus and subclinical keratoconus detection by topographic and tomographic data. Ophthalmology 119(11):2231–2238
9. Hidalgo IR, Rodriguez P, Rozema JJ (2016) Evaluation of a machine-learning classifier for keratoconus detection based on Scheimpflug tomography. Cornea 35(6):827–832
10. Lavric A, Valentin P (2019) KeratoDetect: Keratoconus detection algorithm using convolutional neural networks. Computat Intell Neurosci. https://doi.org/10.1155/2019/8162567
11. Dos Santos VA et al (2019) CorneaNet: fast segmentation of cornea OCT scans of healthy and Keratoconic eyes using deep learning. Biomed Opt Express 10(2):622–641. https://doi.org/10.1364/BOE.10.000622
12. Kamiya K et al (2019) Keratoconus detection using deep learning of colour-coded maps with anterior segment optical coherence tomography: a diagnostic accuracy study. BMJ Open **9**(9). https://doi.org/10.1136/bmjopen-2019-031313

13. Lavric A, Popa C, David C, Paval CC (2019) Keratoconus detection algorithm using convolutional neural networks: challenges. In: 2019 11th International Conference on Electronics, Computers and Artificial Intelligence, pp 1–4. https://doi.org/10.1109/ECAI46879.2019.9042100
14. Shi C, Wang M, Zhu T et al (2020) Machine learning helps improve diagnostic ability of subclinical keratoconus using Scheimpflug and OCT imaging modalities. Eye and Vis 7:48. https://doi.org/10.1186/s40662-020-00213-3
15. Kuo B-I et al (2020) Keratoconus screening based on deep learning approach of corneal topography. Trans Vis Sci Tech 9(2):53. https://doi.org/10.1167/tvst.9.2.53
16. Lavric A, Popa V, Takahashi H, Yousefi S (2020) Detecting keratoconus from corneal imaging data using machine learning. IEEE Access 8:149113–149121. https://doi.org/10.1109/ACCESS.2020.3016060
17. Thulasidas M, Teotia P (2020) Evaluation of corneal topography and tomography in fellow eyes of unilateral keratoconus patients for early detection of subclinical keratoconus. Indian J Ophthalmol 68(11):2415–2420. PMID: 33120630. https://doi.org/10.4103/ijo.IJO_2129_19
18. Koc M et al (2020) Topometric and tomographic evaluation of subclinical Keratoconus. Ophthalmic Epidemiol 27(4):289–297
19. Kanimozhi R, Gayathri R (2020) Keratoconus detection based on corneal morpho-geometric analysis using correlation. In: 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), pp 203–206. https://doi.org/10.1109/ICISS49785.2020.9316066
20. Hallett N et al (2020) Deep learning based unsupervised and semi-supervised classification for Keratoconus. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp 1–7. https://doi.org/10.1109/IJCNN48605.2020.9206694
21. Mahmoud HAH, Mengash HA (2021) Automated Keratoconus detection by 3D corneal images reconstruction. Sensors (Basel) 21(7):2326. https://doi.org/10.3390/s21072326
22. Feng R et al (2021) KerNet: a novel deep learning approach for keratoconus and sub-clinical keratoconus detection based on raw data of the pentacam system. IEEE J Biomed Health. https://doi.org/10.1109/JBHI.2021.3079430
23. Lavric A et al (2021) Keratoconus severity detection from elevation, topography and pachymetry raw data using a machine learning approach. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3086021
24. Zaki WM, Daud MM, Saad AH, Hussain A, Mutalib HA: Towards automated Keratoconus screening approach using lateral segment photographed images. In: 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), pp 466–471 (2021). https://doi.org/10.1109/IECBES48179.2021.9398781

# Conversion of NAM to Normal Speech Based on Stochastic Binary Cat Swarm Optimization Algorithm


**T. Rajesh Kumar, G. N. Balaji, D. Vijendra Babu, Soubraylu Sivakumar, K. Kalaiselvi, and G. R. Suresh**

**Abstract** Speech recognition plays an important role in a variety of applications for mobile communication. User communication devices for contact necessitate a broad vocabulary recognition scheme, greater precision, and a real-time, low-power schema. The power consumption and memory bandwidth of miniaturized battery-controlled devices are important. People's handheld devices often demand more effort, due to the speech challenge. As a result, a valuable technology based on the Stochastic Binary Cat Swarm Optimization Algorithm (SBCSO) is proposed in this research study to transform the non-audible murmur to normal voice. From the input murmured speech signal, the characteristics such as spectral skewness, spectral centroid, pitch chroma, and Taylor-Amplitude Modulation Spectrum are extracted and trained in the Deep Convolutional Neural Network (DCNN) classifier. The proposed stochastic binary cat swarm optimization algorithm is used to train DCNN classifier for speech recognition. To boost the results in metric analysis, the stochastic gradient descent algorithm and a Binary Cat Swarm Optimization Algorithm (BCSOA) are combined. In order to boost the experimental results in metric analysis, the stochastic gradient descent algorithm and BCSOA are combined in this research paper. The proposed algorithm performance is validated in terms of true positive rate, false positive rate and classification accuracy, and it showed better improved in speech recognition.

T. R. Kumar (✉)
Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India
e-mail: t.rajesh61074@gmail.com

G. N. Balaji · G. R. Suresh
Department of CSE, SRM TRP Engineering College, Trichy, Tamil Nadu, India

D. V. Babu
Department of ECE, Aarupadai Veedu Institute of Technology, Chennai, India

S. Sivakumar
Computer Science Engineering, Sri Ramachandra Faculty of Engineering and Technology, Porur, Chennai, Tamil Nadu, India

K. Kalaiselvi
Department of CSE, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_24

**Keywords** Binary cat swarm optimization algorithm · Deep convolutional neural network · Mobile communication · Speech recognition · Stochastic gradient descent approach

## 1 Introduction

For the translation of audio signals or human speech to text, ASR systems are used. Furthermore, in many implementations, such as spoken text extraction, Integrated Voice Response (IVR) programs, and dictation interfaces, ASR is used. For the above-mentioned applications, however, a clear speech recognition system is needed. Moreover, notable advances have been made in some of the frameworks for automatic speech recognition [1]. ASR performance is determined by many influences, including the nature of input speech and forms and actual speaker features, such as accent, style and speech intensity, psycho-physical state of the speaker, anatomy of the vocal tract etc. [2, 3]. In speech technology, Whisper plays an important role because of its substantial variance relative to the normal phoned speech, which induces noisy form, absence of glottal sounds, and less Signal to Noise Ratio (SNR) [4, 5]. Murmur is a basic verbal communication system that is widely used under different situations. Whispered speech is created by adjusting the vocal cords while engaged in glottis development of constricted restraint, inducing vocal area excitation. To train DCNN for speech recognition, the proposed BCSOA algorithm is employed. The SBCSOA is a combination of the stochastic gradient descent and BCSOA. In order to improve device performance, the SBCSOA-DCNN is optimized for speech recognition.

## 2 Related Works

Due to recent advances in computer hardware technology and machine learning method, Deep Neural Network (DNN) is used successfully in the ASR framework [6]. Additionally, Secret Markov Model (HMM) with DNN approaches (DNN-HMM) was used in various studies to build series of audio observations augmented by HMM with Gaussian Mixture Model (GMM-HMM) in different occupations, such as massive vocabulary and very large scale datasets [7, 8]. In addition, different techniques have been developed to boost audio divergence through model adaptation, other sensing approaches, such as throat microphone, and transformation of features. Audio-visual technique was also applied to isolate word recognition under the condition of murmur speech [9]. The cat swarm optimization algorithm was introduced in this work, has only been used in function optimization problems. The aim of this study was to create a cat swarm optimization algorithm that can be used to solve clustering problems. The CSO algorithm has been shown to be effective in solving optimization problems [10]. CNN may be used to classify the speech signals

instantly and predict the murmured speech spoken by speech disables or physically challenged speakers.

## 3 Methodology

In order to produce non-audible whisper to ordinary voice, an efficient procedure, called binary cat swarm optimization algorithm is proposed. Initially, the input speech signal is pre-processed by filtering, and the required characteristics are extracted in the function extraction procedure. The system built by the introduction of the Taylor series into the regular AMS is Taylor-AMS. DCNN [11], in which the training algorithm is performed using Stochastic Binary Cat Swarm Optimization, which is built using stochastic gradient descent algorithm and BCSOA, performs speech recognition after the feature extraction method. The developed model then classifies the speech signal into ordinary speech based on extracted characteristics.

### *3.1 The Input Signal NAM Received Through Gadgets*

By considering the signals gathered from gadgets as the dataset, the speech signal is known. Let us assume the dataset as Y defined as the $n$ number of input speech signals and as defined in Eq. (1).

$$\mathrm{Y} = \{C_i\}\ ;\ \ i \in \{1, 2, \ldots, n\} \tag{1}$$

where, Y indicates dataset, $C$ specifies the input speech signal, and $n$ signifies total amount of speech signals present in dataset. However, the input video $C_i$ is selected from the dataset to perform the speech recognition.

### *3.2 Pre-processing*

To allow smooth processing of an input speech signal, the input speech signal $C_i$ is selected and given to the pre-processing level. Here, the Wiener filter and elimination of silence are used to pre-process speech signal data. The silence deduction, however, is the pre-processing procedure for eliminating silence and the voiceless divisions of the speech signal and the Wiener filter is used to filter noise from the input signal and $C_i^*$ is denoted as the pre-processed output.

### 3.3 Feature Extraction

The pre-processed signal is $C_i^*$ forwarded to the feature extraction module until the speech signal is pre-processed. The feature extraction is done here on the basis of Taylor-AMS, pitch chroma, spectral centroid, spectral skewness, and the mixture of the Dirichlet method to produce highly important features for improved speech signal recognition. In addition, the signal complexity is minimized, thus reducing the range of characteristics. Furthermore, the precision associated with classification ensures efficient extraction of the feature. As below, the function extraction phase followed in this paper is outlined. Furthermore, the precision associated with classification ensures efficient extraction of the feature.

### 3.4 SBCSO-Based Deep CNN

Once the appropriate features get extracted, speech recognition is done based on DCNN. The feature vector is taken as input of DCNN for the conversion of non-audible murmur to a regular speech. However, the DCNN [12] tunes the biases and weights of classifier for improving the classification accuracy. In addition, the Deep CNN is trained by developed optimization method, named SBCSO that is designed newly by incorporating stochastic gradient descent algorithm and CSOA. DCNN is utilized to convert the non-audible murmur to the normal speech to achieve optimal speech recognition output. The Deep CNN architecture consists of three separate layers, such as the convolutional layer, the pooling layer, and the completely linked layer. Each layer executes its own operations in the Deep CNN architecture. Each layer executes its own operations in the Deep CNN architecture. The convolutional layer is used here to measure the function diagram, and then the pooling layer performs sub-sampling. Finally, using the completely linked layer, the classification is performed. Classification precision is strengthened as the convolutional layers in the Deep CNN classifier are expanded. In addition, the first-layer neuron is paired with the following layer of individual neurons.

**Convolutional Layer.** Through considering the convolutional filter that links neurons from the previous layer to the next layer by trainable weights, the convolutional layer is used to generate patterns for femur points. Assume that the convolutional layer input is expressed in Eq. (2)

$$\left(U_m^n\right)_{g,h} = \left(T_m^n\right)_{g,h} + \sum_{z=1}^{e_1^{z-1}} \sum_{v=-e_1^j}^{e_1^n} \sum_{x=-e_2^j}^{e_2^n} \left(\varpi_{m,z}^n\right)_{x,y} * \left(F^{\omega}\right)_{g+v,h+x} \tag{2}$$

where, $\left(U_m^n\right)_{g,h}$ refer to the output of $n^{th}$ convolutional layer or the feature map, and $*$ refer to convolutional operator. $\varpi_{m,z}^n$ signifies the weight, and the term $T_m^n$ indicates the bias of $n^{th}$ convolutional layer.

**ReLU Layer:** For applying the non-saturating activation function, ReLU is the rectified linear unit. Here, by setting them to zero, the negative values are efficiently extracted from the activation diagram. As the input of the next $k^{th}$ convolutional layer, the output generated by $n-1^{th}$ layer is forwarded and expressed in Eq. (3)

$$U_m^n = A\left(U_m^{n-1}\right) \tag{3}$$

**Pooling Layer.** The pooling layer is often considered a non-parametric layer, and is used to execute a fixed operation that ignores bias and weight consideration.

**Fully Connected Layers.** The pooling layer output is taken as the input of the fully linked layer and the output from the fully linked layer is given in Eq. (4).

$$V_m^n = \sigma\left(U_m^{n-1}\right),\ U_m^n = \sum_{z=1}^{e_1^{z-1}} \sum_{\nu=-e_1^j}^{e_1^n} \sum_{x=-e_2^j}^{e_2^n} \left(\varpi_{m,z}^n\right)_{x,y} * \left(F^{\omega}\right)_{g+\nu,h+x} \tag{4}$$

where, the term $\left(\varpi_{m,z}^n\right)_{x,y}$ refer to weights. The weight values are computed for tuning Deep CNN such that the optimal weight is retrieved.

## *3.5 Training of Deep CNN Based on Stochastic Binary Cat Swarm Optimization*

The preparation of the DCNN classifier is performed to achieve a better weight factor using the established SBCSOA. The SBCSOA is a mixture of stochastic gradient descent techniques, Binary Cat Swarm Optimization for successful collection of optimum weights. The parametric characteristics obtained from the optimization algorithm described above allow efficient classification efficiency based on the input image. The stochastic algorithm is efficient since it is linear for training data and is capable of approximating the true gradient over time for each training data.

BCSO Algorithm, on the other hand was split into two sub-models based on two main behavioral characteristics of cats. Both are referred to as "seeking mode" and "tracing mode," respectively. In CSO, we first determine how many cats we'll need in the iteration, and then we use the cats to solve problems in CSO. Every cat has its own location, which is made up of D dimensions, velocities for each dimension, a fitness value representing the cat's accommodation to the fitness function, and a flag indicating whether the cat is searching or tracing. The optimal location of one of the cats will be the final solution. Thus, the stochastic gradient descent algorithm update

equation is updated using CSO. This move then helps the approach, with better efficiency, to be more effective. Below is the algorithmic method of the proposed SBCSOA.

*Step 1: Initialization.* The zero vectors are initialized in the first step, which is represented as $Z$, and select the training sample randomly represented as, $\left(H_l^h,\ M_l^h\right)$. The term $H_l^h$ denotes the feature vector of $l^{th}$ training sample with dimension $\eta$, and the term $M_l^h$ refers to category of training data.

*Step 2: Fitness function computation.* The fitness is estimated for finding optimal solution for speech recognition. Moreover, it is estimated by minimum error value, and solution with respect to minimal error is taken as best solution. The fitness is formulated using Eq. (5).

$$MSE = \frac{1}{\kappa}\left[\sum_{o=1}^{\kappa} U_{t\arg et} - U_m^n\right] \tag{5}$$

where, $\kappa$ refer to training samples, the terms $U_{t\arg et}$ and $U_m^n$ signifies the target and estimated output of classifier.

*Step 3: Update the solution using proposed SBCSOA.* SMP stands for searching memory pool, SRD stands for seeking selection of the chosen dimension, CDC stands for counts of dimension to modify, and self-position consideration stands for self-position consideration (SPC) are four main factors of seeking mode in CSO. According to Chu et al. Let FS be the fitness, then $\mathrm{FS_b} = \mathrm{FS_{max}}$ if the fitness function's objective is to find the smallest solution; otherwise, $\mathrm{FS_b} = \mathrm{FS_{min}}$. Tracing mode is a sub-model for simulating the cat's tracing of goals that is defined in the Eqs. (6) and (7).

$$V_d = V_d + k \times c \times (x_{best,d} - x_d) \tag{6}$$

where, $\mathrm{x_d}$ is the position of the cat 'k', constant is 'c' the cat position is $\mathrm{x_{best,d}}$, Vd is the velocity of the cat at every dimension 'd'.

$$x_d = x_d + V_d \tag{7}$$

To incorporate these two modes into the algorithm, we define a mixture ratio (MR), which specifies how searching and tracing modes should be combined. Clustering data and looking for the right cluster centre using the CSO algorithm are the two key aspects of the CSO Clustering suggested in this study [13, 14]. The population of data to be clustered, the number of clusters 'k', and the number of copies would be the inputs for clustering CSO. The steps involved in clustering CSO are shown below.

*Step i): Initialization of cluster centre in DCNN*

Pick k points at random from the data to serve as the initial cluster centre.

*Step ii): Swarming the data in to Clusters.*

Sort the data into clusters based on the cluster centre that is nearest to you. The Euclidean distance is used to measure the distance between the data and the cluster core. Equation can be used to define the distance between x and y using the Euclidean distance theorem, as defined in Eq. (8)

$$Ed(x, y) = |x - y|^2 = \sqrt{\sum_{j=1}^{n}(x_j - y_j)^2} \tag{8}$$

*Step iii): Determination of the (SSE) Sum of Squared-Error.*
The function of SSE is defined in Eq. (9)

$$SSE = \sum_{j=1}^{m}\sum_{x \in C}(||x - c_j)^2 \tag{9}$$

where as 'm'- clusters count, $c_j$ – centre of cluster j and x = data and it is the members of C.

*Step iv): Cat Swarm Optimization based cluster optimization.*

The CSO algorithm represents cat as cluster centre and the new cluster centre will be the speech solution set, which results with SSE lesser than the previous iterations. It has two mode combination and named as binary cat swarm. They are seeking and tracing mode.

Initially, the parameters such as SMP, SRD and SPC are defied in seeking mode. Copy cluster centre (i) location as many times as SMP for i = 1 to k (number of cluster centres). Determine the changing value (SRD*cluster centre (i)) and determine the j value. Roughly positive or negative cluster centres with changing value for m = 1 to SMP. Calculate the SSE after calculating the distance and clustering the results. Choose a new cluster centre rette wheel nominee. Now update the cluster centre and SSE. This is the seeking mode algorithm in Cat Swarm Optimization Clustering [15].

*Step v): Recheck the feasibility.* Based on fitness value, the feasibility is again computed, hence if the generated new solution is best than previous one, the old one is replaced by the new solution.

*Step vi): Termination.* The steps described above are repetitive before the best solution for speech recognition is obtained. Therefore, the speech signal characteristics are collected when the new signal arrives for speech recognition, and are forwarded to the classifier to process the characteristics in terms of new function dataset and create class mark for input speech signal. In recognizing words that are subject to many uses, the evolved speech recognition technique is successfully verified.

## 4 Simulation Results

The experimentation of developed model is developed in MATLAB tool with the windows 10 OS, 4 GB Ram, and the Intel I3 processor. TIMIT is used to conduct the experiment, and the overview of the datasets is given below: The TIMIT corpus includes a read speech to enable the automated processing of the speech recognition system. TIMIT consisted of 630 speakers of speech signal, which is rich in phonetics, orthographic time-aligned, and 16-bit speech waveform word transcriptions of 16 kHz for the actual utterance. The TIMIT amount transcripts are checked with the preparation and evaluation sub-sets with the required dialectal coverage and phonetics criteria.

### *4.1 Discussion*

Using three metrics, such as Accuracy, True Positive Rate (TPR), and False Positive Rate (FPA), the efficiency of the established approach is evaluated. Also, compared with existing techniques, such as Stochastic-Whale Optimization Algorithm + Deep Convolutional Neural Network (Stochastic-WOA + DCNN) [15], Random sampling-based Deep Neural Network (DNN) [1], DCNN [2], multi-task DNN acoustic models [4] and Deep Learning-Recurrent Neural Network (RNN) [8], the performance of the developed technique is evaluated.

**Analysis Using Sample Signal based on K-fold.** The analysis of approaches using a sample signal through K-Fold by considering accuracy, FPR, and TPR parameters is illustrated in Fig. 1. The analysis of techniques based on accuracy parameter is deliberated in Fig. 1a. For K-Fold = 5, the accuracy measured by Stochastic-WOA + DCNN is 0.952, Deep-CNN is 0.944, Random sampling-based DNN is 0.756, multi-task DNN acoustic models is 0.852, Deep learning-RNN is 0.713, and the proposed SBCSOA-DCNN is 0.96. In Fig. 1b, the analysis using the FPR parameter is deliberate. The FPR determined by Stochastic-WOA + DCNN is 0.02 when K-Fold

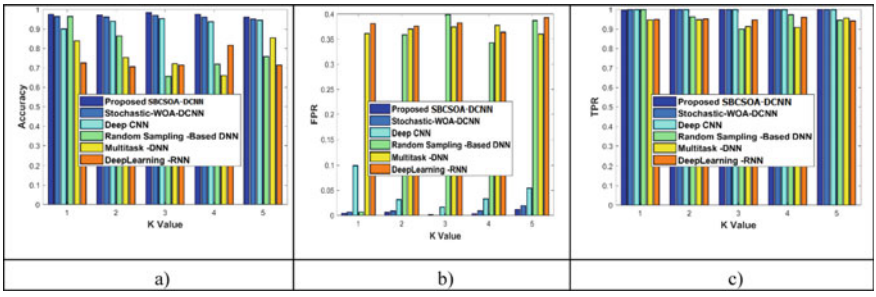



**Fig. 1** Analysis of K-Fold methods for a sample signal **a** Accuracy, **b** FPR, and **c** TPR

= 5, Deep-CNN is 0.055, Random sampling-based DNN is 0.388, multi-task DNN acoustic models are0.361, Deep learning-RNN is 0.394, and the implied SBCSOA-DCNN is 0.010. In Fig. 1c, the analysis using the TPR parameter is intentional. When K-Fold = 5, the TPR of methods computed by Stochastic-WOA + DCNN is 0.981, Deep-CNN is 0.973, Random sampling–based DNN is 0.944, multi-task DNN acoustic models is 0.955, Deep learning-RNN is 0.940, and the proposed SBCSOA-DPM is 0.998.

**Analysis Using Sample Signal through Holdout.** The analysis of methods through sample signal using holdout by considering accuracy, FPR, and TPR parameters is illustrated in Fig. 2. The analysis of approaches based on accuracy parameter is deliberated in Fig. 2a. When training data percentage is 90, accuracy estimated by Stochastic-WOA + DCNN is 0.967, Deep-CNN is 0.937, Random sampling–based DNN is 0.964, multi-task DNN acoustic models is 0.714, Deep learning-RNN is 0.666, and the proposed SBCSOA-DCNN is 0.984. The analysis using FPR parameter is deliberated in Fig. 2b. For 90% training data, FPR computed by Stochastic-WOA + DCNN is 0.014, Deep-CNN is 0.051, Random sampling–based DNN is 0.013, multi-task DNN acoustic models is 0.378, Deep learning-RNN is 0.387, and the proposed SBCSOA-DCNN is 0.013.

The analysis using TPR parameter is deliberated in Fig. 2(c). When training data percentage is 70, the TPR of methods computed by Stochastic-WOA + DCNN is 0.984, Deep-CNN is 0.968, Random sampling–based DNN is 0.999, multi-task DNN acoustic models is 0.930, Deep learning-RNN is 0.938, and the proposed SBCSOA-DCNN is 0.938. Both the K-fold and Hold out validation for the metrics of Accuracy, FPR and TPR are hereby tabulated in the Table 1 for the references.

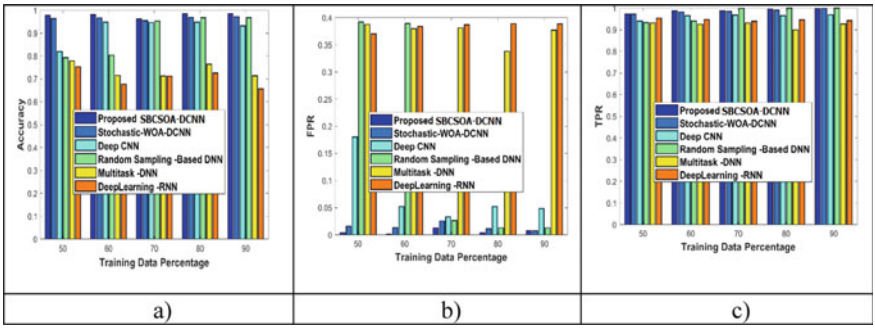



**Fig. 2** Analysis of methods using sample signal by varying the training data percentage **a** Accuracy, **b** FPR, and **c** TPR

**Table 1** Comparative discussion on sample signal for K-Fold and Holdout analysis

| Validation based on | Metrics | Stochastic-WOA + DCNN [15] | Deep-CNN [2] | Random sampling–based DNN [1] | Multi-task DNN acoustic models [4] | Deep learning-RNN [8] | Proposed SBCSOA-DCNN |
|---|---|---|---|---|---|---|---|
| K-Fold | Accuracy | 0.952 | 0.944 | 0.756 | 0.852 | 0.713 | 0.96 |
| | FPR | 0.02 | 0.055 | 0.388 | 0.361 | 0.394 | 0.010 |
| | TPR | 0.981 | 0.973 | 0.944 | 0.955 | 0.940 | 0.998 |
| Hold out | Accuracy | 0.967 | 0.937 | 0.964 | 0.714 | 0.666 | 0.984 |
| | FPR | 0.014 | 0.051 | 0.013 | 0.378 | 0.387 | 0.013 |
| | TPR | 0.984 | 0.968 | 0.999 | 0.930 | 0.938 | 0.938 |

## 5 Conclusion

An enhanced form of speech recognition based on Stochastic Binary Cat Swarm Optimization Algorithm based DCNN is proposed in this research work to convert non-audible murmurs to natural speech. Initially, the input signal is pre-processed using filtering to remove the noise in the signal. Next, the feature extraction is carried out based on Taylor AMS, pitch chroma, spectral centroid, and spectral skewness for further processing. Then, the speech recognition is carried out using extracted features using Deep CNN, which is trained using developed optimization algorithm, termed SBCSOA. The incorporation of stochastic gradient descent methodology and Binary Cat Swarm Optimization is the SBCSOA. Therefore, based on extracted functionality, developed SBCSOA-based DCNN classifies input signal as natural speech. Using the K-fold and holdout metric, the efficiency of the developed model is validated for accuracy, FPR, and TPR. Comparatively, the proposed model dominated other methods, which is validated using Accuracy, FPR, and TPR. In future, speech prediction will be achieved through consideration of another meta-heuristic algorithm and dataset.

## References

1. Deypir Zoughi T, Homayounpour MM, Deypir M (2020) Adaptive windows multiple deep residual networks for speech recognition. Expert Syst Appl 139:112840
2. Rajesh Kumar T, Suresh GR, Kanaga Subaraja S, Karthikeyan C (2020) Taylor AMS features and deep convolutional neural network for converting non-audible murmur to normal speech. Comput Intell 36(3):940–963. Wiley Publishers
3. Karthikeyan C, Sreedevi E, Kumar N, Vamsidhar E, Kumar TR, Babu DV (2020) Cost optimization in neural network using whale swarm algorithm with batched gradient descent optimizer. In: IOP Conference Series: Materials Science and Engineering
4. Agrawal P, Ganapathy S (2019) Modulation filter learning using deep variational networks for robust speech recognition. IEEE J Sel Top Signal Process 13(2):1–10
5. Kumar TR, Suresh GR, Raja S (2019) Conversion of non-audible murmur to normal speech based on GR-GMM using non-parallal training adaptation method. In: Proceedings of ICISS-2019. IEEE Xplore

6. Kumar ER, Rao AK (2019) Suicide prediction in twitter data using mining techniques: a survey. In: Proceedings of ICISS, pp 122–131. IEEE Xplorer
7. Bommadevara HSA, Sowmya Y, Pradeepini G (2019) Heart disease prediction using machine learning algorithms. Int J Innov Technol Explor Eng 8(5):270–272
8. Kumar S, Singh R (2021) Review and analysis of optimization algorithms of digital filters design. Turkish J Comput Math Educ 12(7):1798–1806
9. Kumar TR, Suresh GR, Raja S (2018) Conversion of non audible murmur to normal speech based on full-rank Gaussian mixture model. J Comput Theor Nano Sci (JCTN) 15(1):185–190
10. Ghaffarzadegan S, Bořil H, Hansen JH (2017) Deep neural network training for whispered speech recognition using small databases and generative model sampling. Int J Speech Technol 20(4):1063–1075
11. Mannepalli K, Sastry PN, Suman M (2017) Accent recognition system using deep belief networks for Telugu speech signals. Int J Speech Technol 19(1):87–93
12. Sreedevi E, Prasanth Y (2017) A novel class balance ensemble classification model for application and object oriented defect database. J Adv Res Dyn Control Syst 18:702–726
13. Anila M, Pradeepini G (2017) Study of prediction algorithms for selecting appropriate classifier in machine learning. J Adv Res Dyn Control Syst 9(Special Issue 18):257–268
14. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67
15. Shariff MN, Saisambasivarao B, Vishvak T, Rajesh Kumar T (2017) Biometric user identity verification using speech recognition based on ANN/HMM. J Adv Res Dyn Control Syst 9(12 Special issue):1739–1748

# Convolutional Neural Network Models for Throat Cancer Classification Using Histopathological Images


**Ravindranath Kadirappa, Gadipudi Amaranageswarao, and S. Deivalakshmi**

**Abstract** Cancer is a condition where the abnormal growth of cells is observed, initially identified as tumors. Cancer can happen to any part of the body, early detection of cancer can increase the life expectancy of patients in this paper throat cancer is considered for the study, cancer can start at any part of the throat and later stage can spread to the lungs and liver. Samples are extracted through the fine needle biopsy method and by using Hematoxylin, Eosin (H&E) stain images are obtained with three magnitudes original, 10×, 20× from an electron microscope. Convolution Neural Networks (CNN) with Support Vector Machine (SVM) and Transfer Learning (TL) are used to classify the images as either cancer or normal both techniques have shown promising results with an average accuracy of 93%. Transfer learning with 10× data has provided better results with an average accuracy of 94.18% as TL slightly outperformed SVM and itself with other magnitudes.

**Keywords** Throat cancer · CNN · SVM · Transfer learning · Histopathological images

## 1 Introduction

Throat cancer broadly comes under the category of the neck and head cancer, it has common features of oropharyngeal and mouth cancer affecting both children and adults. Throat has many parts and cancer can start in any one of them, for example, Oropharyngeal cancer, Hypopharyngeal, Laryngeal cancer. Majorly throat cancer starts at the mouth and starts spreading to the throat region. Reasons for cancer include tobacco usage, heavy alcohol consumption [1], vitamin deficiency, HPV infection, prolonged exposure to chemicals. Symptoms [2] of throat cancer are sore

R. Kadirappa (✉) · G. Amaranageswarao · S. Deivalakshmi
Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli, India
e-mail: 408119010@nitt.edu

S. Deivalakshmi
e-mail: deiva@nitt.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_25

throat with cough for a long duration, change in voice (hoarseness), lump in the neck, problem in swallowing and/or breathing etc.

A biopsy is one of the potential method used to detect the disease at the early stage in this method a tissue sample is collected and is examined under a microscope to find the cancer cells this is done by fine needle surgery or endoscopy since cancer is a penetrating disease throat cancer, can be classified into four stages [3] in stage 1 and stage 2 cancer cells are confined to a small area in the organ whereas in stage 3 the neighboring organs like lymph nodes get infected in the Stage 4 cancer would have been spread to other parts like neck, head, lungs, and liver. Considering this traversing nature of cancer, there is a serious need to detect cancer at an early stage, that helps in the diagnosis of the diseases and can be controlled from further diffusing to other parts of the body, early detection of throat cancer can increase life expectancy by 5 years [4, 5].

Computational methods can be highly helpful to detect cancerous images at high accuracy and help as an aiding tool for physicians to suggest an accurate diagnosis. In this paper, a deep learning tool, Convolution Neural Network (CNN) is used to classify the images as normal and cancerous with better accuracy, specificity, and sensitivity from the histopathological images collected from patients with throat cancer symptoms. Samples are collected by fine needle biopsy method and by using a microscope, images are extracted with normal, 10-time(10×) and 20-time(20×) magnification respectively.

### *1.1 Extraction of Images*

Histopathological Images are obtained through the fine needle biopsy method and samples are prepared from Hematoxylin and Eosin [6–11] (H&E) stain by using this dye, a region in tissue is stained in a different color. Hematoxylin stain nucleus, nucleic acid, and other cellular components with blue color whereas Eosin stain other parts of tissue with red color. Samples are obtained from patients with symptoms of throat cancer, a total of 120 patients were examined out of which 100 patients were having throat cancer and 20 patients were normal according to ground truth results and these images are used for classification using 7 CNN techniques as discussed in Sect. 3.

### *1.2 Objective*

Objective of the paper is to evaluate performance analysis of different existing CNN architectures by using Histopathological images. Also compare the evaluation metrics of existing CNN architectures.

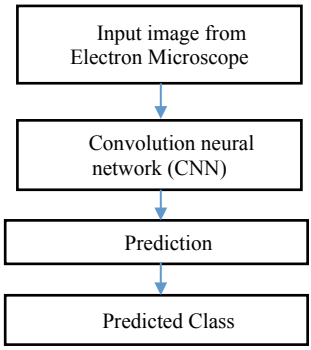



**Fig. 1** Flow chart of the proposed work

## 2 Methodology

Extracted and stained images are used as input, to the Convolution Neural Network (CNN) with transfer learning [12–14] for classification. The flow chart of the proposed work is shown in the Fig. 1. Based on the feature maps obtained from CNN, Support Vector Machine (SVM) developed by Vapnik [15] is mainly used for binary classification.

In this paper SVM is used to classify the image as either cancerous or normal with an objective to classify the data from the database into two separate classes with less error. SVM is uses a single hyperplane given in Eq. 1, x indicates feature maps, w indicates weights with x as input vector and b is bias point. The test datasets are tested and the class is predicted with an average accuracy of 93%, for prediction SVM is used.

$$f(w, x) = w \times x + b \quad \text{where x} \in \mathrm{R}^{\mathrm{m}} \tag{1}$$

## 3 Convolution Neural Network (CNN) for Classifying Images

Convolution neural network is a subset of neural networks, CNN is a powerful method widely used in image segmentation, object recognition, regression, image & video recognition, and Image analysis & classification, recommendation systems, natural language processing, etc. the most general steps involved in CNN for classification is shown in Fig. 2 and discussed in detail.

In Deep learning algorithms input data may be a multidimensional array and kernel also is a multidimensional array of parameters adapted by a learning algorithm, these multidimensional arrays are considered as tensors. In this paper two-dimensional

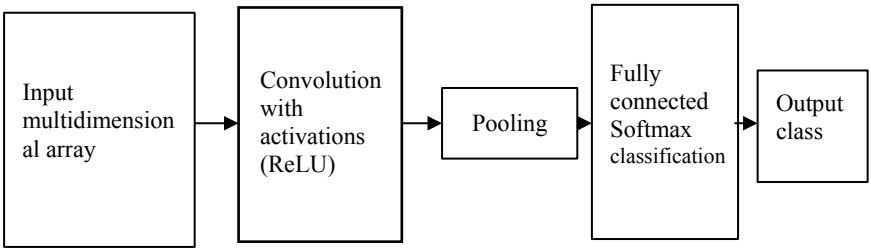



**Fig. 2** Steps involved in CNN architecture

image (I) is considered as input hence two-dimensional kernel (K) is considered to extract learning parameters, the convolution operation is defined in Eq. 2 and sine convolution is commutative it can be equivalently written as Eq. 3.CNN incorporates three steps that aids learning algorithm to improvise its results (a) sparse interaction (b) parameter sharing (c) equivariant representation

$$S(i, j) = (I \times K)(i, j) = \sum_m \sum_n I(m, n)k(\mathrm{i} - m, j - n) \tag{2}$$

$$S(i, j) = (K \times I)(i, j) = \sum_m \sum_n I\,(i - m, j - n)k(\mathrm{i}, j) \tag{3}$$

CNN has three major steps in the first stage convolutions are performed in parallel to produce a set of linear activations in the second stage output from the first stage is used as input to nonlinear activations like rectified linear unit (ReLU) in the third stage pooling function is used to further modify the layer output. The pooling function replaces the net output at a certain location with the statistical summary of the neighborhood, for example, max-pooling is used in this paper where the center value of the filter is replaced with the highest value in the considered rectangular or square filter of fixed dimension. This improves computational efficiency as the next layer has fewer inputs to process.

A fully connected layer, forms the last stage of the convolutional neural network making CNN a more sophisticated one, at this stage CNN can be considered as a traditional neural network, the details captured in the pooling step is flattened to form a vector data at the end of this stage features are ready to be classified with a fair degree of accuracy in the next steps the classification can be done with high precision. After multiple iterations, the particular neuron gives high priority to particular features with the highest weight as the features belong to that particular class and neglecting other features as it belongs to other classes.

To classify the features support vector machine (SVM) and transfer learning (TL) is used and results are discussed in detail in Sect. 4. The final step of CNN involves classifying the images from the dataset into a group of classes with high accuracy in this paper softmax and SVM techniques are used to classify the images. Both are having a small difference in objective function like SVM tries to classify by having

a large margin between data points of different classes where as softmax tries to classify by minimizing cross-entropy or maximizing log-likelihood value [16].

## 4 Results and Discussion

Images obtained from microscope are used for classification initially images are labeled as cancer and normal. These labeled images are used for training, validation, and testing using the CNN techniques a total of 7 techniques were used to extract features and classify the images as either normal or cancerous. Classification is done by CNN combined with support vector machine (SVM) and CNN combined with Transfer Learning (TL) results are tabulated in Table 1.Convolution neural network mainly consists of a convolution layer, rectified linear unit (RELU) layer, pooling layer, fully connected layer, softmax layer and classification layer [17, 18].

Images with fixed size ($224 \times 224 \times 3$) where $224 \times 224$ is the size of image and 3 indicate the channel size, is given as input to the network, as the convolution process the size of the image is reduced in the consecutive layers finally at the fully connected layer the features are flattened for classification into many output classes in our case it is two-class problem. The output is evaluated by the evaluation metrics given below and accuracy is obtained by the classifier, comparing with ground truth results. To classify the features two methods were employed by using a support vector machine and transfer learning discussed in detail.

### *4.1 Performance Evaluation Metrics*

To test the performance of the classifier we used evaluation metrics such as accuracy, sensitivity, specificity, precision, F-Measure, G-Mean given in Eqs. (4)–(9) for the values of confusion matrix given in Fig. 3. With the performance evaluation metrics, a detailed study is performed on Deep learning techniques and results are tabulated in Table 1 and analyzed.

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \tag{4}$$

$$\text{Sensitivity} = TP/(TP + FN) \tag{5}$$

$$\text{Specificity} = TN/(TN + FP) \tag{6}$$

$$\text{Precision} = TP/(TP + FP) \tag{7}$$

$$\text{F} - \text{Measure} = 2 \times (((TP/(TP + FP)) * (TP/(TP + FN)))/((TP/(TP + FP)) + (TP/(TP + FN)))) \tag{8}$$

**Table 1** Results of CNN techniques

| Methodology | Database | Accuracy (%) | Specificity | Sensitivity | Precision | DSC | G-mean |
|---|---|---|---|---|---|---|---|
| Alexnet with SVM | ORG | 88.89 | 0.67 | 0.90 | 0.96 | 0.9375 | 0.7784 |
| | 10× | 89.19 | 0.67 | 0.93 | 0.93 | 0.9354 | 0.7897 |
| | 20× | 91.67 | 0.80 | 0.93 | 0.96 | 0.9508 | 0.8650 |
| Alexnet with TL | ORG | 88.89 | 1.00 | 0.88 | 1.00 | 0.9393 | 0.9411 |
| | 10× | 89.19 | 1.00 | 0.88 | 1.00 | 0.9393 | 0.9411 |
| | 20× | 91.67 | 1.00 | 0.93 | 1.00 | 0.9677 | 0.9682 |
| InceptionV3 with SVM | ORG | 94.44 | 0.80 | 0.96 | 0.96 | 0.9677 | 0.8798 |
| | 10× | 97.30 | 1.00 | 0.96 | 1.00 | 0.9841 | 0.9842 |
| | 20× | 97.22 | 1.00 | 0.96 | 1.00 | 0.9836 | 0.9837 |
| InceptionV3 with TL | ORG | 91.67 | 1.00 | 0.91 | 1.00 | 0.9538 | 0.9548 |
| | 10× | 89.19 | 1.00 | 0.88 | 1.00 | 0.9393 | 0.9411 |
| | 20× | 86.11 | 1.00 | 0.85 | 1.00 | 0.9230 | 0.9258 |
| Googlenet with SVM | ORG | 86.11 | 0.50 | 0.90 | 0.93 | 0.9206 | 0.6731 |
| | 10× | 91.89 | 1.00 | 0.91 | 1.00 | 0.9538 | 0.9548 |
| | 20× | 94.44 | 1.00 | 0.93 | 1.00 | 0.9677 | 0.9682 |
| Googlenet with TL | ORG | 94.44 | 1.00 | 0.93 | 1.00 | 0.9687 | 0.9692 |
| | 10× | 94.59 | 0.83 | 0.96 | 0.96 | 0.9677 | 0.8980 |
| | 20× | 88.89 | 1.00 | 0.88 | 1.00 | 0.9375 | 0.9393 |
| Resnet50 with SVM | ORG | 96.00 | 1.00 | 0.95 | 1.00 | 0.9767 | 0.9770 |
| | 10× | 96.00 | 0.80 | 1.00 | 0.95 | 0.9756 | 0.8944 |
| | 20× | 96.00 | 0.80 | 1.00 | 0.95 | 0.9756 | 0.8944 |
| Resnet50 with TL | ORG | 95.83 | 1.00 | 0.95 | 1.00 | 0.9756 | 0.9759 |
| | 10× | 95.83 | 0.80 | 1.00 | 0.95 | 0.9743 | 0.8944 |
| | 20× | 95.83 | 1.00 | 0.95 | 1.00 | 0.9756 | 0.9759 |
| Resnet101 with SVM | ORG | 92.00 | 0.75 | 0.95 | 0.95 | 0.9523 | 0.8451 |
| | 10× | 96.00 | 1.00 | 0.95 | 1.00 | 0.9767 | 0.9770 |
| | 20× | 96.00 | 0.80 | 1.00 | 0.95 | 0.9756 | 0.8944 |
| Resnet101 with TL | ORG | 88.89 | 0.67 | 0.90 | 0.96 | 0.9375 | 0.7784 |
| | 10× | 97.30 | 1.00 | 0.96 | 1.00 | 0.9841 | 0.9842 |
| | 20× | 97.22 | 0.85 | 1.00 | 0.96 | 0.9830 | 0.9258 |
| VGG16 with SVM | ORG | 97.22 | 1.00 | 0.96 | 1.00 | 0.9841 | 0.9842 |
| | 10× | 91.89 | 1.00 | 0.91 | 1.00 | 0.9538 | 0.9548 |
| | 20× | 91.67 | 0.71 | 0.96 | 0.93 | 0.9491 | 0.8304 |
| VGG16 with TL | ORG | 94.44 | 1.00 | 0.93 | 1.00 | 0.9687 | 0.9692 |
| | 10× | 97.30 | 1.00 | 0.96 | 1.00 | 0.9841 | 0.9842 |

(continued)

**Table 1** (continued)

| Methodology | Database | Accuracy (%) | Specificity | Sensitivity | Precision | DSC | G-mean |
|---|---|---|---|---|---|---|---|
| | 20× | 94.44 | 1.00 | 0.93 | 1.00 | 0.9677 | 0.9682 |
| VGG19 with SVM | ORG | 92.00 | 1.00 | 0.91 | 1.00 | 0.9545 | 0.9555 |
| | 10× | 92.00 | 0.75 | 0.95 | 0.95 | 0.9523 | 0.8451 |
| | 20× | 96.00 | 1.00 | 0.95 | 1.00 | 0.9767 | 0.9770 |
| VGG19 with TL | ORG | 91.67 | 1.00 | 0.90 | 1.00 | 0.9523 | 0.9534 |
| | 10× | 95.83 | 0.80 | 1.00 | 1.00 | 0.97561 | 0.9759 |
| | 20× | 95.83 | 0.80 | 1.00 | 0.95 | 0.9744 | 0.8944 |

| | Predicted class Positive | Predicted class Negative |
|---|---|---|
| Actual class Positive | True positive (TP) | False-negative (FN) |
| Actual class Negative | False-positive (FP) | True negative (TN) |

**Fig. 3** Confusion matrix for binary classification

$$\text{G - Mean} = \sqrt{(TP/(TP + FN) * TN/(TN + FP))} \tag{9}$$

**Accuracy.** CNN models accuracy obtained lies between 86.11 to 97.3 an average accuracy of deep learning techniques with TL and SVM are 93.52% and 93.098 respectively hence deep learning techniques with TL has slightly outperformed SVM by 0.423 average percentage for the augmented dataset. As discussed in the previous section data is extracted from samples with three magnitudes viz., original, 10×, 20× and their average accuracy with TL are 92.26, 94.18, 92.85 respectively and with SVM are 92.38, 93.46, 94.71 respectively.

**Sensitivity.** Indicates (true positives predicted/all true positives) with this metric we calculate the portion of patients that are correctly identified as cancerous to total patients with cancer results are depicted in Table 1 and for quick reference, the sensitivity of the model has varied from 1 to 0.85 whereas the average sensitivity of the deep learning techniques with TL varies from 1 to 0.85 and SVM varies from 1 to 0.90. Average sensitivity with TL of original, 10× and 20× data are 0.92, 0.94, 0.95 respectively, average sensitivity with SVM of original (ORG), 10× and 20× data are 0.94, 0.95, 0.97 respectively with sensitivity metric deep learning with SVM has outperformed TL.

**Specificity.** Indicates the class left by sensitivity it explains how many samples are correctly identified as normal cases comparing with the total normal cases in the data results are tabulated in Table 1. The specificity has varied from 1 to 0.5 in the case of SVM and from 1 to 0.67 with TL due to data skewness the specificity has observed very low with average specificity with TL, SVM being 0.94, and 0.86 respectively. Specificity has been better observed when data is magnified and the TL technique

is used. Average specificity for original, 10× and 20× with TL are 0.95, 0.92, 0.95 and with SVM it has been observed as 0.82, 0.89 and 0.87 respectively.

**Precision.** Gives the measure on out of positively predicted images how many are positive. Precision should be as high as possible, in the models used precision has varied from 1 to 0.93 as tabulated in Table 1, for the dataset average precision in the case of SVM and TL are 0.97 and 0.99 respectively. Considering the average precision for the dataset using SVM is 0.97, 0.98, 0.97 for original 10× and 20× respectively and using TL we have obtained 0.99, 0.99, 0.99 for the original 10× and 20× data overall precision is obtained with high values.

**F – measure.** Is the harmonic mean, in turn attacking the outliers for the dataset. f –measure is tabulated in Table 1, metric varies from 0.98 to 0.92. Using f – measure we can calculate precision and sensitivity(recall), the average f – measure for both SVM and TL is 0.96 but it has varied for the dataset with values for original, 10×, and 20× using TL are 0.96, 0.97, 0.96 respectively. SVM has average values of 0.96, 0.96, 0.97 for original, 10× and 20× respectively.

**G-mean.** Geometric mean is used to get the typical value; it has varied from 0.98 to 0.67 values are tabulated in Table 1. The average g-mean value using TL is 0.94 compared to 0.90 of SVM, for original, 10×, and 20× magnitude using SVM we have obtained 0.87, 0.91, 0.92 respectively whereas using TL for the original, 10× and 20× magnitudes we have obtained 0.93, 0.95, 0.94.

## 5 Conclusion and Future Scope

Throat samples of different magnitude (original, 10×, 20×) were used in the study, deep neural convolutional neural networks with transfer learning are used to train and test the samples in one case, and in another case, a support vector machine was used to classify the images as normal or cancer condition. The results obtained were promising with an average accuracy of 93%, performance evaluation metrics are tabulated in Table 1. Evaluation metrics have shown better results when the 10× magnitude images are used rather than original images. The transfer learning technique has been slightly preferred compared to SVM as the results are more promising with an average learning time of 98 min 16 s.

### *5.1 Future Scope*

In current paper only existing architectures are used to perform comparative analysis and imbalanced dataset is used for analysis. In future data augmentation can be used to handle imbalanced data and novel architectures can be implemented and results can be compared with existing architectures.

## References


1. Williamson JS, Biggs TC, Ingrams D (2012) Laryngeal cancer: an overview. Trends Urol Men's Health 3. https://doi.org/10.1002/tre.295
2. Jones TM, De M, Foran B, Harrington K, Mortimore S (2016) Laryngeal cancer: United Kingdom national multidisciplinary guidelines. J Laryngol Otol 130:S75–S82
3. Patel SG, Shah JP (2005) TNM staging of cancers of the head and neck: striving for uniformity among diversity. CA Cancer J Clin 55(4):242–58. quiz 261–2, 264. PMID: 16020425. https://doi.org/10.3322/canjclin.55.4.242
4. Lee L-A, Cheng A-J, Fang T-J, Huang C-G, Liao C-T, Chang JT-C, Li H-Y (2008) High incidence of malignant transformation of laryngeal papilloma in Taiwan. Laryngoscope 118(1):50–55
5. Chu EA, Kim YJ (2008) Laryngeal cancer: diagnosis and preoperative work-up. Otolaryngol Clin North Am 41(4):673–695
6. Bhattacharjee K, Girish HC, Murgod S, Alshame AMJ, Dinesh BS (2018) Connective tissue stains: a review article. Dusunen Adam 9: 809–818
7. Alturkistani HA, Tashkandi FM, Mohammedsaleh ZM (2015) Histological stains: a literature review and case study. Glob J Health Sci 8. https://doi.org/10.5539/gjhs.v8n3p72
8. Gurcan MN et al (2009) Histopathological image analysis: a review. IEEE Rev Biomed Eng 2:147–171
9. Bancroft JD, Gamble M (2008) Theory and Practice of Histological Techniques. Elsevier Health Sci, New York
10. Patil RC, Mahesh PK (2015) Study of hematoxylin and eosin dye and its alternatives in oral cancer detection. In: Proceedings of the International Conference on Current Trends in Engineering, Science and Technology, ICCTEST
11. Rao RS, Patil S, Majumdar B, Oswal RG (2015) Comparison of special stains for keratin with routine hematoxylin and eosin stain. J Int Oral Health JIOH 7(3):1–5
12. Ribani R, Marengoni M (2019) A survey of transfer learning for convolutional neural networks. In: 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), Rio de Janeiro, Brazil, pp 47–57. https://doi.org/10.1109/SIBGRAPI-T.2019.00010
13. Pan SJ, Yang Q (2009) A survey on transfer learning IEEE transactions on knowledge and data engineering. https://doi.org/10.1109/TKDE.2009.191
14. Weiss K, Khoshgoftaar TM, Wang D (2016) A survey of transfer learning. J Big Data 3:9. Weiss et al. https://doi.org/10.1186/s40537-016-0043-6
15. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
16. Tang Y (2013) Deep learning using linear support vector machines. arXiv preprint arXiv:1306.0239
17. Agarap AF (2019) An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification. arXiv:1712.03541v2
18. Kim S, Yu Z, Kil R, Lee M (2015) Deep learning of support vector machines with class probability output networks. Neural Netw 64:19–28

# Deep Learning Based Pneumonia Infection Classification in Chest X-ray Images Using Convolutional Neural Network Model


**Jyoti Nayak and Devbrat Sahu**

**Abstract** Pneumonia infection is a life threating entity which is caused by bacteria, viruses and fungi inflames the lungs making it difficult to breath. To recognize the infection, various stages are employed which includes lung examination through X-ray images. However, the structure of non-uniform infection represents vague details and may lead to poor detection causing improper treatment. Various image processing techniques had been applied in this domain but their performance were restricted due to the varying feature representation of lung infection. In this direction, a convolutional neural network model has been developed to classify between normal and pneumonia infected X-ray lung images. The performance of the model has been thoroughly examined and validated through the standard performance measures. The model achieved the classification accuracy of 99.06%, F1-score of 99.01% and sensitivity of 99.29%. These standard scores describe the model is potentially strong and outperforms with the state-of-the-art methods in this domain.

**Keywords** Computer-added diagnosis · Convolutional neural network · Deep learning · Lung X-ray images · Medical imaging · Pneumonia infection

## 1 Introduction

Pneumonia is a type of infection that majorly affects one or both lungs. This infection causes air sacs to distress the lungs with pus or fluid which further transformed into cough, fever and sometimes difficulty in breathing. A variety of virus, bacteria, and fungi are responsible to produce the pneumonia which can weaken the immune system and sometimes known for death. It may include person from every age group, but responsible for approximately 15% of total deaths of the children who aged below 5 years, and in 2017 it killed 8,08,694 children [1]. The symptom of pneumonia is closely related to the COVID symptoms and thus cannot be avoided which is causing panic situation in the whole world. However, early diagnosis and relevant

J. Nayak (✉) · D. Sahu
Department of Computer Science and Engineering, Raipur Institute of Technology, Mandir Hasaud, Chhatouna, Chhattisgarh 492001, India
e-mail: jnayak200@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_26

treatment of pneumonia can be helpful to save the life. In case of bacterial pneumonia antibiotics prescribed by the doctor can heal a person in few days, but in case of viral pneumonia it needed close examination of chest X-ray and treatment may last for many weeks. Any inaccurate observation even in the early stages of diagnosis may lead the person suffers from high risk which can be a life threatening experience to the person. Therefore, the procedure to examine the severity of infection can be assisted using latest intelligent computing technique to make the prediction of infection accurate which surely minimizes the chances of human death. Chest X-ray analysis is a kind of image analysis which identifies and locate the infection made by expert. Apart from the medical X-ray image understanding, numerous image processing and optimization practices have been engaged to comprehend the diverse data patterns in many application domains [2, 3]. However, the reported performance has been investigated under the standard conditions and depends upon application area. Continuing with the advancement in these fields, recent growth in machine and deep learning practices have leveraged the performance of the Artificial Intelligence (AI) based [4–15] AI techniques itself has the potential to provide advantage to some of the most challenging medical problems including pneumonia detection in the early stages with to facilitate the medical field and for the betterment of our society. Among the recent AI based methods, deep learning approaches mostly learn significant features to build a robust, more accessible and soft operative solutions which can also be employed in pneumonia detection and classification problem in chest X-ray images. In this paper, a deep learning based Convolutional Neural Network (CNN) model has been developed which not only precisely categorize between normal and pneumonia uninfected images but also leads to the applications of similar domain. The proposed approach is illustrated in the Fig. 1.

The organization of the rest of this paper is systematized as follows: Sect. 2 reflects the relevant studies in the existing domain while Sect. 3 express the implementation details and the experimental outcomes; finally, Sect. 4 reveals the performance of proposed work and Sect. 5 describes the final observations and consists the insights for future directions.

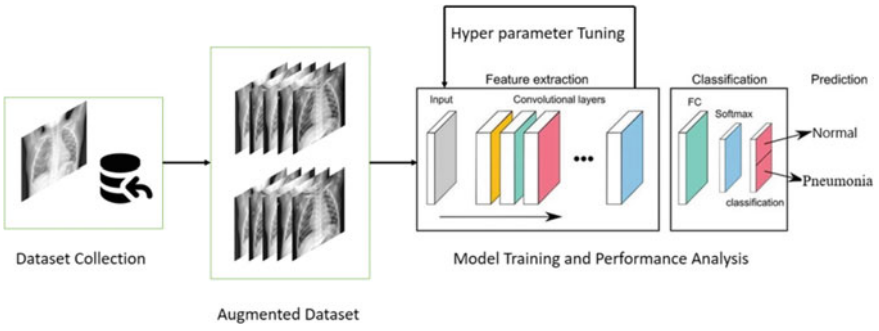



**Fig. 1** Process flow of the proposed convolutional neural network based model to categorize the X-ray lung images into normal and pneumonia infected classes

## 2 Related Works

Gupta et al. utilized second order gray-level co-occurrence matrix (GLCM) to examine the important features in pneumonia classification [16]. Total of 14 features were obtained and then normalized and outliers were removed which assisted to obtain the accuracy of 89.5%. Variation in the pixel intensities, the position and the change in the target orientation may degrade the detection performance. To deal with this issues, Ambita et al. had combined the support vector machine (SVM) with local adaptive regression kernel (LARK) which resulted a 98% score for both precision and recall metrics [17]. Based on the examination of non-rigid deformable features, Chandra et al. obtained the significant features and segmented lungs regions to detect pneumonia [18]. The approach was examined under multilayer perceptron, logistic regression technique, random forest, sequential minimal optimization and classification through regression technique. Among the mentioned approaches, highest accuracy of 95.63% was obtained using logistic regression classifier. Khatri et al. had applied earth mover's distance (EMD) technique to find out the pixel variation in X-ray images to describe the pixel difference between pneumonia and non-pneumonia cases [19]. In another work of significant feature examination which represents the accurate detection of pneumonia cases, Yu et al. had employed the graph embedded feature reconstruction and developed a CNN model called CGNet [20]. The exact identification of lesion region is advantageous because it can offer a significant information for the treatment selection. In this direction, Ni et al. applied a deep learning framework to obtain the specific set of features and processed the quantitative lesion in context of COVID-19 pneumonia [21]. But, being a retrospective study, the working behavior of the model on actual data had not been validated as mentioned. Similarly, the performance is suffered from the data imbalance problem.

Chouhan et al. employed a transfer learning based deep learning model to extract significant features from the lung X-ray images[22]. Features were intended to be extracted using neural network model trained on ImageNet followed by the classifier. The obtained result in pneumonia recognition shows the performance in terms of accuracy of 96.4% and recall of 99.62%. Similarly, Eid et al. combined ResNets, to obtain the hierarchical features from X-ray images and boosting system to choose the salient features, followed by the support vector machine (SVM) for classification [23]. Likewise, Rahman et al. had applied four different pre trained deep CNN model namely SqueezeNet, ResNet18, AlexNet, and DenseNet201 [24]. Out of which, DenseNet reveals classification accuracy of 95% and 96% sensitivity/recall score. With the weighted classifier, Hashmi combined the Xception, DenseNet121, ResNet18, MobileNetV3, and InceptionV3 models, performed partial data augmentation and attained the pneumonia detection accuracy of 98.43% with 98.26% of precision score [25]. With 5 different custom CNN model, Fathurahman experimented the functions of 1D convolutions and achieved 94% classification accuracy. However, the design of proposed work had not been discussed [26].

The mentioned approached have either been assisted using traditional image processing or machine learning methods. Some of them have utilized the deep

learning based transfer learning model. Though, CNN based mechanism have also been used but due to hidden or complex architecture, the actual performance of the models has not been revealed. Valuable features are the one which may help a system to recognize the pneumonia in X-ray images. But, feature extraction in the mentioned work has not been potentially brought out. Similarly, the design issues of deep learning based model have not been discussed in context of their training and testing behavior. In this paper, a CNN based model to classify lung X-ray images into normal and pneumonia category has been proposed which offers less complex design and validated through the qualitative and quantitative assessment scores.

## 3 Materials and Methods

In order to perform classification against the normal and pneumonia infected lung X-ray images, following phases have been considered whose process flow are given in the Fig. 1.

### *3.1 Dataset Collection*

For the proposed working towards the classification using CNN model, a high number of data samples are required to learn reliable and significant features. For this purpose, one such public image dataset "Chest X-Ray Images (Pneumonia)" has been referred [27]. This dataset contains 5856 images distributed in train, test and validation sets of total 1.15 GB in size. All the image samples are already encoded in the gray scale representation. The sample grid of the referred dataset has been shown in the Fig. 2.

### *3.2 Dataset Augmentation*

This phase is required to increase the size of data samples thus making the model to learn more but valuable features. Techniques of image processing including rotation, horizontal flipping, vertical flipping, noise, and shearing operations have been applied as a data augmentation process and combined with the original sample thus resulted a total of 33,136 images. This huge number of data samples are sufficient for the stated purpose and will be inputted to the developed model.

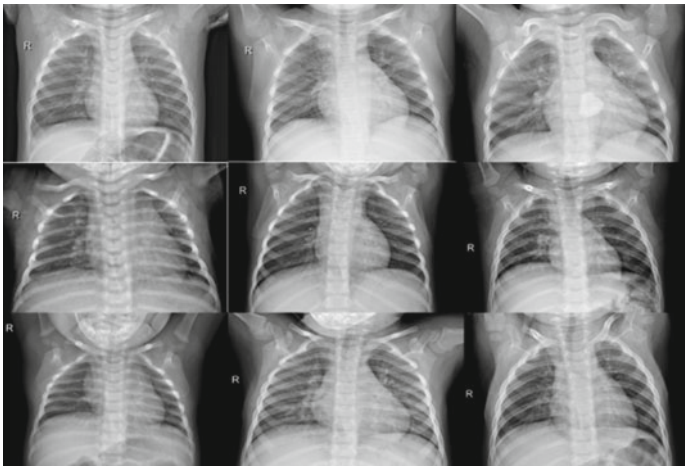


**Fig. 2** Sample grid of referred dataset containing Chest X-ray images

### *3.3 Model Design and Training*

The proposed model which is intended to extract more number of significant and robust features has been equipped with convolution process. As the dataset has more number of inputs, a total 10 layers of convolutional process in which a pair of 16, 32, 64, 128 and 256 kernels have been applied for the defined image size of 180 * 180 in all the three channels. Further, to get the desired features from the images and to normalize the reliable pixel information, max pooling strategy has been applied. This strategy reduces the computation complexity and focuses on the high level feature extraction. But, they are unable to process for each batch. Keeping the facts in the design of such model, regularization techniques in the form of batch normalization technique has been used which standardizes the inputs to a layer for each batch of inputs. Additionally, it has the stabilizing effect to the learning and assists to minimize the number of epochs to train the network. Further, model may have additional but trivial number of features extracted during learning process. To keep them out, dropout strategy has been employed with a dropout rate of 20% in the initial pair and then adopted 70%, 50% and 30% respectively in the next three blocks. Rectified linear unit (relu) is employed in all the relevant layered approach as an activation function except the last layer of dense layer where softmax function has been applied to represent the class probability over the two defined classes of normal and pneumonia affected image. The padding has been kept on a single shift while the kernel which is applied to obtain the features is set to 3 * 3 throughout the learning process. Apart from this, we have selected adam optimizer which has been known to get optimal performance in the network training. The model architecture has been illustrated in Fig. 3 for which description has been provided. The training loss which occurs during the learning process has also been considered and targeted to have minimum loss

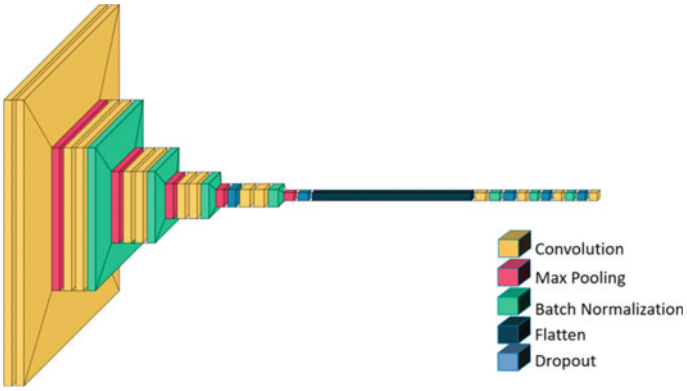



**Fig. 3** Proposed convolutional neural network model to obtain the significant features during its training phase. Various hyper parameter has also been regulated to tune up the performance of the model during the training

especially in the case of imbalance data. It has to be done, otherwise the developed model will be biased towards the dominating class. The considered loss function *(L)* with the weighted loss function is described below.

$$L_{CrossEntropy}(x_i) = -(y_i \log(f(x_i)) + (1 - y_i) \log(1 - f(x_i))) \tag{1}$$

For the sample input x and output y, the Eq. (1) has been described. For the average cross entropy loss using training set T whose size is N with its positive ($p$) and negative ($n$) example set can be defines as:

$$L_{CrossEntropy}(T) = -\frac{1}{N}(\sum_{pExample} \log(f(x_i)) + \sum_{nExample} \log(1 - f(x_i)) \tag{2}$$

For the imbalance data samples, the mentioned loss defined in Eq. 2 has been further modified using weighted loss entity and defined in Eq. (3)

$$L^{W}_{CrossEntropy}(x) = -W_p(ylog(f(x) + W_n(1 - y) \log(1 - f(x))) \tag{3}$$

Model training is influenced from its design strategy, tuning the hyper-parameters, and selection of optimizers. The overall training presentation is characterized through two important factors which are the accuracy and loss curve over 2,862,353 parameters in the mentioned architecture. However, in the model training other hyper parameters such as learning rate, choice of optimizers among seven different optimizers, selection of batch size parameter, and decision on selecting the number of epochs have been exhaustively inspected and applied to confirm the reliable feature extraction while making the network learning stability in terms of training and loss.

# 4 Discussion

## *4.1 Performance Evaluation*

Considering the Fig. 4(a) which describes the model performance using the learning rate of 0.1, RMSProp optimizer and batch size of 8 which does not ensure the learning of significant feature even with the maximum learning accuracy. Similarly, the loss in Fig. 4(b) is not stable and showing the instability during learning. It is simply indicating that the tuning hyper parameter is a bit complex task and required close observation even if the model achieves higher accuracy. The sample output from the proposed trained network which is given in Fig. 4 is one of the worst output which was examined during training. However, with the improvement of the model n the Fig. 5(a-b) where the training behavior of the proposed model is stable and considered as its final performance. In this, the learning rate of 0.001, optimizer is adam and batch size 32 has been selected for the final observation. The stability of the learning network, minimum difference between training and validation accuracy

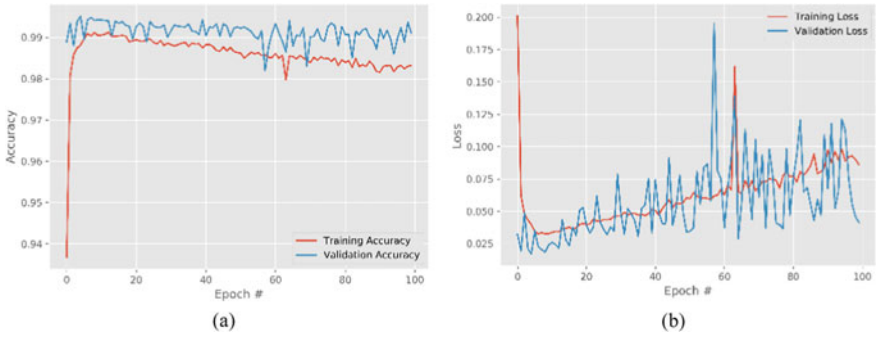



**Fig. 4** Model training using learning rate of 0.1, batch size 8 and RMSProp optimizer **a** Accuracy **b** Loss

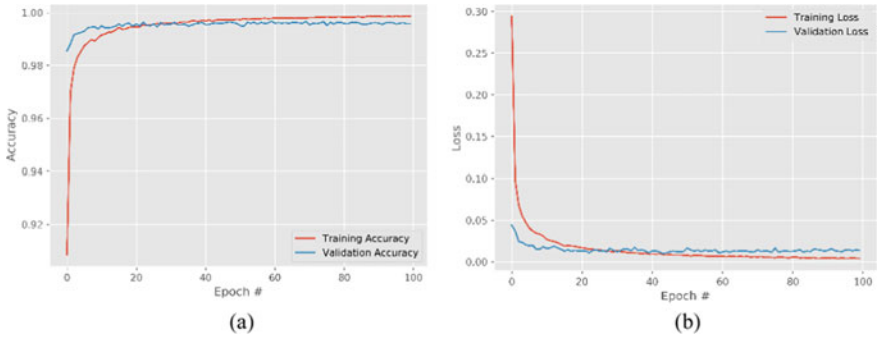



**Fig. 5** Fine-tuned final model whose performance has been observed stable **a** Accuracy **b** Loss

**Table 1** Quantitative classification performance of proposed CNN model and its comparison with the relevant state-of-the-art technique

| Method | Accuracy | F1 | Precision | Sensitivity |
|---|---|---|---|---|
| Gupta et al. [16] | 89.5 | 91.8 | 89.7 | 96.1 |
| Ambita et al. [17] | – | – | 98.00 | 98.00 |
| Chandra et al. [18] | 95.63 | 95.45 | 97.47 | 93.63 |
| Yu et al. [20] | 99.01 | 98.97 | 100.00 | 98.98 |
| Chouhan et al .[22] | 96.40 | – | 93.28 | 99.62 |
| Eid et al. [23] | 98.13 | 98.16 | – | – |
| Rahman et al. [24] | 95.00 | 93.82 | 95.00 | 96.00 |
| Hashmi et al. [25] | 98.43 | – | 98.26 | 99.00 |
| Fathurahman et al. [26] | 94.00 | – | – | – |
| Proposed model | 99.06 | 99.01 | 98.87 | 99.29 |

and loss graph approves the behavior of the model is optimum and considered as the reliable and final model to classify between normal and pneumonia affected lung images.

## *4.2 Comparison with the State-Of-The-Art Method*

To validate the performance achieved using qualitative measurement as represented in the Fig. 5, quantitative evaluation has also been made which includes the analysis of standard evaluation measures such as accuracy, precision, sensitivity, and f1 score. F1 score is significant because it designates the probability score in case of imbalanced dataset. With all attained scores for mentioned metrics, performance of the proposed model has been compared with other state-of-the-art techniques and given in Table 1. It specifies that the proposed method outperforms other allied studies in the domain. Further, the test results on the classification using the final model has been represented in Fig. 6 where, the red color text displays the wrong prediction where the blue color text indicates the true prediction. The test samples have been taken randomly and does not represent repetition.

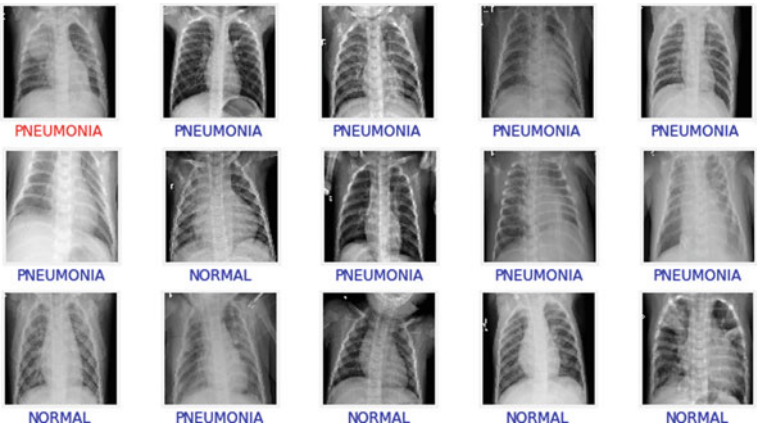



**Fig. 6** Test performance of the proposed CNN model on the referred dataset

## 5 Conclusion

The developed CNN based model categorizes the X-ray of lungs into normal and pneumonia infected images with the accuracy of 99.06%. To make the performance of this model robust, it is also examined and validated through F1-score which is an assessment parameter majorly used in unbalanced dataset. Quantitatively, model attained F1-score of 99.01%, and which is an optimal score as compared with other standard techniques in the presented domain. However, the performance of the model can be further enhanced and tested with different optimizers and unseen data. In context of future directions, this model may be stretched to categorize the different causes of pneumonia infection such as other virus, bacteria and fungi.

## References

1. WHO (2021) World Health Organization, Reports on Pneumonia. https://www.who.int/news-room/fact-sheets/detail/pneumonia.
2. Dewangan DK, Sahu SP (2020) Real time object tracking for intelligent vehicle. In: 2020 1st international conference on power, control computer technology, ICPC2T 2020, pp 134–138. https://doi.org/10.1109/ICPC2T48082.2020.9071478
3. Chaudhuri A, Sahu TP (2021) Feature selection using Binary Crow Search Algorithm with time varying flight length. Expert Syst Appl 168:114288. https://doi.org/10.1016/j.eswa.2020.114288
4. Sahu SP, Dewangan DK (2021) traffic light cycle control using deep reinforcement technique. In: International conference on artificial intelligence and smart systems (ICAIS), 2021, pp 697–702. https://doi.org/10.1109/ICAIS50930.2021.9395880

5. Dewangan DK, Sahu SP (2021) Road detection using semantic segmentation-based convolutional neural network for intelligent vehicle system. In: Data engineering and communication technology. Lecture notes on data engineering and communications technologies. Springer, Singapore, pp 629–637
6. Dewangan DK, Sahu SP (2021) Deep learning-based speed bump detection model for intelligent vehicle system using raspberry pi. IEEE Sens J 21(3):3570–3578. https://doi.org/10.1109/JSEN.2020.3027097
7. Dewangan DK, Sahu SP (2021) RCNet: road classification convolutional neural networks for intelligent vehicle system. Intell Serv Robot 14(2):199–214. https://doi.org/10.1007/s11370-020-00343-6
8. Chaudhuri A, Sahu TP (2020) PROMETHEE-based hybrid feature selection technique for high-dimensional biomedical data: application to Parkinson's disease classification. Electron Lett 56(25):1403–1406. https://doi.org/10.1049/el.2020.2517
9. Dewangan DK, Sahu SP (2021) Driving behaviour analysis of intelligent vehicle system for lane detection using vision-sensor. IEEE Sens J 21(5):6367–6375. https://doi.org/10.1109/JSEN.2020.3037340
10. Dewangan DK, Sahu SP (2021) PotNet: pothole detection for autonomous vehicle system using convolutional neural network. Electron Lett 57(2):53–56. https://doi.org/10.1049/ell2.12062
11. Dewangan DK, Sahu SP (2021) Predictive control strategy for driving of intelligent vehicle system against the parking slots. In: 2021 5th international conference on intelligent computing and control systems (ICICCS), pp 10–13. https://doi.org/10.1109/ICICCS51141.2021.9432362
12. Ojha A, Sahu SP, Dewangan DK (2021) Vehicle detection through instance segmentation using mask R-CNN for intelligent vehicle system. In: 2021 5th international conference on intelligent computing and control systems (ICICCS), pp 954–959. https://doi.org/10.1109/ICICCS51141.2021.9432374
13. Banjarey K, Sahu SP, Dewangan DK (2021) A survey on human activity recognition using sensors and deep learning methods. In: 2021 5th international conference on computing methodologies and communication (ICCMC), pp 1610–1617. https://doi.org/10.1109/ICCMC51019.2021.9418255
14. Pardhi P, Yadav K, Shrivastav S, Sahu SP, Dewangan DK (2021) vehicle motion prediction for autonomous navigation system using 3 dimensional convolutional neural network. In: 2021 5th international conference on computing methodologies and communication (ICCMC), pp 1322–1329. https://doi.org/10.1109/ICCMC51019.2021.9418449
15. Sahu SP, Dewangan DK, Agrawal A, Sai Priyanka T (2021) Traffic light cycle control using deep reinforcement technique. In: 2021 international conference on artificial intelligence and smart systems (ICAIS), pp 697–702. https://doi.org/10.1109/ICAIS50930.2021.9395880
16. Gupta A, Gupta A, Verma V, Khattar A (2020) Texture feature extraction : impact of variants on performance of machine learning classifiers : study on chest X-Ray – pneumonia, vol 1. Springer, Heidelberg
17. Ambita AAE, Boquio ENV, Naval PC (2020) Locally adaptive regression kernels and support vector machines for the detection of pneumonia in chest x-ray images. Springer, Heidelberg
18. Chandra TB, Verma K (2020) *Pneumonia* detection on chest x-ray using machine learning paradigm. Springer, Singapore
19. Khatri A, Jain R, Vashista H, Mittal N, Ranjan P, Janardhanan R (2020) Pneumonia identification in chest x-ray images using EMD. Springer, Singapore
20. Yu X, Wang S, Zhang Y (2021) CGNet: a graph-knowledge embedded convolutional neural network for detection of pneumonia. Inf Process Manag 58(1):102411
21. Ni Q, et al (2020) A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images, pp 6517–6527
22. Chouhan V, Singh SK, Khamparia A, Gupta D, De Albuquerque VHC (2020) A novel transfer learning based approach for pneumonia detection in chest x-ray images. Appl Sci 10(2):559
23. Eid MM, Elawady YH (2021) Efficient pneumonia detection for chest radiography using resnet-based SVM. Eur J Electr Eng Comput Sci 5(1):1–8

24. Rahman T, Chowdhury MEH, Khandakar A (2020) Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. Appl Sci 10(9):3233
25. Hashmi MF (2020) Efficient pneumonia detection in chest xray images using deep transfer learning. Diagnostics 10(6):417. https://doi.org/10.3390/diagnostics10060417
26. Fathurahman M, Fauzi SC, Haryanti SC (2020) Implementation of 1D-convolution neural network for pneumonia classification based chest X-ray image. Springer, Heidelberg
27. Mooney P (2018) Chest X-Ray Images (Pneumonia)-Version 2. https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia.

# Deep Learning Model for Reduction COVID-19 Spreading Through Tracking Students' Commitment to Wearing a Face Mask


**Ramy Said Agieb**

**Abstract** The world lives in these days a state of panic and fear that it has not experienced before, due to the spread of the COVID-19 which required specific measures. Among these measures include wearing face masks for protection, and limiting the spread of infection among individuals. Therefore, work has been done in this research to establish a system that works to continuously follow up students in studying halls to ensure that the face mask is worn and issued a warning for those who take off the face mask or wear it in the wrong way. This is done automatically without the need for human intervention. The core of this research is to classify between people with mask and without the mask. The classification is based on using Deep Learning (DL) methods. The model of classifications produces by training the dataset using Convolution Neural Network (CNN), TensorFlow (TF), and OpenCV libraries.

**Keywords** Convolution Neural Network · COVID-19 · Deep Learning · Face Mask · OpenCV · TensorFlow

## 1 Introduction

The world is facing a challenge that has never been faced before, which is the emergence of the COVID-19 epidemic. The impact of this epidemic extended to all different aspects of life and led to isolating people all over the world in an attempt to reduce the death rate of this epidemic, which led to the emergence of the term social distancing [1]. Despite the emergence of more than one vaccine that can protect humans at certain rates, the need to use the face mask as the best protection remains to reduce infection rates and increase protection from this epidemic [2]. Among the places most exposed to injury are student gatherings, specifically inside the lecture halls, and therefore, the extent of students'commitment to using the face mask must be monitored and take a decision with any mistake [3].

R. S. Agieb (✉)
Faculty of Engineering, Department of Electrical Engineering, MTI University, Cairo, Egypt
e-mail: drragieb@eng.mti.edu.eg; ramysaidagieb@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_27

In this research, we try to employ the progress and development achieved in the field of DL to achieve follow-up and observe the extent of students' commitment in the lecture halls to the use of masks. Artificial intelligence is seen as the umbrella that drives the transformation of computers to think and act like a human being, and it has developed through the field of machine learning in which accurate learning has also developed. The term may comparatively be applied to any machine that shows attributes related to a human brain, for example, learning and essential thinking. ML enables systems to normally take in and improve for a reality without being unequivocally redone. ML bases on the improvement of PC programs that will get information and use it to look out it in seclusion. DL copies the operations of the human cerebrum in preparing information for use in distinguishing objects, perceiving discourse, interpreting dialects, and deciding. Profound learning AI can learn without human management, drawing from information that is both unstructured and unlabeled [4].

In fact, the main difference between humans and computers is the human learn from experiences and computers need to tell what they do (following instructions). Therefore, the starting point is finding the computers learning from its experience's especially from its data and that what machine learning do. ML has three main types of Supervised learning, unsupervised learning, reinforcement learning. ML has a lot of applications like guessing the price of a house, detecting spam e-mail or guesses the value of anything, or classification objects based on previous known and classified data and it's called supervised learning. Customer segmentation in Customer relationship management (CRM), Image compression, Bioinformatics are some applications in unsupervised learning types that search for already undetected examples in an informational collection with no prior marks and with at least human oversight. Reinforcement learning is the third type of ML wherein the clients don't have to manage the model. All things being equal, it permits the model to take a shot at its own to find examples and data that was beforehand undetected. It, for the most part, manages the unlabeled information; some of its applications are game-playing, robots in a maze. DL inspired by the structure of human behave. In terms of DL, these structures are called artificial neural networks. The features of different objects are breakdown by neural networks without human invariant which required more mass volume data, computation power, and training time to train and produce the model because of the features extracted automatically based on the neural network. DL algorithms take large volumes of data as input, analyze the input to extract features out of an object, and identifies similar objects while ML algorithms extract patterns based on labeled sample data. So DL used to Handel unstructured data like image classification and recognition, autonomous driving cars, and robot navigation [5].

DL inspired by the structure of human behave. In terms of DL, these structures are called artificial neural networks. The features of different objects are breakdown by neural networks without human invariant which required more mass volume data, computation power, and training time to train and produce the model because of the features extracted automatically based on the neural network. DL algorithms take large volumes of data as input, analyze the input to extract features out of an object, and identifies similar objects while ML algorithms extract patterns based on labeled sample data. So DL used to Handel unstructured data like image classification and

recognition, autonomous driving cars, and robot navigation. Therefore, to produce a more accurate system the model is produced based on DL [6].

This paper is organized a follows. Section 2 describes the related works. Section 3 describes the method for the proposed system. Section 4 discusses the result. Section 5 presents the conclusion of the research.

## 2 Related Works

DL is the most form of AI nearest to human intelligence and his experience in learning from what he knows. DL is considering as Deep Neural Networks (DNN) which has a large number of hidden layers and each layer has more than one node. DL develops algorithms of DNN for use with the training dataset and produced models to apply in the systems and produced the output. The probability distribution by this model M to the output y conditional distribution of input x as a function of trained parameters Θ is shown in Eq. 1. The output is depending on the activation function f as shown in Eq. 2 where "w" refer to the weight and "b" to the bias. In the backward process use many function to decrease the error between the predicted and expected output. Even if there are many models of Neural Network (NN) but it works with the same steps as following [7]. Table 1 compares between NN models used in DL to select the type used in the proposed system.

a) Calculate the weighted sum of inputs (Forward- propagation).
b) Add the biased of the layer.
c) Determine the activation neuroses in the next layer.
d) The results are fed into the output layer.
e) Back propagation and Gradient Descent
f) Repeated until finding the optimum values of weighted in each layer.

$$P_M = (y/x; \Theta) \tag{1}$$

$$f\left(\sum_{i=1}^{n} x_i \cdot w_i + b\right) \tag{2}$$

**Table 1** NN applications, pros, and cons summarized

| Model | Applications | Pros | Cons |
|---|---|---|---|
| Autoencoder neural network (ANN) [8] | Object dimension reduction | Fit data through multiple filters | Needs more processing time |
| Denoising Autoencoder DAE [9] | Forecasting | The dataset continuously updates | More noise and error in model |
| Restricted Boltzmann Machine RBM [10] | Modeling systems | Create a model based on a sample from the dataset | Precarious to training model well |
| Deep Belief Networks DBN [11] | Human emotion recognition | Hidden layers used in efficient way | More hardware requirements |
| Generative Adversarial Network GAN [12] | Image perfection | Content is visualized in real form | Enormous computerized calculation |
| Self-Organizing Maps SOM [13] | The priority of selection between objects | Information is effectively deciphered and perceived | Slow preparing training model |
| Recurrent neural networks RNN [14] | Music genre recognition | Ready to display long haul arrangement conditions | Needs systems with more memory size |
| Long Short-Term Memory LSTM [15] | Control process | No requirement for boundary calibrating | Required additional hardware units |
| Convolution Neural Network CNN [16] | Image processing | Automatically detects the important features | A huge amount of training dataset |

## 3 Methods

The proposed system depends on using the infrastructure already found in most of the universities and schools as shown in Fig. 1. Streaming videos will be directed to Network Video Recorder (NVR) and then forward to the computerized connections. The computerized system contains a DL model and Decision piece. The model has the ability to switch between different cameras. DL detects any student who takes off the face mask and sends his picture to the decision part to direct it into an alarm

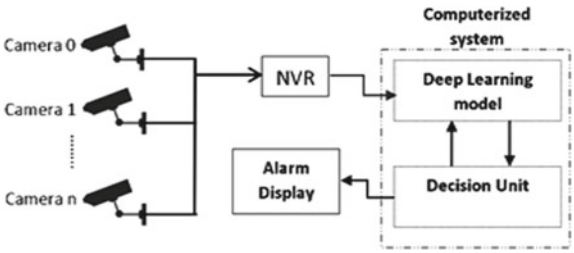



**Fig. 1** The proposed detected and tracking system

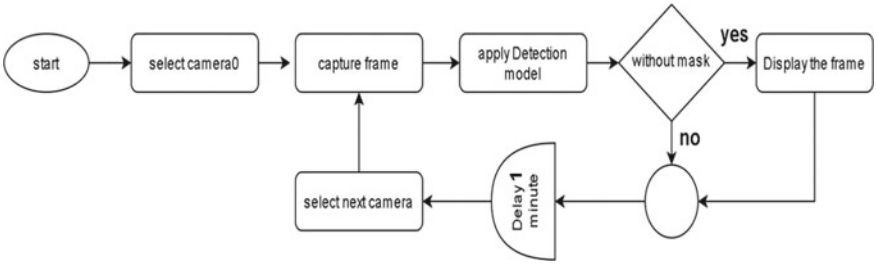



**Fig. 2** Proposed model Flow Chart

display unit like a LED. Figure 2 shows the flow chart of the proposed system used in detecting and tracking students in the case of a mask off.

Although there are plenty of pre trained models available as Alexnet, Resnet, VGG16, and MobileNetV2, I decided to construct a very simple and basic CNN for simplicity to work on any infrastructure and retrained according to students in each institute in the future.

## *3.1 CNN in Depth*

Convolutional Neural Networks (CNN) is neural organizations most normally used to investigate images. A CNN gets a picture as a contribution to the type of a 3D grid. The initial two measurements relate to the width and stature of the picture in pixels while the third one compares to the RGB estimations of every pixel. CNN's consist of the following sequential modules Convolution, ReLu activation function, Pooling, Fully Connected layers, and Output stage (each one can also comprise a couple of layers). Convolution and pooling repeated many times to optimize the network.

**Convolution Layer.** The role of convolution layers is based on representing the elements of the picture's matrix in a simple form (take a batch from the input picture). Convolutional layers take the three-dimensional information grid we referenced previously and they pass away (convolutional kernel) over the image, applying it to a little window of pixels all at once (i.e. $3 \times 3$ pixels) and moving this until the whole picture has been filtered as shown in Fig. 3. The convolutional activity computes the spot result of the pixel esteems in the current step alongside the loads characterized in the channel. The yield of this activity is the last tangled picture. The basis of image classification CNNs is that as the model trains what it truly does is that learns the values used by the filter matrices that empower to separate significant highlights (hued zones, shapes, surfaces, and so forth) in the picture. Every convolutional layer applies one new clear out to the convoluted photo of the previous stage that can extract one more feature. So, as we stack extra filters, the more functions the CNN can extract from a photograph. The output of layer l consists of m feature denoted by Y is given in Eq. 3. Where K represents the filter, B the bias matrix, i

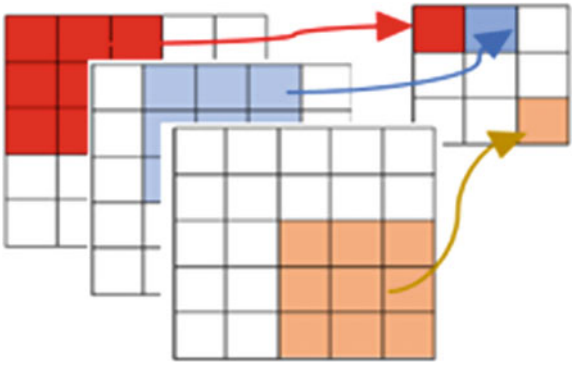


**Fig. 3** Represent image by convolution layer

represent the feature map in layer l, and j is the feature map in layer (l–1).

$$Y_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{m_1^{(l-1)}} K_{i,j}^{(l)} * Y_j^{(l-1)} \tag{3}$$

**ReLu Activation Function Layer.** The next stage following the Convolution stage is the ReLu activation function and is applied after each convolution layer. It introduced non-linearity into the neural network model. It avoids over fitting through generalized the model to accept more inputs. The output of the layer shown Eq. 4

$$Y_i^{(l)} = \max(0, Y_i^{(l-1)}) \tag{4}$$

**Pooling Layer.** Pooling is where the CNN down samples the convolved picture by lessening the number of measurements of the component map. It does so to scale back time interval and therefore the computing power. During this process, it saves the most essential function information. There are a number of techniques that can be used for pooling. The most frequent ones are Average pooling and max pooling. Average Pooling is transferring the normal of the selected block are related to a certain feature in the map. Maximum Pooling (or Max Pooling) is transferring the maximum number in the block area related to a certain feature in the map, was selected in this research.

**Fully Connected Layer.** The next stage after pooling, this stage consists always of more than one layer. The layers take the decision in the classification process. The decision based on the extracted features from the picture according to on training model. The output of the layer when (l–1) is fully connected will be giving by Eq. 6, otherwise will be given by Eq. 5. Where w represents the weight from different nodes.

$$y_i^l = f\left(\sum_{j=1}^{m1^{(l-1)}} w_i^{(l)} y_i^{(l-1)}\right) \tag{5}$$

$$y_i^l = f\left(\sum_{j=1}^{m1^{(l-1)}} \sum_{r=1}^{m2^{(l-1)}} \sum_{s=1}^{m3^{(l-1)}} w_{i,j,r,s}^{(l)} (y_i^{(l-1)})_{r,s}\right) \tag{6}$$

**Output Layer.** The foundation of image classification CNNs is that as the model learns the values used by the filter matrices that enable it to distinguish significant highlights (hued regions, shapes, surfaces, and so on) in the picture.

### *3.2 Tensorflow (TF), Keras, Opencv*

TF and Keras are used to initialize CNN. TF is the magic open-source tool developed by Google, then it used to implement more applications in ML and DL based on the Data Flow Graph (DFG). DFG initialized through two units Nodes and Edges. Nodes describe the mathematical operations while edges represent tensors (multi-dimensional arrays). TF gives the flexibility to make different structures for NN with various activation functions. Keras is a library run on the TF and used for complicated numerical computation with minimal structure. For real-time processing, OpenCV is utilized in this research. OpenCV applied to frames from streaming video and, read images, and to object detection.

### *3.3 Trained the Network*

DL model train used data with labels "mask" and "without a mask" downloaded from [17]. The model was written and executed using Google Collaboratory. DL model selected to work based on CNN in the classification process according to summarized information in Table 1. The first step is to determine the faces in each frame by using Haar Feature-based Cascade Classifiers [18]. Then use CNN to classify the faces with masks or without masks. The first step is to determine the faces in each frame. Start training by visualization the dataset. The dataset is portioned randomly into two parts training data and validation data. The dataset was pertained by 80% to training and 20% to validation. The Sequential CNN model builds with three layers such as Conv2D, MaxPooling2D, Flatten, Dropout and Dense. In the last layer, we use the 'softmax' function to output a vector that gives the probability of each of the two classes. The loss function used is 'adam' optimizer and 'binary_crossentropy'. The number of output filters in the convolution is equal to 32, kernel size equal (3, 3), down samples by (2, 2).

## 4 Results, and Discussion

The results of the proposed system are divided into two parts. The first part related to the training phase while the second is the real testing of the system inside the lecture hall. The classifier model was produced by using CNN. The first process is done on Google collaborate after setting training parameters. Google collaborate provided the hardware resources for training the classifier model. Table 2 shows the parameters used. The face with the mask was labeled by zero where without masks were labeled by one. Figure 4 shows the visualization accuracy and loss values of the proposed model. The values calculated and tested in the training and validation phases. The accuracy of the classifier model for the validation process arrives at 97.4%, can track up to 5 faces without mask. On the other hand, the model wasn't overfitting which improves the accuracy then [19].

The second part of the results appeared during real tests of the system. Figure 5 showed a certain screen shoot frame from a streaming video of one of the cameras in the lecture hall. The proposed system succeeded to detect the face without a mask inside the frame and remark it with a red rectangle border to display it on the alarm screen as shown in Fig. 5.

On the other hand, another test on many groups of different gender which gave results accuracy nearer to results as mention before. When applied the model on a

**Table 2** Training parameters and its values

| Parameter | Values | Parameter | Values |
|---|---|---|---|
| Batch size | 10 | Validation | 20% |
| Epochs | 40 | Filters | 32 |
| Training | 80% | Learning rate | 0.0001 |

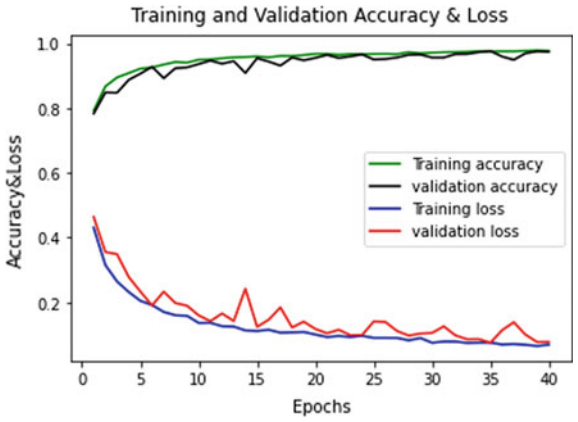



**Fig. 4** Training and validation accuracy & loss


**Fig. 5** Real test inside lecture hall

group of kids and in the case of partial masks the accuracy arrives at 96%. The model was compared to some pre-trained models like VGG16, and MobileNetV2. In the case of VGG16, the accuracy less than the proposed model by 1.5% and take more time. In the case of MobileNetV2, the accuracy is more than the proposed model by 0.2% but takes more time.

The lack in the model's results when the unmasked faces increase according to the characteristics of the used cameras like Resolution, Video Compression, and Frame Rate, which will be improved in the future. Another factor that may be used to increase the accuracy is the used dataset, which if it increased and diversified especially in semi unmask, the accuracy will be increased. The pattern of students' distribution in the lecture hall also affected the accuracy, in the future more experiments will be done to arrive at the optimal distribution. The light of the hall especially on the corners affects the results, so another road for experiments is how to deal with the light distribution in the hall. The processing system may need to upgrade by increasing the processor speed and memory systems.

## 5 Conclusion

In this research, I propose a system to reduce the spread of COVID-19 between the students inside the lecture hall. The system based on the DL model especially CNN. TF, Keras, and OpenCV were used to implement the CNN to train the classifier model which applied to determine the student's face with a mask or without a mask. The training process is done using Google collaborate. The proposed model satisfied accuracy equal to 97.4% in the validation phase. The system tested in real and live environments and succeeded to make his jobs.

## References


1. World Health Organization (2020) Coronavirus Disease (COVID2019) Situation Reports, World Health Organization, Geneva, Switzerland
2. Chen N, Zhou M, Dong X, et al (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. 4e Lancet 395(10223):507–513
3. Wu Z, McGoogan JM (2020) Characteristics of and important lessons from the coronavirus disease 2019 (COVID19) outbreak in China. JAMA 323:1239–1242
4. Agieb R (2020) Machine learning models for the prediction the necessity of resorting to ICU of COVID-19 patients. Int J Adv Trends Comput Sci Eng (IJATCSE) 9(5):6980–6984
5. Phung R (2019) A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. Appl Sci (Switzerland) 9(21):4500
6. Dawani J (2020) Hands-on mathematics for deep learning: build a solid mathematical foundation for training efficient deep neural networks, 1st edn. Packt Publishing
7. Graupel D (2019) Principles of artificial neural networks: basic designs to deep learning, 4th edn. Packt Publishing, Birningham
8. Huang F, Zhang J, Zhou C, Wang Y, Huang J, Zhu L (2019) A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. Springer-Verlag GmbH Germany part of Springer Nature
9. Jarrah M, Salim N (2019) A recurrent neural network and a discrete wavelet transform to predict the Saudi stock price trends. Int J Adv Comput Sci Appl 10(4):155–162
10. Alphonse AS, Shankar K, Rakkini J (2020) A multi-scale and rotation-invariant phase pattern (MRIPP) and a stack of restricted Boltzmann machine (RBM) with preprocessing for facial expression classification. J Ambient Intell Hum Comput. 12:3447–3463
11. Hassan MM et al (2019) Human emotion recognition using deep belief network architecture. Inf Fusion 51:10–18
12. Maeda H, Kashiyama T, Sekimoto Y (2020) Generative adversarial network for road damage detection. Computer-Aided Civil and Infrastructure Engineering –willy online library
13. de Souza AA (2021) Simple hemogram to support the decision-making of COVID-19 diagnosis using clusters analysis with self-organizing maps neural network. In: Soft computing. Springer, Heidelberg
14. Zhong C et al (2019) Inland ship trajectory restoration by recurrent neural network. J Navig 72:1359–1377
15. Ghimire S et al (2019) Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. Appl Energy 253:113541
16. Lossau T et al (2019) Motion estimation and correction in cardiac CT angiography images using convolutional neural networks. Comput Med Imaging Graph 76:101640
17. Kagel. https://www.kaggle.com/andrewmvd/face-mask-detection
18. Haarcascades. https://github.com/opencv/opencv/tree/master/data/haarcascades.
19. Yuan Z (2020) Face detection and recognition based on visual attention mechanism guidance model in unrestricted posture. Sci Program 2020:1–10

# Design and Analysis of New Ultra Low Power CMOS Based Flip-Flop Approaches


**Naga Raju Jangam, Likhitha Guthikinda, and G. P. Ramesh**

**Abstract** The flip flops are essential part of the clocking circuits in complementary metal oxide semiconductor circuit based designs. The adiabatic flip flops are more useful in digital systems for clock switching applications. The clock switching approach with energy restoration is most popular and prominent for reducing power dissipation in ultra-low power based digital system designs. These flip flops playing key role in the design of energy efficient adiabatic clock switching methods and these flip flops works on the basis of adiabatic principle. In this research paper, we have simulated and analyzed the behavior of energy restoration flip flops. These are observed as one end condition capture and differential condition capture flip flops. Both the flip- flops are more useful in energy recovery methods. For better enhance results, we can use clock gate switching method together with energy restoration method. These are verified using cadence 180, 90 nm library technologies. Finally, we obtained desired results after simulation.

**Keywords** Adiabatic flip flops · Complementary metal oxide semiconductor · Pull up · Pull down networks · Power dissipation · Time delay · Very large-scale integration

## 1 Introduction

In CMOS transistor digital circuits, flip-Flops are most important component clock switching applications [1]. They are conscientious for synchronous, asynchronous method approaches. These are mainly used to obtain low power consumption in digital circuit designs [2, 3]. The flip flop is used to obtain the output load in digital circuits. This load is referred as clock load, which directly influence the power utilization. So by adding a process element we can overcome power utilization losses of

N. R. Jangam (✉) · L. Guthikinda
Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India
e-mail: nagaraju.jangam@mlrinstitutions.ac.in

G. P. Ramesh
St. Peters Institute of Higher Education and Research, Avadi, Chennai 600054, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_28

flip flops. It is important to enhance the performance of flip-flop network; as the modernization of digital circuits the power is limited. Hence the design of flip-flop circuits is more critical because of their large capture time and setup time [4, 5]. Re-use of power or energy recovery in the design is a low power method that can build to obtain low power distribution. It can implement by limiting the power as a resistor that will flow across the region. Therefore, there will be a lower value for power outages in parts of the region [6]. These circuits use renewable energy or sinusoidal electricity. Here, we have used the re-use of adiabatic power or techniques in a network enclosed by a natural power-based signal and it is necessary also. The principle used here is the power supply circuits and reuse power at the load. This occurs during each performance cycle from input node to output load.

Adiabatic flip-flop design utilizes power recovery from a clock network, which leads to a significant reduction in distribution of power. This watch can be made of highly efficient and durable materials. These results are simulated on cadence digital simulator tool [7, 8]. In this design we addressed on flip flop power recovery through single ended conditional capturing energy recovery (SCCER). By adding clock gating technique in flip-flop circuit we estimated its power recovery, timing constraints. This approach reduces over all power dissipation, delay constraints [9, 10].

## 2 Related Works

Raghav and Bartlett [15] has analyzed power clock generation by utilizing step charging circuits. In this literature, the overall adiabatic system, and the impact of adiabatic load on the energy dissipation of four phase power clock generate were analyzed. Kumar and Kumar [16] has designed a new architecture of energy recycling for lower power applications. In this literature, the architecture was designed on the basis of SR flip-flops, ultra-low power diode, single bit full adder, and adiabatic logic inverter. Biswas et al. [17] implemented and designed a two-bit magnitude comparator using several adiabatic logic gates and then the performance of the designed architecture was compared with other circuits in terms of transistor count and power consumption. Dhananjay and Salman [18] has implemented a charge based approach for mounting a correlation based analysis on the basis of SIMON core. Maheshwari et al. [19] addressed the issues of existing works, and developed a new approach to detect the circuit's invalid operations. In this literature, author developed a model containing adiabatic logic gates and trapezoidal power-clock.

## 3 Design of Adiabatic Flip-Flops

The most commonly chosen technology is the attractive and effective CMOS standard technology for the design of a low power circuit due to power outages. Generally, CMOS uses two networks, one is Pull up network (PUN) and pull down network (PDN). The power utilization in CMOS circuits is observed due to charging and discharge capacitances. These are associated with clock gating circuits and it is equal to double of circuit's electrical power utilization, alternating of frequency clock. Other factors in the region increase the energy consumption of the system but demand for low power. According to adiabatic technique, the discharge capacitance is charged with the help of continuous current source, standard CMOS structures [11]. In PMOS transistor the pull up network resistance must be R and the output capacitance as Co. The Constant current source is similar to the voltage ramp where we can observe, power consumption varies with the variation of resistor R [12, 13].

The power distribution is reduced by reducing the resistor R of the PMOS circuit. This can be done without power dissipation and power stored during the charging process at the capacitance of the load and it is connected to the source supply. It can be achieved with changing the source direction. This makes it attractive to adiabatic logic as it regenerates its energy, hence it is also called regenerative energy. The allocation used should be able to reverse the charge back on the delegation to this provision should be the same. Instead of using conventional supply it utilizes a source supply [14]. They are more design parameters are to be considered in CMOS design approach. It is implemented by two building methods. First, the implementation should lead to the construction of source supply, clock generating circuit. Second, it must observe other rules such as, transistor should be in stable state and provide minimum pulse power.

## 4 Experimental Analysis of Adiabatic Flip-Flops

### *4.1 SCCER Flip-Flop Approach*

Flip flop circuits are more useful in digital circuits. We analyzed adiabatic SCCER-flip-flop especially for ultra-low power applications. SCCER stands for one conditional scan SCCER stands for one of the conditions of termination of power acquisition. There are two circuits under this SCCER. They are.

1. SCCER flip-flop
2. Clock gating implemented in SCCER flip-flop.

There are usually FPTG flip-flops representing the four-phase transmission gates, which are commonly used in digital circuits. But these SCCER flip flops work better locally and work harder than FPTG flip-flops. Both of these flip-flops are very similar to each other and are commonly used. The goal behind this is an adiabatic system.

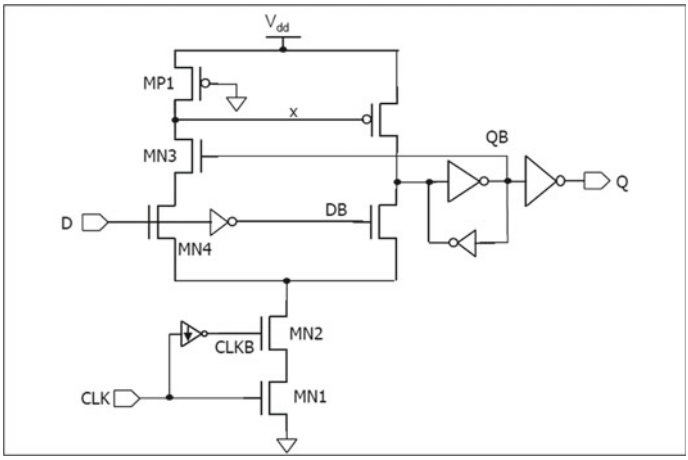



**Fig. 1** SCCER flip flop

Power recovery here is done by installing a clock. From this we can see that it has the character of gaining power. Here are some storage items and interiors which are connected with constant source supply. The main circuit shown in Fig. 1.

As if most inverters contain drag and drop transistors here SCCER flip-flops also contain. The MP1 transistor in the region is a drag transistor which means PMOS. The NMOS transistors MN1, MN2, MN3and MN4 are called as transceiver transistors. Output is observed at node QB, which controls the MN3 transistor. This provides conditional photography. On the right side there is a vertical test method is also used to overcome conditional capture process. The transistor MN3 is connected on top ofMN4. It reduces the payment sharing of the network. This is due to the charging period power loss at transistor MN3, which has already charge to maximum range. Hence the contribution of the pay-sharing process is not possible. Due to this phenomenon the flip-flop will works in both sleep mode and active mode.

## *4.2 Clock Gating SCCER Flip-Flop*

As the name implies, the difference between SCCER flip-flop with and without clock is a basic circuit that contains clock inputs and not gate and clock using SCCER flip-flop there is also no power supply or gate which is there as an alternative. Easily clock closures are done by replacing the inverter on the SCCER flip-flop with the NOR gate on the SCCER and the clock. A large amount of energy wasted due to clock gating logic circuits. This design approach divides entire clock scheme circuits from other parts of the digital circuits. This approach is followed with D-flip-flop concept. Clock gating implementation of SCCER flip-flop is represented in Fig. 2.

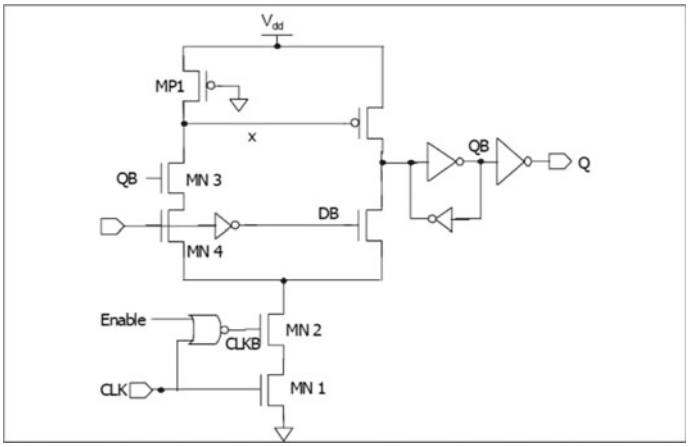



**Fig. 2** Clock gating implemented in SCCER flip-flop

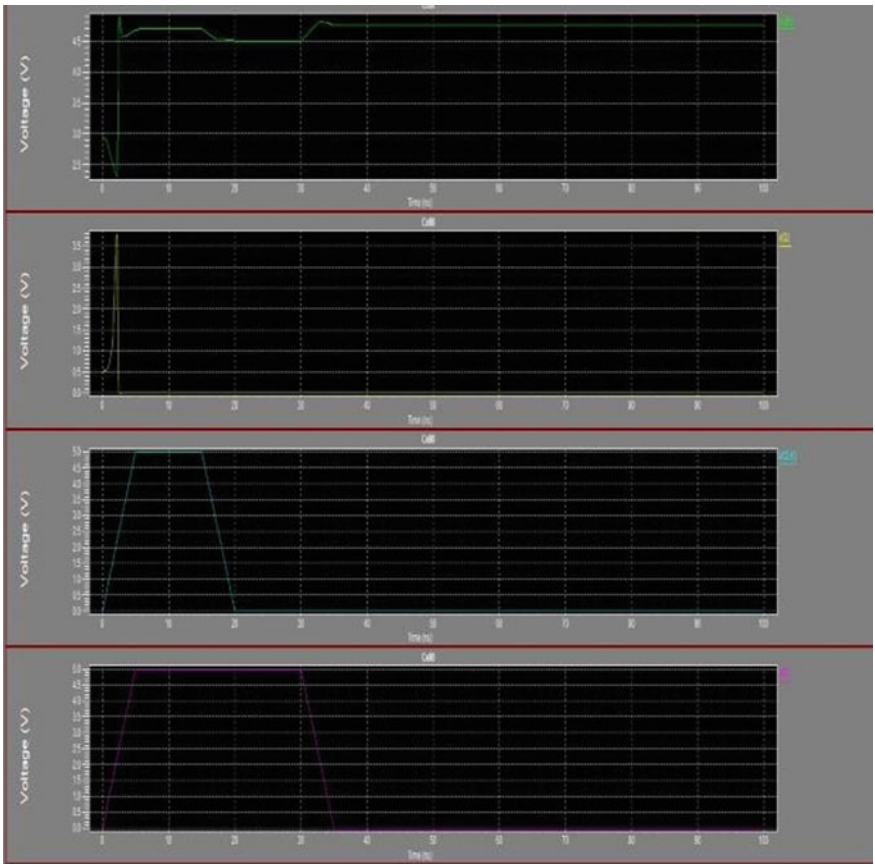



**Fig. 3** Simulation result of SCCER flip-flop

### *4.3 Experimental Results*

These SCCER leaflets are been utilized to achieve ultra-low power consumption and lower delay signals. The simulation results of SCCER flip-flop and clock gating implementation of SCCER are denoted in Figs. 3 and 4.

Power dissipation can be minimized by controlling cock gating circuit in sleep mode. Power dissipation values from the circuit will be displayed in the Tables 1 and 2. The clock gating circuit will utilize more amount of power consumption, and we compare power dissipation with existing papers [16] and [17].

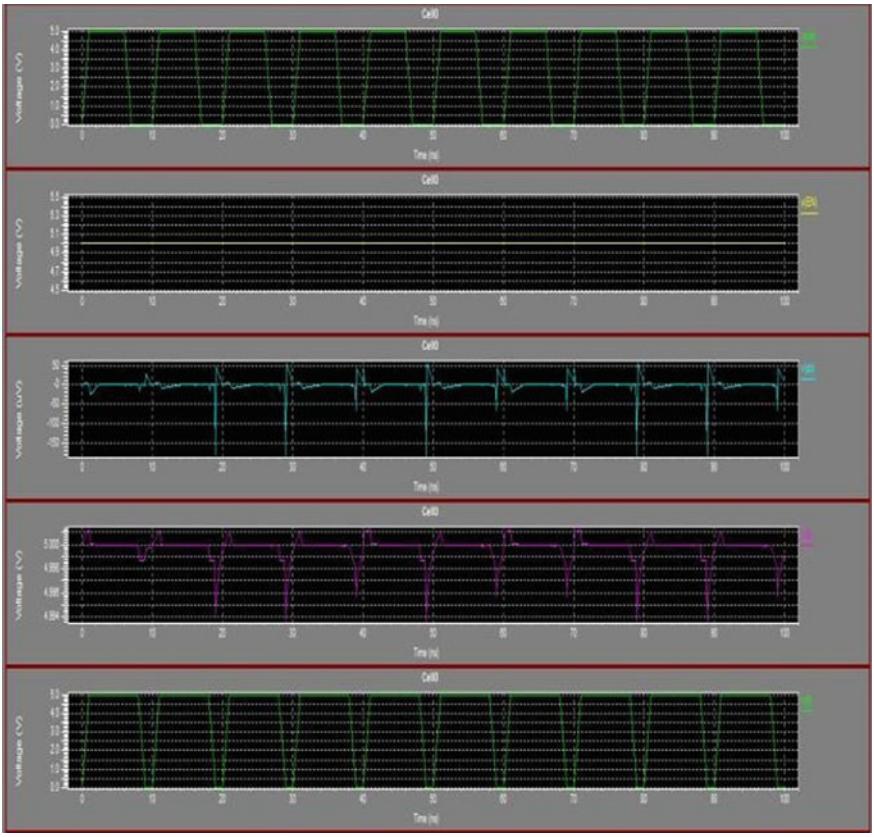


**Fig. 4** Simulation result of clock gating implemented using SCCER

**Table 1** Power dissipation of SCCER flip- flop in sleep mode

| Switching activity | Existing system Power dissipation [16] | | Proposed system Power dissipation | |
|---|---|---|---|---|
| | P (Vdd) ( μw) | P (clk) ( μw) | P (Vdd) | P (clk) ( μw) |
| 0 | 12 | 13.1 | 7.7 | 3.2 |
| 50 | 49.1 | 10.1 | 9.8 | 3.1 |
| 100 | 99.8 | 8.99 | 13.5 | 3 |

**Table 2** Power dissipation of SCCER flip- flop in Active mode

| Switching activity | Existing system Power dissipation [17] | | Proposed system Power dissipation | |
|---|---|---|---|---|
| | P (Vdd) (μw) | P (clk) (μw) | P (Vdd) (μw) | P (clk) (μw) |
| 0 | 12.1 | 13.1 | 11.6 | 11.2 |
| 50 | 52.4 | 10.9 | 48.1 | 10.1 |
| 100 | 110.8 | 9.7 | 93.9 | 8.1 |

## 5 Conclusion

In this research, Tables 1 and 2 demonstrate power output of SCCER flip-flop without using gating circuits in sleep and active modes. For the various switching functions, the energy dissipated by the clock and other circuits power are separated by the remaining flip-flop. Table 1 show the power output of SCCER flip-flop in sleep mode. Table 2 shows the power dissipation of SCCER flip-flop in active mode. It is observed that the power dissipation of the clock and the power consumption of all other circuits are less in the case of sleep mode. In sleep mode the SCCER flip-flop has better performance than other flip-flops. In active mode SCCER also has better flip-flop performance. The clock gating system is implemented without overloading the clock. SCCER flip-flops are preferred for adiabatic time closure in digital building systems because of their good power and delay signals.

## References

1. Wu Q, Pedram M, Wu X (2000) Clock-gating and its application to low power design of sequential circuits. IEEE Trans Circ Syst I 47(3):415–420
2. Matsuzaki Y, Hakoshima H, Seki Y, Kawabata S (2020) Quantum annealing with capacitive-shunted flux qubits. Jpn J Appl Phys 59(SG):SGGI06
3. Frank MP (2003) Common mistakes in adiabatic logic design and how to avoid them. In: Proceedings of the international conference on embedded systems and applications, held in Las Vegas, Nevada, 23–26 June 2003. CSREA Press, pp 216–222
4. Ojha P, Rana C (2015) Design of low power sequential circuit by using adiabatic techniques. Int J Intell Syst Appl 7(8):45–50

5. Li F, Takeshita Y, Hasegawa D, Tanaka M, Yamashita T, Fujimaki A (2021) Low-power high-speed half-flux-quantum circuits driven by low bias voltages. Supercond Sci Technol 34(2):025013
6. Hu J, Xu T, Li H (2005) A lower-power register file based on complementary pass-transistor adiabatic logic. IEICE Trans Inf Syst 88(7):1479–1485
7. Bhaaskaran VSK (2016) Low power VLSI circuit designusing energy recovery techniques. In: Design and modeling of low power VLSI systems, vol 6. IGI Global, pp 128–164
8. Samanta S (2009) Adiabatic computing: a contemporary review. In: International conference on computers and devices for communication (CODEC), pp 1–4 (2009)
9. Zhang W, Zhou D, Hu X, Hu J (2008) The implementations of adiabatic flip-flops and sequential circuits with power-gating schemes. In: IEEE 51st midwest symposium on circuits and systems, MWSCAS 2008, Knoxville, Tennessee, pp 767–770
10. Raju JN (2016) CMOS voltage controlled oscillator (VCO) design with minimum transistors. Int J Res Trends Innov (IJRTI) 1:88–91
11. Karthik R (2017) Implementation of flip- flops using QCA tool. J Fund Appl Sci 10:2332–2341
12. Prasad SVS, Savithri TS, Krishna IVM (2017) Performance evaluation of SVM kernels on multispectral LISS III data for object classification. Int J Smart Sens Intell Syst 10(4):829–844
13. Prasad SVS (2018) Double block zero padding acquisition algorithm for GPS software receiver. J Autom Mob Rob Intell Syst 12(4):58–63
14. Prasad SVS, Savithri TS, Krishna IVM (2014) A new technique for color based image segmentation using support vector machines. In: International conference on medical imaging, m-health and emerging communication systems (MedCom), Greater Noida, India, pp 189–192
15. Raghav HS, Bartlett VA (2020) Investigating the influence of adiabatic load on the 4-phase adiabatic system design. Integration 75:150–157
16. Kumar D, Kumar M (2021) VLSI implementation of wave shaping diode based adiabatic logic (WSDAL). Int J Electron 108(4):589–606
17. Biswas S, Mukherjee DN, Panda S, Maji B (2021) Power efficient magnitude comparator using adiabatic logic and gate diffusion technique. In: Advances in medical physics and healthcare engineering. Springer, Singapore, pp 207–213
18. Dhananjay K, Salman E (2021) Charge based power side-channel attack methodology for an adiabatic cipher. Electronics 10(12):1438
19. Maheshwari S, Bartlett VA, Kale I (2021) A VHDL-based modeling approach for rapid functional simulation and verification of adiabatic circuits. IEEE Trans Comput-Aided Des Integr Circ Syst 40(8):1721–1725. https://doi.org/10.1109/TCAD.2020.3022334

# Design and Comparative Analysis of Microstrip Patch Antenna by Using Various Materials in HFSS


**G. P. Ramesh, Pallavi, Hanifa Abdullah, and B. D. Parameshachari**

**Abstract** Microstrip patch Antenna is most widely used in the design of wireless Applications. The designed antenna works on a resonating frequency of 2.4 GHz which is suitable for Wireless Local Area Network (WLAN). Now-a-days flexible substrates including polyethylene, polyester and polyamide have become more obligatory in order to provide increased flexibility in wearable sensors. The intensity behind using polyethylene, polyester and polyamide materials is to check the tendency of the material in terms of Return Loss and VSWR. Small wearable antennas are used in receiving or transmitting communication and also used in IOT and medical systems in order to enable new applications through wireless connectivity. These three materials have significant properties for use in wearable sensors. In this research work Microstrip patch antenna is designed and simulated using HFSS software with flexible polyethylene, polyester and polyamide materials which provides an outcome of most efficiently used different kinds of antenna materials at a thickness of 3.6 mm. We exposure a selection of most flexible materials by comparing their voltage standing wave ratio (VSWR) and Return Loss (RL).

**Keywords** Bandwidth · Dielectric constant · Microstrip patch antenna · Return loss · VSWR · Wireless Local Area Network

G. P. Ramesh (✉) · Pallavi
St. Peter's Institute of Higher Education and Research, Chennai, India
e-mail: rameshgp@yahoo.com

H. Abdullah
University of South Africa, Pretoria, South Africa
e-mail: abdulh@unisa.ac.za

B. D. Parameshachari
GSSS Institute of Engineering and Technology for Women, Mysuru, India
e-mail: paramesh@gsss.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_29

## 1 Introduction

An antenna means a device that usually converts electrical power into radio waves and contrariwise. Within the research work the foremost popular antenna is employed i.e., is Microstrip Patch Antenna, which are getting more in used nowadays Microstrip patch antenna mandatorily consists of a radiating patch that's present on one side of the dielectric substrate and also the bottom plane on the others side. Because of increasing interest in research of antenna for communication systems associated with the body, there was an institution of IEEE 802.15 standardization group so as to develop standardize applications that are mostly intended for on-body, off-body, and in-body communication. Particularly, the Microstrip patch antennas are good candidates for body-worn applications, as they mainly radiate perpendicularly to the planar structure, and also their ground plane efficiently shields the body tissues [1–6]. For example, the permittivity and the thickness of the dielectric substrate mainly determine the bandwidth and also the efficiency performance of the planar antenna [7]. The types of wearable materials used to design Microstrip Patch Antenna are: Polyester, Polyamide, and Polyethylene. The compact wide beam micro strip antenna also implemented for the CNSS application [8]. Moreover, the novel Microstrip antenna is used for the photonic crystal process with terahertz frequency [9]. In the current field, miniaturization, fuse deposition model, and hexagonal Microstrip antenna is implemented to improve the ultra-wide band applications [10–12].

## 2 Related Works

Guorong et al. [13] proposed frequency doubling Microstrip patch antenna for wireless strain detection. The patch antenna dimension at different length was performed effectively in the different dielectric constant. But, the frequency increment value (1 MHz) was not sufficient to get a sufficient gain. Anandkumar and Sangeetha [14] presented the aperture coupled Microstrip path antenna for RADAR applications. The proposed work design provided the 5 dB gain, 4.5 dB directivity, and 1.5 GHz frequency. However, the VSWR of the patch antenna didn't provided the effective range which affected the radiation pattern. Subha et al. [15] proposed suppress of surface wave propagation in Microstrip patch antenna for wireless communication application. The hindrance of the surface wave propagation was not reduced in the bandgap structure. Rajeshwari and Anbalagan [16] developed the android based mobile application for designing the Microstrip patch antenna. In this work, nonlinear vector decomposed neural network was implemented to calculate the performance parameters. The frequency of the antenna was represented as 60 GHz with 1.14 $\times$ 1.94 size of the length and width. However, the proposed antenna gain value is low (4.8 dB) which needs to be improved. Venkatesh and Naveen Kumar [17] proposed Nanocomposite based miniaturized S band antenna for novel substrate materials. The

frequency range of the antenna is operated at 2.2–3.8 GHZ frequency. However, the substrate size is high (4 × 4 × 0.157) which increase the area of the entire antenna.

## 3 Design of an Antenna

In this design, we are concentrating on a rectangular Microstrip patch antenna which consists of a rectangular patch of length [Lp] and width [Wp] of the patch. The proposed antenna works on the wireless local area network (WLAN) frequency of 2.4 GHz (2400–2484 MHz) which is shown in Fig. 1. The Substrate material used are polyester, polyethylene, polyamide used as a dielectric material with its dielectric constant are 4.1–5.2, 2.3 and 2.5 and thickness of 3.6 mm. The proposed antenna consisting of the dielectric substrate, patch as well as Microstrip feed line. The rectangular patch is separated from the ground plane with some Dielectric substrate with the above-shown dimensions of the system.

The proposed antenna is predicated on the idea of the simulated design in HFSS simulation software which is shown in Fig. 2. The simple structure and configuration

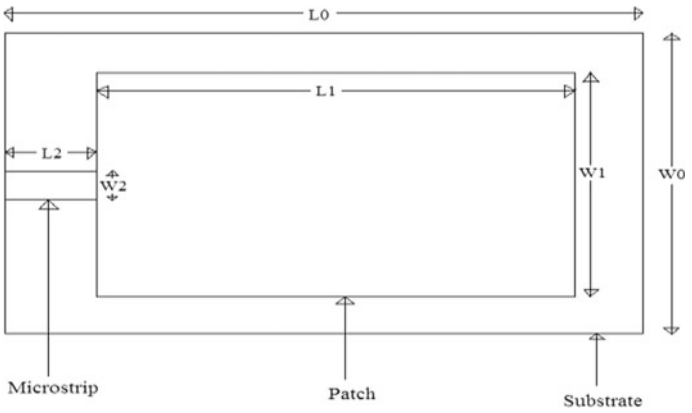



**Fig. 1** Proposed antenna design

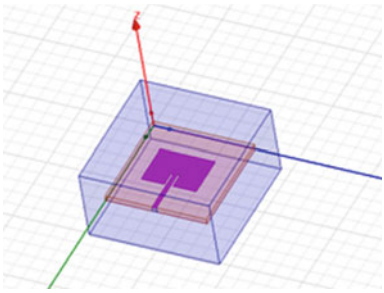


**Fig. 2** Antenna design in HFSS

**Table 1** Dimensions of the proposed antenna

| Parameters | Dimensions (mm) |
|---|---|
| Lg (Length of the ground) | 70 |
| Lp (Length of the patch) | 28 |
| Wg (Width of the ground), | 70 |
| Wp (Width of the patch | 38 |
| T (Thickness,) | 3.65 |
| Fr (Resonant frequency) | 2.4 GHz |

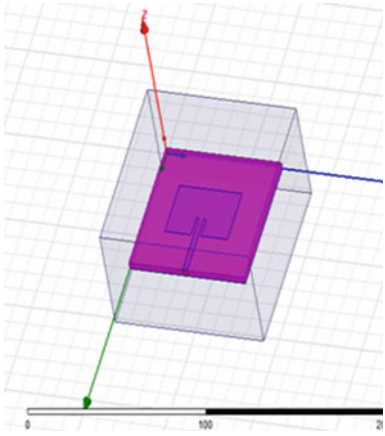


**Fig. 3** Design of microstrip patch antenna using polyamide material

of the antenna and the low profile of the proposed antenna make the fabrication process easy and also suitable for application in the WLAN. Finally, the simulated results obtain better results when compared to the other system. The dimension of the proposed antenna is given in Table 1 which contains the antenna parameters.

Here, the Return Loss & VSWR are,

RLαsignal that is reflected as a result of an impedance mismatch.
RL = 10log10 (Power Incident/Power Reflected)
VSWR = |Maximum Voltage|/|Minimum Voltage|

## 4 Types of Material Used

### *4.1 Polyamide Material*

It is a form of wearable material which is used in the design of an antenna. The foremost famous style of polyamide material is nylon which is shown in Fig. 3. It's having medium heat retention abilities. Polyamide materials were also used in

parachute materials during World War-II. Polyamides may additionally be synthesized from dinitriles using acid catalysis via an application of the Ritter reaction. This method is usually used for the preparation of nylon 1, 6.

To design the Rectangular micro strip patch antenna by using Polyamide material in HFSS software, we have used dielectric constant as 2.5 and thickness as 3.6 mm with resonant frequency at 2.4 GHz. The proposed antenna return loss and VSWR are shown in Fig. 4(a) and 4(b).

**Experimental Results Obtained:** The Return loss is –21 db and VSWR is 1.4 at 2.4 GHz.

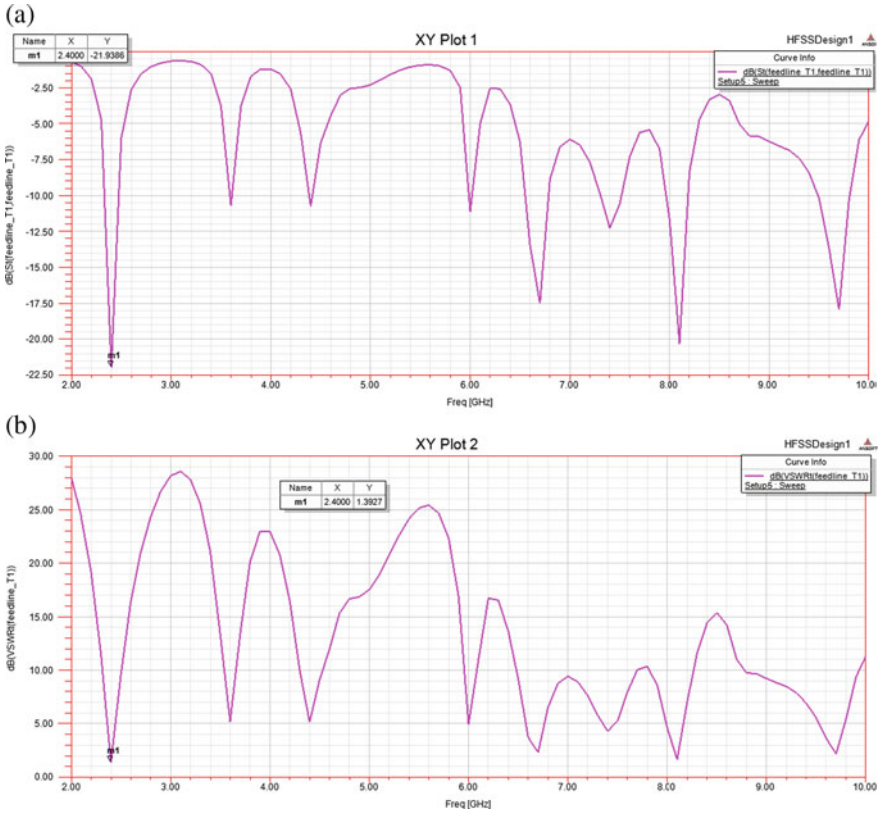



**Fig. 4** Microstrip patch antenna using polyamide material **a** Return loss **b** VSWR

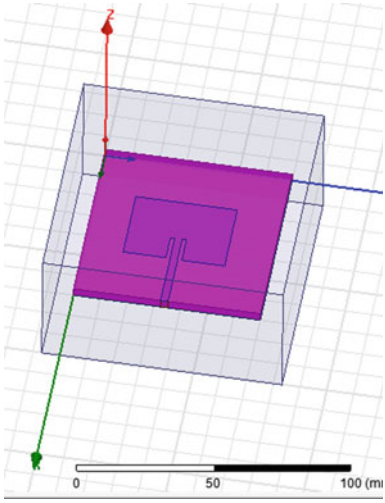



**Fig. 5** Design of microstrip patch antenna using polyester material

## *4.2 Polyester Material*

Polyester was selected because the antenna substrate thanks to its excellent dielectric characteristics and low moisture regains value. The promising properties, like their flexibility, light weight and value effectiveness, enable effortless integration of the proposed antenna into clothes like polyester jacket. Polyesters are one in every of the economically most vital classes of polymers, driven especially by PET, which is counted among the commodity plastics; in 2000 around 30 million tons were produced worldwide. The design of Microstrip Patch Antenna using Polyester Material is shown in Fig. 5.

To design the Rectangular micro strip patch antenna by using Polyester material in HFSS software, we have used dielectric constant between 4.1–5.2 and thickness as 3.6 mm with resonant frequency at 2.4 GHz. The proposed polyester material antenna return loss and VSWR are shown in Fig. 6(a) and 6(b).

**Experimental Results Obtained:** The Return loss is –18 db at 7.6 GHz and VSWR is 0.8 at 9.8 GHz.

## *4.3 Polyethylene Material*

Polyethylene is the most widely used thermoplastic polymer. Polyethylene substrate has become an important candidate substrate material for the development of future flexible devices as well as conformal and wearable wireless antenna systems. It is a flexible substrate for conformal antennas for body-worn applications. The design of Microstrip Patch Antenna using Polyethylene Material is shown in Fig. 7.

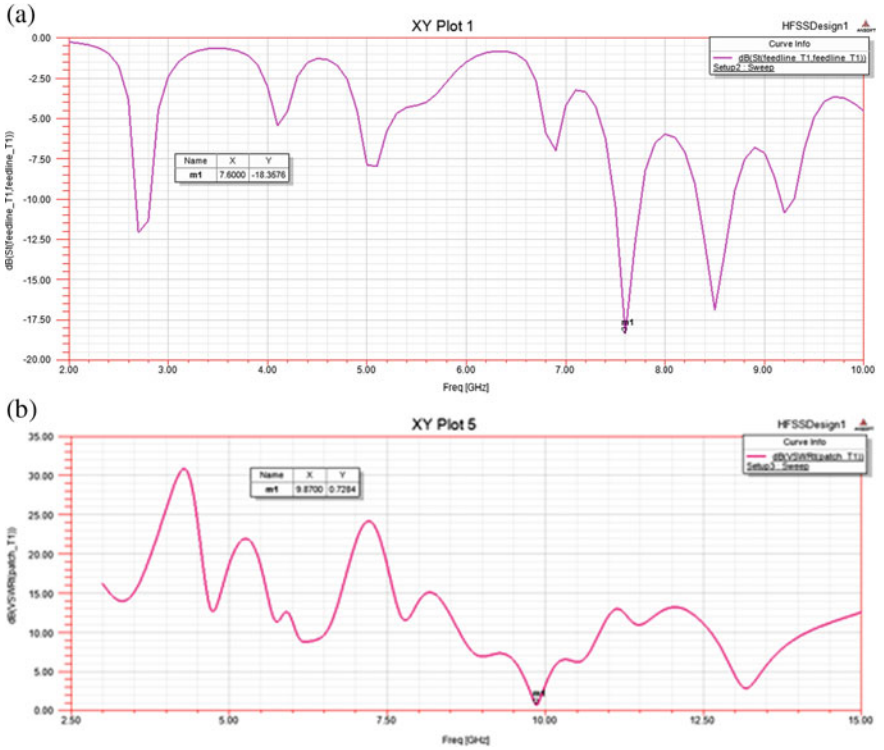



**Fig. 6** Microstrip patch antenna using polyester material **a** Return loss **b** VSWR

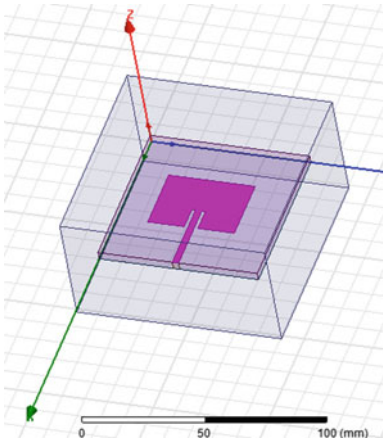



**Fig. 7** Design of microstrip patch antenna using polyethylene material

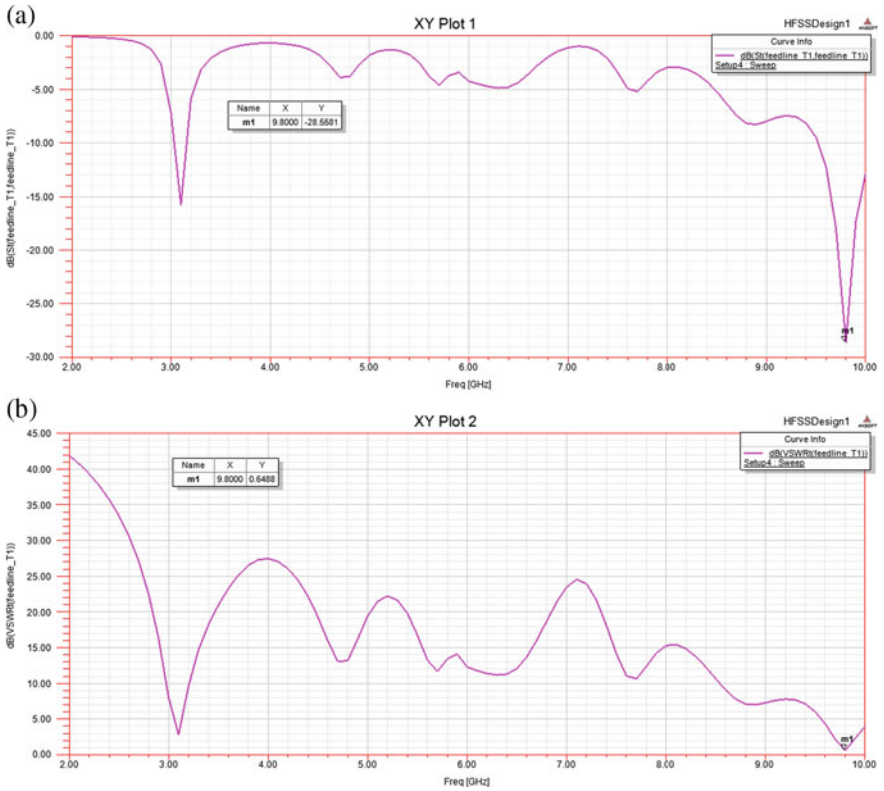



**Fig. 8** Microstrip patch antenna using polyethylene material **a** Return loss **b** VSWR

To design the Rectangular micro strip patch antenna by using Polyester material in HFSS software, we have used dielectric constant at 2.3 and thickness as 3.6 mm with resonant frequency at 2.4 GHz. The proposed polyethylene material antenna return loss and VSWR are shown in Fig. 8(a) and 8(b).

**Experimental Results Obtained:** The Return loss is –29 db at 9.8 GHz and VSWR is 0.6 at 9.8 GHz.

## *4.4 Comparison Between All the Material*

## *4.5 Comparative Analysis*

The Comparison of return loss and VSWR for the different materials and the comparison of return loss, gain, frequency, and VSWR for the different antenna designs are

**Table 2** Comparison of return loss and VSWR for the different materials

| Materials | Parameters | |
|---|---|---|
| | Return loss | VSWR |
| Polyester | ~−18 db (at 7.6 GHz) | ~0.8 (at 9.8 GHz) |
| Polyethylene | ~−29 db (at 9.8 GHz) | ~0.6 (at 9.8 GHz) |
| Polyamide | ~−21 db (at 2.4 GHz) | ~1.4 (at 2.4 GHz) |

**Table 3** Comparison of return loss, gain, frequency, and VSWR for the different antenna designs

| Antenna designs | Return loss (dB) | Gain (dB) | Frequency (GHz) | VSWR |
|---|---|---|---|---|
| Guorong et al. [13] | −24 | 4.2 | 2.8 | 1.8 |
| Anandkumar and Sangeetha [14] | −19 | 5 | 1.5 | 1.6 |
| Subha et al. [15] | −22 | 4.3 | 2.2 | 1.9 |
| Rajeshwari and Anbalagan[16] | −18 | 4.8 | 1.4 | 1.4 |
| Proposed antenna | −29 | 6.5 | 9.8 | 0.6 |

given in Tables 2 and 3 respectively. In Polyamide, the VSWR and return loss is improved in the proposed work which is compared with the conventional related works [13–16]. When compared to the conventional antenna designs, the proposed antenna worked effectively inters of return loss, gain, frequency, and VSWR.

## 5 Conclusion

For polyester material we get return loss value & VSWR at 7.6 & 9.8 GHz frequency. It means here frequency deviation occurs. Also for polyethylene material we get return loss value & VSWR at 9.8 GHz frequency. It means here also there is a deviation in frequency. But for polyamide material we get return loss value at exact 2.4 GHz frequency. It means here there is no frequency deviation, distortion and also we get good result for impedance matching. The designed antenna is a low-profile antenna hence it is very compact, easy to fabricate, and it is also fed by a Microstrip feed line at the desired resonant frequency of 2.4 GHz which makes it in becoming a more attractive structure for current as well as future WLAN applications. From our analysis of different materials, we have observed that the polyamide was best suited for our design. Hence, if we have to go for wearable Textile antenna then we can prefer to use polyamide material.

**Future Scope.** This research work can be taken as a reference for the design of antennas using different materials. Suppose, if we want to go for a wearable antenna then we can take this research work as a reference and choose the type of material to be used to get the better performance of the antenna.

## References


1. Lee H, Li ES, Jin H, Li C, Chin K (2019) 60 GHz wideband LTCC micro strip patch antenna array with parasitic surrounding stacked patches. IET Microw Ant Propag 13(1):35–41
2. Mumtaz S, Bo A, Al-Dulaimi A, Tsang KF (2018) Guest editorial 5G and beyond mobile technologies and applications for industrial IoT (IIoT). IEEE Trans Ind Inf 14:2588–2591
3. Sun H, Zhang Z, Hu RQ, Qian Y (2018) Wearable communications in 5G: challenges and enabling technologies. IEEE Veh Technol Mag 13:100–109
4. Brebels S, Ryckaert J, Boris C, Donnay S, De Raedt W, Beyne E, Mertens RP (2004) SOP integration and codesign of antennas. IEEE Trans Adv Pack 27:341–351
5. Hertleer C, Rogier H, Member S, Vallozzi L, Langenhove LV (2009) A textile antenna for off-body communication integrated into protective clothing for firefighters. IEEE Trans Adv Pack 57:919–925
6. Zhang L, Wang Z, Psychoudakis D, Volakis JL (2012) Flexible textile antennas for body-worn communication. In: Proceedings of IEEE international workshop on antenna technology, Tucson, ZA, USA, 5–7 March 2012, pp 205–208
7. Pei R, Wang J, Leach M, Wang Z, Lee S, Lim EG (2016) Wearable antenna design for bio information. In: Proceedings of IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB), pp 1–4
8. Pan ZK, Lin WX, Chu QX (2014) Compact wide-beam circularly-polarized microstrip antenna with a parasitic ring for CNSS application. IEEE Trans Ant Propag 62(5):2847–2850
9. Kushwaha RK, Karuppanan P, Malviya LD (2018) Design and analysis of novel microstrip patch antenna on photonic crystal in THz. Physica B 545:107–112
10. Oraizi H, Hedayati S (2012) Miniaturization of microstrip antennas by the novel application of the Giuseppe Peano fractal geometries. IEEE Trans Ant Propag 60(8):3559–3567
11. Prakash C, Senthil P, Sathies T (2021) Fused deposition modeling fabricated PLA dielectric substrate for microstrip patch antenna. Mater Today: Proc 39:533–537
12. Sawant KK, Kumar CS (2015) CPW fed hexagonal micro strip fractal antenna for UWB wireless communications. AEU-Int J Electron Commun 69(1):31–38
13. Song G, Zhang B, Lyu Y, Sun T, Wang X, He C (2021) Application of frequency doubling in micro-strip patch antenna for wireless strain detection. Sens Actuators A 321:112403
14. Anandkumar D, Sangeetha RG (2020) Design and analysis of aperture coupled micro strip patch antenna for radar applications. Int J Intell Netw 1:141–147
15. Subha TD, Subash TD, Jane KC, Devadharshini D, Francis DI (2020) Study and analysis of suppress of surface wave propagation in microstrip patch antenna. Mater Today: Proc 24:2414–2423
16. Rajeswari P, Anbalagan P (2020) Design and deployment of android based mobile application for performance analysis of micro strip patch antenna. Microprocess Microsyst 77:103111
17. Venkatesh CK, Kumar SN (2018) Development of nanocomposite based miniaturized S-band antenna for novel substrate materials. Mater Today: Proc 5(4):10621–10628

# Design of an Efficient Mobile Communication and an Armament System for Women Safety


**Sesha S. Sankar, Valavala Sandeep, K. S. Viswesh, S. Vigneshwar, and C. B. Rajesh**

**Abstract** To ensure the fast and rapid development of smart cities, one of the foremost matter of concern is safety for women and children which has been a critical concern for the country. Even though many regulations, laws, devices and applications have been developed that do function to an extent, there hasn't been any decrease in crime rate against women. It is sad that even though equal rights are there for both men and women, the latter are always in fear because of the former, especially in India where women are considered goddesses. We have come with an approach which makes the chances of saving the victim better. We have employed the use of a powerful Raspberry Pi for this. The Raspberry Pi can capture images and videos of the assaulter which would be safely uploaded to Google Drive or Cloud for later retrieval while producing evidence. A smart push button-based model is developed which makes the victim call for help through her fingertips. The button system along with a GPS/GSM Module helps in alerting the emergency contacts or the police by sending an alert message, the latitude and longitude coordinates, along with the location on Google Map. A shock gun model is also attached which helps the victim to sustain until help arrives. A buzzer system is also included which can alert nearby people, with a loud sound to gather their attention. The entire device would be encapsulated such that the button can be pressed from a regular usage handbag.

**Keywords** Buzzer · Global positioning system · Google drive · Push button · Raspberry Pi · Shock gun model

## 1 Introduction

In the 21st century, India has been labelled as one of the most unsafe places to live in for women and children. Throughout the two decades, we saw so many movements that empowered women, change in laws that should have instilled fear among the men. This means that not only do women require extra safety, but the oppressors who

S. S. Sankar · V. Sandeep · K. S. Viswesh · S. Vigneshwar · C. B. Rajesh (✉)
Department of Electronics and Communication Engineering, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: cb_rajesh@cb.amrita.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_30

are 99.9% of the time men, need to be shown their place by bringing their patriarchal and animalistic ego down. Our project thrives on this major motivation to instill fear among the men who have been oppressing and committing heinous crimes for years, so that they wouldn't even think to commit one and to supply such a safety system for women that will not only keep them away at risk, but also give them the bravery and upper hand to defend themselves in case of aggressive advancements by the assault. This can be fit in a handbag for women who are travelling. The Raspberry Pi also works with a powerful GPS GSM Module, that will effectively track the location of the person in trouble and send their location via a help message on Telegram, E-Mail and WhatsApp of the emergency contacts or police so that they can arrive to their help seeing the GPS coordinates along with its Google Map marker location. Meanwhile, the victim in trouble can make use of two self-defense mechanisms and flee from the location until help arrives. Not only is safety prioritized, but the entire evidence in the form of image detected and video captured by the Raspberry Pi will safely be backed up to a storage spot such as Database, Google Drive or Cloud which can be accessed later for verifying.

## 2 Literature Review

Bharathwaj [1] has proposed a method to get the location of the police station which is situated the nearest to the victim and alert them through a message when the women/victim is in shear distress, there are some flaws which needs to be clear, though this gets the location of the police station which is near to the victim there will be a minimum amount of delay that is occurring due to the official themselves. Viswanath [2] had built a device so compact that it can be worn on the feet of the victim which consists of a light blue bean microcontroller, when the sound is measured by the sensor about 4–5 times and conveys the message through message or email. In this paper, it was mentioned that the speed of the light blue bean is minimum. Monisha [3] had built a device using buttons for meeting the emergency by single/double pressing the buttons while, pressing the buttons during such situation is not at all possible.

VamsiPriya [4] had built a portable device and a professionally written code for the entire project that made Code readable and more professional on the contrary, the microcontroller chosen for the same is very primitive one that is the Arduino Uno that can't do multitasking and the range of the same is mentioned to be 10 m that is very less. Rasool [5] was able to successfully implement a lot of test case real world situation with one single written Program. Also, they made use of common household bat as a self-defense mechanism but, the major disadvantage of the developed system is it leads to confusion when the person is severely suffering from a health issue and it has been explicitly mentioned that the circuitry is affected by weather. Krishnamurthy [6] had achieved in obtaining the GPS coordinates dynamically and updating the same every 30secs and it will be sent to the emergency helper numbers. Even if the women are safe, the device will be active state and conveys a false message to the registered

helper which may lead to waste of time etc. and this whole thing fails if the device goes into aero plane mode or into a network less area.

Prashanth [7] developed an application with particular features for security and authentication purpose. The person will get access to it only if he is authenticated, he can use the app. Even though, it is a security application providing their credentials and authenticating will be tough at a distress situation. Mehta [8] built a device which upon shaking the device gets activated and the location is sent to register emergency contacts. The major flaw in this system is that it may lead to confusion when mobile trembles accidently. Verma [9] developed a device for tracking the pulse rate and the temperature of the body. If the sensor reading is in abnormal state, then it automatically sends a location police station. To protect themselves the person must press a buzzer and the location is sent to their contacts. The fact to be noted in the paper is only the people who are close to the victim can help the affected person.

Dandamudi [10] developed a CNN based system aerial image processing which classifies if the women are in distress or not and creates an alert message. The main disadvantage of the system this is very difficult to be employed in everyday life and it may lead to very costly livelihood. Ravisankar [11] built an Android based security and tracking system for children which involves having a mobile phone and an app has been developed to send messages and email to the helper but it is not possible to have mobile phone in every situation. Mohan [12] had developed a wearable gadget which track the location and sends to the helper but it has no other feature and the feature of instant recovery has not been planned. [13] and [14] gives a general idea on what women are facing these days and how the problem can be solved using IoT and provides a wide range of ideas and facts on the problem and an IoT solution for the same.

## 3 Hardware Components

The device is designed in a way that, when the button is pressed by the user, it would track the location where the system is located and would also capture the face of the person, who is assaulting detects their gender, save it as a video and is uploaded to GDrive. The location tracked is been sent to the user's family/friends through telegram and Gmail. When the receiver wants to get location again after sometimes, he/she can send a specific keyword, so that the device can track the new location and send it again. The main hardware components used here are Raspberry Pi a GPS module, a camera module [15, 16].

### *3.1 Raspberry Pi and GPS Module*

Raspberry Pi is a device which is like a mini computer with high computational capability. The main reason we need to use raspberry pi is because of the high

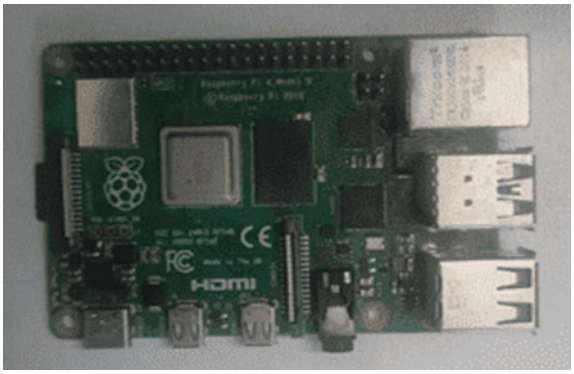


**Fig. 1** Raspberry Pi

**Fig. 2** GPS/GSM module

computational need when using programs for face detection, gender detection and so on. The Python 3 programming language is used to code the program for this project. The GPS module is a module which uses Global Positioning Satellite to get the location of the device and sends many data such as the latitude, the longitude, timestamp of the satellite and few more data of the location where the device is present. In this project we require only latitude and longitude coordinates of the device, which is depicted in Figs. 1 and 2.

## *3.2 Camera Module and 555 Timer Based Buzzer System*

The camera module is a module which is used to attach an external camera to the device. In this project the camera module is attached to raspberry pi and is placed in a location such that the camera is hidden and can capture the face of the assaulter. The camera module is used to get the face of the assaulter. This helps the victim to alert for help from nearby people is designed. This circuit consists of set and reset buttons, in which once the reset button is pressed the circuit becomes inactive, then once the set button is pressed, it becomes active by ringing the buzzer. The camera module and timer circuit are represented in Figs. 3 and 4.

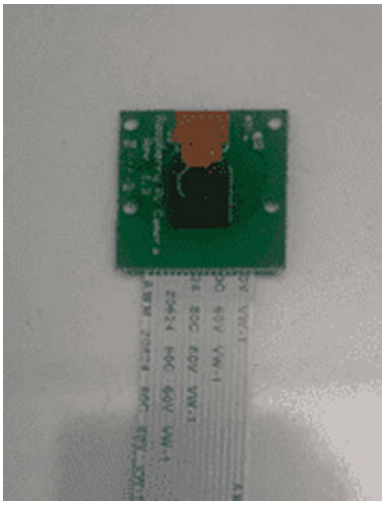


**Fig. 3** Camera module

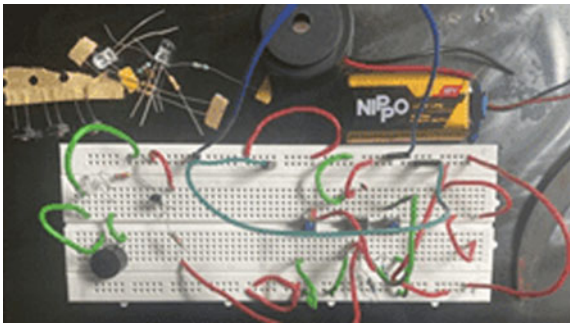



**Fig. 4.** 555 timer circuit

## *3.3 Shock Circuit*

This is for the self-defense of the victim before the help arrives which is very much necessary for the victim to survive. This circuit consists of a push button which once pressed, makes contact with the battery, thus charging the circuit. Once the shock circuit is done, it gets discharged and once again the button needs to be pressed to generate shock. In the circuit shown below, the circuit becomes complex if implemented on breadboard so we have soldered it together as shown in Fig. 5.

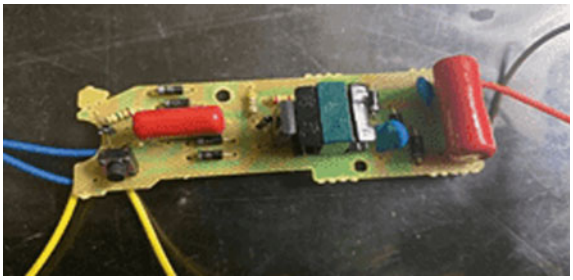


**Fig. 5** Shock circuit

## 4 Software Algorithm

The main program of this project involves many fractions, in which each function performs a specific task. These specific tasks are: Location detector, Sending the location through telegram and mail, getting image from camera and detecting the face from it and saving it, detecting gender from the image saved, Uploading the continuously saved image as video to the cloud, Multitasking/multithreading to perform two tasks separately at the same time. The two tasks are: Finding and sending location, detecting face and saving it in cloud.

### *4.1 Location Detector*

Serial, time, string, pynmea2 are the functions imported for this program into the raspberry pi. "Serial". The location tracking is started with the code "dataout = pynmea2. NMEAStreamReader" And the data is uploaded with the code "newmsg = pynmea2. "

### *4.2 Sending Message*

The message is sent through both mail and telegram. The message should contain "https://www.google.com/maps/#latitude,#longitude,14z", where latitude and longitude coordinates are entered in place of #latitude# and #longitude. "Send_message.".#text# is the message which is send.

### *4.3 Detecting Face*

For detecting the face, OpenCV is mainly imported. Numpy is also imported for the program. The file "*haarcascade_frontalface_default.xml*" is downloaded and is imported to this program. With the function "cap = cv2.Video Capture (0)" The camera module is accessed and the image is captured. The image is converted into grayscale with the code "gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)". The image is converted into grayscale so that the face can be detected easily than having colored images. With the function "faceCascade.detectMultiScale()", the presence of face is detected and its coordinates in the image is marked and is saved as a video.

### *4.4 Saving in Cloud*

Creating a backup storage forum in Google Drive, Cloud or Database for the image detected and video captured, in order to show evidence. One of the main aspects of this project comes down to providing evidence of the attacker's presence or actions. For this purpose, the video captured and face detected by the Raspberry Pi, has to be stored in a safe place like a Database, Cloud or Google Drive. This can be accessed later by uploading the video and image directly to Google Drive using Python. The essentials for this are: Creating an OAuth Client ID and Secret Key along with an authorized URI using one's google account, storing the key in the same directory as the python program, installing Pydrive package for connection to Google Drive.

### *4.5 Multitasking/Multithreading*

When each task is within the same program and is done at the same time, then this type of multitasking is called multithreading. #FunctionName# is the name of the function which is processed separately aside from main program. Note that the main program is also a separate process which is independent of the newly created thread.

## 5 Experimental Results

When the push button is pressed the entire program gets implemented. When the program runs, the GPS/GSM will work simultaneously to gather the location and a google map link will be sent to the mail of the contacts registered by the user. If the helper doesn't have a Gmail account to receive mail, the same can be obtained by the Telegram bot that we have built. The results are shown in Figs. 6, 7, 8, and 9.

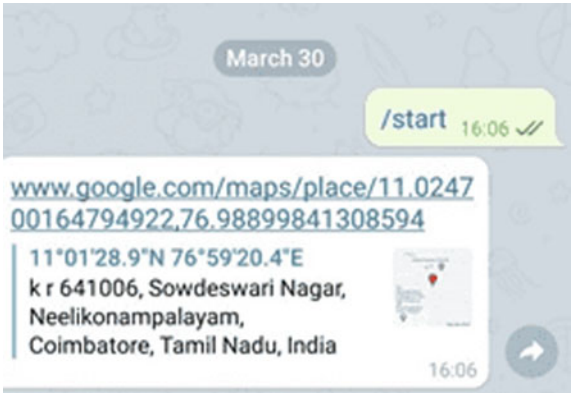



**Fig. 6** Telegram message

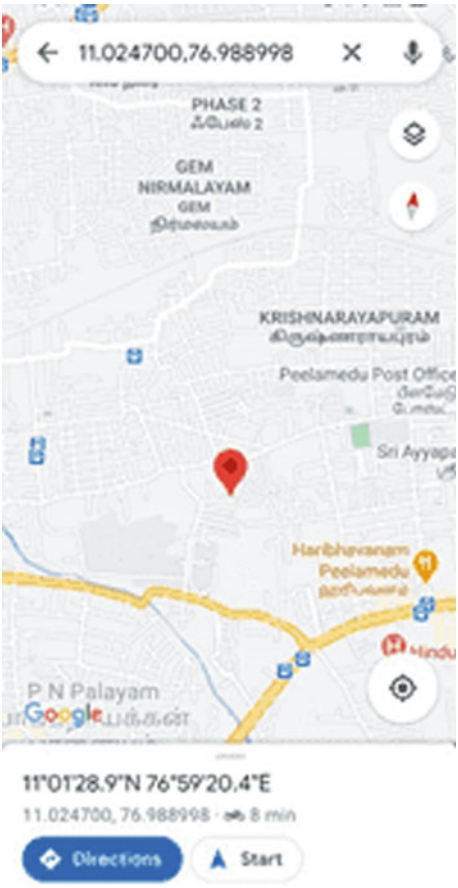



**Fig. 7** Google map

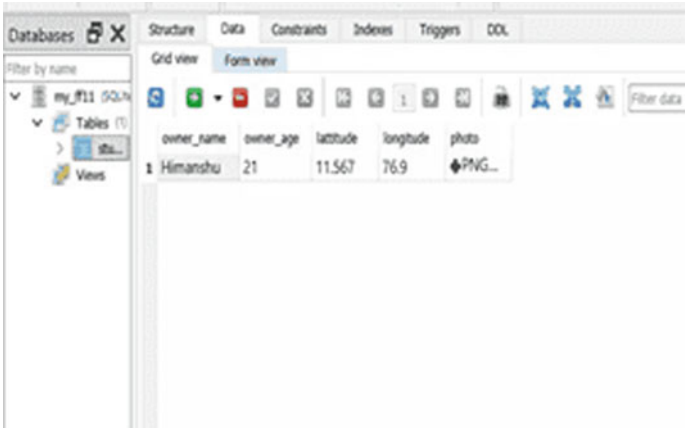



**Fig. 8** Database for storing the received data


**Fig. 9** Uploading files to Google Drive

## 6 Conclusion

Our women safety system aims to supply maximum support to the victim in the panic situation. Our shock gun system will help the victim in serving until the help arrives. The GPS system helps in sharing the location coordinates which makes up the main aim of this project. Unlike other similar works, we derive the GPS coordinates and display it on google maps. Buzzer system, camera model with drive backup are

handy add ones to our project. With the technology growing every day there is a lot of future scope for this project to even reduce the time gap for the help to come. The implementation of individual circuits is big, because of which encapsulation of the entire circuit and making them into one small sized circuit is tedious. The programming is multi-fold and complex because of the usage of multiple concepts which needed different syntaxes. Unpredictable weather conditions might affect the circuitry. The Shock Circuit has a disadvantage that the law doesn't legally allow its output voltage and licensing for a Stun Gun. In future work, with the help of VLSI technology, the circuits can be made much smaller and easily portable as a wearable device. Even though this project offers a good mechanism for safety and self-defense, the moral compass of the assaulter cannot be predicted. There are many possible situations of behavior of an assaulter, which must be taken in consideration for further development and to ensure 100% security. Further, this can also lead to creation of an interactive application on the mobile phone which will send help from the nearest police station to the women in trouble.

## References

1. Bhardwaj N, Aggarwal N (2014) Design and development of – suraksha. IEEE Int J Inf Comput Technol 4. ISSN 0974–2239
2. Viswanath N, Pakyala NV, Muneeswari G (2016) Smart foot device for women safety. In: IEEE Conference. ISBN: 978-1-5090-0931-2
3. Monisha DG, Monisha M, Pavithra G, Subhashini R (2016) Women safety device and application – FEMME. Indian J Sci Technol. ISSN: 0974–5645
4. Priya NLV, Sai KB, Naveen PB, Manaswini GJ, Krishna DSR, Krishna BSG (2018) Self defense system for women safety with location tracking and SMS alerting through GPS & GSM networks. World Wide J Multidisc Res Dev. E-ISSN: 2454–6615
5. Rasool SM, Fatima SB, Farooqi SA, Hassan SS, Saber S (2020) Location tracking based women safety system with electric shock. IEEE Int J Inf Comput Technol. ISSN: 0950–0707
6. Krishnamurthy V, Saranya S, Srikanth S, Modi S (2017) M-WPS: mobile based women protection system. In: IEEE international conference on energy, communication, data analytics and soft computing. ISBN: 978-1-5386-1887-5
7. Prashanth DS, Patel G, Bharathi B (2018) Research and development of a mobile based women safety application with real- time database and data-stream network. In: IEEE international conference on ciruits power and computing technologies. ISBN 978-1-5090-4967-7,
8. Mehta S, Janawade S, Kittur V, Munnole S, Basannavar S (2017) An android based application or women safety. Int J Eng Sci Comput
9. Vemia PK, Sharma A, Varshney D, Zadoo M (2018) Women safety device with GPS, GSM and health monitoring system
10. Dandamudi AGB, Vasumithra G, Praveen G, Giriraja CV (2020) CNN based aerial image processing model for women security and smart surveillance. In: Paper presented at the proceedings of the 3rd international conference on smart systems and inventive technology, ICSSIT 2020, pp 1009–1017. https://doi.org/10.1109/ICSSIT48917.2020.9214143
11. Ravisankar K, Vanitha V (2018) Android-based security and tracking system for school children. https://doi.org/10.1007/978-981-10-8228-3_50
12. Mohan R, Raj CV, Aswathi P, Bhavani RR (2018) UAV based security system for prevention of harassment against woman. In: 2017 International conference on intelligent computing, instrumentation and control technologies, ICICICT 2017

13. Hemasagar K, Kumar RM, Manojna VS, Priya BK (2020) Ingenious bracelet for women safety and security using IoT. In: Sixth international conference on emerging research in computing, information, communication and applications, ERCICA 2020. Nitte Meenakshi Institute of Technology, Bangalore
14. Bavya MP, Raghunandan MV (2018) A study on problems faced by working women of BPO sector in Mysuru. Int J Mech Eng Technol 9(1):399–406
15. Venkatesh K, Parthiban S, Kumar PS, Kumar CV (2021) IoT based Unified approach for Women safety alert using GSM. In: 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV). IEEE, pp 388–392
16. Tejesh BSS, Mohan Y, Kumar CA, Paul TP, Rishitha RS, Durga BP (2020) A smart women protection system using internet of things and open source technology. In: 2020 international conference on emerging trends in information technology and engineering (ic-ETITE). IEEE, pp 1–4

# Design of an Efficient IoT System for Poultry Farm Management


**G. Rajakumar, K. Lakshmi Narayanan, R. Santhana Krishnan, Y. Harold Robinson, M. Subramanian, and M. Asirvatham**

**Abstract** In the past, chicken meat was very negatively viewed by people. Now, however, very high nutritious chicken meat is consumed by people from all over the world. Environmental parameters such as temperature, humidity, and nutrition are very important in chicken growth. So, society should grow and monitor this grass or this grass at good conditions. The automatic monitoring system is created to overcome several issues that may arise on the chicken farm. Changes in the environment affect the egg production and feed consumption of the bird. Thus, the PEST system's major function is to supervise and manage the environmental conditions such as heat, humidity, oxygen, methane, and water that exist on the poultry farm. If any changes have occurred in the poultry farms' environmental condition, the system sends the relevant information to a mobile device like phones through telephone lines which reduces the human effort and time required. The interactive device has been created using the Arduino Uno microcontroller and various sensors, such as temperature, humidity, gas and water level sensors. As this system, the design takes on a very efficient way to farm poultry that will result in an increase in production and profit. It is high time to start thinking about solar technologies that can generate electricity. Although this technology is made for household, this has been successfully used in commercial farms for the energy saving of batteries and the opportunity to sell the power back to the grid.

G. Rajakumar (✉) · K. L. Narayanan
Department of ECE, Francis Xavier Engineering College, Tirunelveli, India
e-mail: gmanly12@gmail.com

K. L. Narayanan
e-mail: kyelyen@gmail.com; drkln86@gmail.com

R. S. Krishnan · M. Subramanian · M. Asirvatham
Department of ECE, SCAD College of Engineering and Technology, Tirunelveli, India
e-mail: santhanakrishnan86@gmail.com

M. Subramanian
e-mail: m.subramanian86@gmail.com; m.subramanian@gmail.com

M. Asirvatham
e-mail: email2asir@gmail.com

Y. H. Robinson
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India
e-mail: yhrobinphd@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_31

**Keywords** Arduino · Global system for mobile · Microcontroller · Poultry farm · Sensor

## 1 Introduction

The word "poultry" is a word that refers to all domesticated birds; such as chicken, duck, emu, geese, ostrich, pigeon etc. The poultry industry is one of the fastest growing sectors of the livestock industry and contributes a significant growth in 2017 of food supplements. Changing food habits, globalization, industrialization, rising income, urbanization had created a favorable atmosphere for development of poultry sector.India is the third largest producer of eggs and ranks fifth in the world in respect of poultry meat production. Rearing and breeding of hens is called poultry farming [1]. In Poultry Farm environmental parameters like temperature, humidity and gas should be monitored and controlled. The environmental condition should be maintained at the threshold value [2]. These Environmental parameters will affect the growth of the chicken and they are harmful for chickens in digestive, respiratory and behavioral change [3]. The smart chicken farm monitoring involves many stages such as management [4], feeding, planting, quality assurance, environment control, and performance. Chicken meat demand grows continuously due to its price competitiveness, high availability and good taste. Various dishes use chicken meat as its component. In order to supply the chicken meat demand, poultry industry must increase their productivity. Bad environment quality of poultry house will reduce the bird food consumption; increase its stress and death rate. To meet the chicken meat demand in market, environmental parameters of the chicken farm need to be monitored and maintained in order to produce healthy and wealthy chicken.

There is a high demand for agricultural products for food and for food processing industries. These include cassava farming, maize farming and poultry farming. In this research solutions are given for smart chicken farm monitoring which makes the chicken to grow healthier and wealthier. It maintains the farm record systematically and gives information in time. This model will also show the realistic execution of the apparatus which could be placed in the poultry farm for production and profit. Hence, a model has to be designed which will supervise and organize the environmental parameters such as temperature, humidity, gas level and water level in the poultry farm which will increase the growth of the chicken. Thereby, the layout of the project was presented and on which the work will be focused.

This system reduces the human spares in the poultry farm which in turn reduces the cost of poultry management and labor cost. It also monitors the environmental parameters in the chicken farm and transmits the data in real time to increase chicken growth and productivity. To design an Arduino based elegant structure for chicken farm monitoring in which environmental parameters such as temperature, humidity, gas and water level are measured and controlled in real time and the measured parameters are transmitted via wireless technology.

## 2 Related Works

In this segment, addressing the construction of the chicken farm and the investigation of other monitoring aspects. In creating the hybrid model for the management of mobile and wireless sensor networks of poultry farm, there is a system architecture implemented by Imtem et al. [5]. The farm was built so that it could work while on the go. Overall, it's important to think about humidity, temperature, light, and what is going on with our population density to understand how environmental factors can affect the system. Since the chicken are growing in the building, that reduces the dust and ammonia there. To prevent low growth rate, a humidity of 40% or lower must be preserved, and the temperature is at room temperature of 27 °C or below. Tadesse et al. [6] has suggested a few issues to consider, such as the effectiveness growth, the humidity should be regulated. The sum of heat, humidity, and ammonia, which are the three basic factors that birds are subjected to. If there are high levels of humidity and dust, then ammonia can be produced in the building. Once you've hit 30° or more, the humidity should be between 50 and 70%. Lu et al. [7] have clarified this monitoring method to be achieved with a heterogeneous wireless network of ultra-low power wireless sensor nodes. One way to appreciate the freshness of chicken is to use a technology that can sustain the temperature of them during the day. The R&D department comes up with a less-expensive version of it. Changed habits, grow exhausted and appears to grow older is a downside that makes our day-to-day lives more difficult. It also demonstrates that the node which is added to a chicken health monitoring system for detecting infected chicken with the highly pathogenic avian influenza (HPAI) viruses using clinical data at the early stage. Dong and Zhang [8] validated an online monitoring system to be used in a poultry farm based on ZigBee using a CC2430 controller integrated with different sensors such as temperature (TGS4161) and humidity (PWM4232), (SHT75). Besides, the internet of plants allows managers to track the temperature and light intensity of a farm from distant locations through web architecture.

Ayaz et al. [9] has developed a technique to improve a chicken farm's efficiency by being able to use a technology of wireless sensor networks. As part of a wireless sensor network and a fitness one, the chicken quality is considerably improved, which means these improvements are good for the health. This new device is based on a wireless sensor that can detect infected chickens. There are almost certain to boost overall demand, efficiency, and economy. The sensing techniques of wireless sensor networks and mobile networks are used to control and track poultry's environmental parameters. Siwakorn Jindarat and Pongpisitt Wuttidittachotti [10] built a device and a smartphone that can help poultry farms management. This device tracks the surrounding parameters critical for the poultry's welfare, such as humidity, temperature, and climate quality. Once in place, Poultry farms can be regulated by this device from anywhere, at any time. Wathes et al. [11] broke down four main pieces of knowledge for livestock farming. 1. Through monitoring, we can keep track of the poultry farm at a reasonable interval and provide information back to the Arduino controller. 2. A hypothesis, which is rather compact and which, through statistical

analysis, can measure the dynamic responses of each process output to variation of the inputs with the option of estimating, in real time, the offset of error. 3. Each process performance objective value and trajectory need to be defined. 4. The actuator was created to communicate with the model based predictive controller. Liu et al. [12] have identified a fire alarm system in which a real time monitoring system detects the presence of smoke in the air due to a fire, and captures the images by a camera installed inside the room on the off chance that a fire happens. It will have to a device that collects information from some of its peripherals, then manually conveys it by telephone after which it would cancel that telephone call after it talks with another telephone. The feature of the system that is most apparent is the ability to send a fire, smoke, or gas alarm, when there is a detected. If a mobile phone came with a smartphone text messaging application installed, that application would be able to regularly alert the customer during the program and make them recognize when they been smoking in a public place. If a smartphone text messaging application came with a smartphone application installed, that application would be able to regularly notify the consumer on the program and make them recognize when they been smoking in a public place. For those purposes, it is safer for this device to be able to work without any false alarms reported to the firefighter.

## 3 Methodology

The proposed system is formulated with the following modules; Monitoring module, Control module and Alert module. Monitoring module comprises of various sensors such as Temperature sensor (LM35), Humidity sensor (HR202), Gas sensor (MQ 7), Water level sensor, LDR sensor. Control module has Arduino UNO microcontroller; this microcontroller will monitor all the sensors based the input values given to them. GSM modem is used as a alert module, the GSM will send a message to the person who is in-charge to the poultry farm based upon the environmental parameters [13].

### *3.1 Monitoring and Control Module*

The poultry farm is monitored with various sensors such as Temperature sensor (LM35), Humidity sensor (HR202), Gas sensor (MQ 7), Water level sensor, LDR sensor along with Arduino UNO. Sensor analog input is given to the Arduino UNO pins $A_0$ to $A_6$. A threshold value is set for all the sensor observation for controlling the environmental parameters. The microcontroller will send a message to GSM module if the value received from the sensor is above the threshold level, cooling fan, exhaust fan, water sprinkler and motor pump will be switched ON automatically depending on the inputs received from the sensors.

### 3.2 Monitoring Temperature

The poultry farm temperature should be monitored so that the growth of the chicken will not be affected. Temperature sensor used in the proposed system is LM35. LM 35 measures the temperature more accurately compared to other sensor modules. The measuring temperature range is about "−55 to 150 °C"and consumes low power. The poultry farm should be monitored at room temperature of 28 °C. Three conditions are explained to maintain temperature in poultry farm: 1. If the temperature is between 29 to 31 °C, the cooling fan will be ON. Then the GSM modem will send a message "Cooling Fan is ON" to the person who is in-charge. 2. If temperature is above 31 °C, the Water Sprinkler will start to spray the water, so that the farm will be cool. Then the GSM modem will send a message "Water Sprinkler ON" to the person who is in-charge. 3. If the temperature below 28 °C, the GSM modem will send a message "Temperature is Low" to the person who is in-charge [14, 15]. If the temperature is above the room temperature chickens will consume more feed to sustain the body temperature. On controlling the temperature, chicken will consume more feed and better growth is achieved.

### 3.3 Monitoring Humidity

In poultry farm, the environmental parameter humidity is monitored. Humidity also affects the growth of the chicken. Humidity sensor used in the proposed system is HR202. This sensor detects the ambient humidity of the poultry farm. The operating voltage of this sensor is 5 V. The unit of Humidity is Relative Humidity (RH). The poultry farm should be monitored and controlled at 40 RH. If the humidity is above the threshold value (40 RH), the Exhaust fan will be ON. The GSM module will send a message "Fan ON" to the person who is in-charge. If the humidity is below the threshold value (40 RH), the Exhaust fan will be OFF. The GSM module will send a message "Fan OFF" to the person who is in-charge.

### 3.4 Monitoring Air Pollutants

Another Environmental parameter in poultry farm which affects the growth of chicken is emission of harmful gases. The Ammonia will release excessively when a small amount of water falls in the fecal matter which release the excess amount of heat concentration. Ammonia concentration should not be over 25 ppm. In the proposed system, the gas content is monitored using gas sensor MQ 7. MQ 7 is highly sensitive to ammonia gas. This sensor detects the ammonia gas concentration anywhere from 20 to 2000 ppm. The current consumption is about 150 mA. When the concentration of ammonia exceeds above 25 ppm, gas sensor will detect and the Exhaust fan will

be switched ON automatically and a message "Exhaust Fan ON" will be sent to a mobile via GSM modem. Ammonia fumes can be reduced by increasing ventilation and replacing litter. As the gas concentration affects the growth of chicken, it has to be monitored and controlled. This is necessary to monitor and control the ammonia gas because the gas concentration affects the growth of the chicken and leads to increase in mortality rate.

### *3.5 Artificial Lightning*

One more important parameter has to be monitored in the poultry farm is artificial lightning. Chickens are often kept often in 6 h of darkness provided each night. During light, they consume more food. If the intensity goes below 7 lx, lights will be switched ON automatically and a message "Lights ON" will be sent to a mobile [15].

### *3.6 Monitoring Water Level*

To monitor the water level, Magnetic float sensor is used. It is an Electromagnetic On/Off switch two metal contact. If magnet comes close to the contact, current passes into the circuit and the switch closes, motor is switched On. If magnet moves away from the contact, contact demagnetize and the switch opens and hence the motor is switched Off. In this, the float sensor is fixed in a tank at a certain level. If the water level goes below the threshold value, the magnet comes into contact, the current passes through it and the motor is switched On. The water is pumped into the tank. A message "Motor Pump On" will be sent to a mobile. This will reduce the human effort.

## 4 Results and Discussions

### *4.1 Hardware Module*

The hardware module consists of Arduino UNO Board, Temperature sensor, Humidity sensor, LDR sensor, Gas sensor, Floating switch, GSM modem, Cooling fan, Exhaust fan, Motor pump, White light, 4-channel relay and Power supply. Hardware setup is shown in Fig. 1.

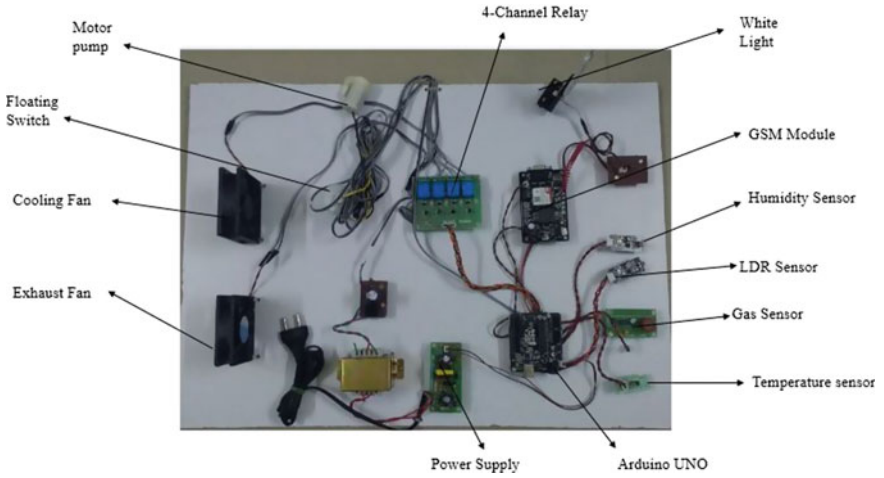



**Fig. 1** Hardware module

The hardware setup will be placed in the poultry house for achieving better growth of chicken. The poultry house will be automatically monitored and controlled which reduces the mortality rate and increases the production. The external 12 v power supply is specified to the Arduino microcontroller but it requires only 5 v power supply.

## *4.2 Monitoring Temperature*

Arduino microcontroller is programmed for three cases to monitor the temperature in the poultry farm for the better growth of chicken. In first case, if the temperature is between 29 to 31 °C, the message "Temperature high Fan ON" will be sent to a mobile via GSM modem, which is depicted in Fig. 2. If the temperature is above 31 °C, the water sprinkler will spray the water in the poultry farm to monitor the temperature and a message "Water spray on" will be sent to mobile. If the temperature is below 28 °C, this will not affect the growth of the chicken and a message "Temperature low Fan off" will be sent to a mobile via GSM modem, which is depicted in Fig. 3.

The sensor will be operated on 5 v. The output of LM 35 sensor is given to the $A_0$ pin of the Arduino board. The Arduino board is used to convey the output.

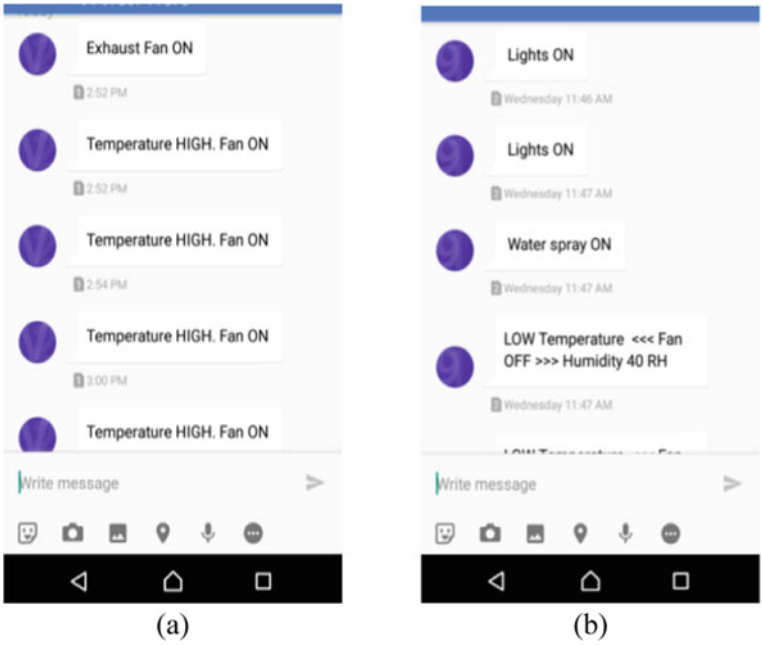



**Fig. 2** **a** Temperatures between 29 to 31 °C, and **b** Temperatures above 31 °C

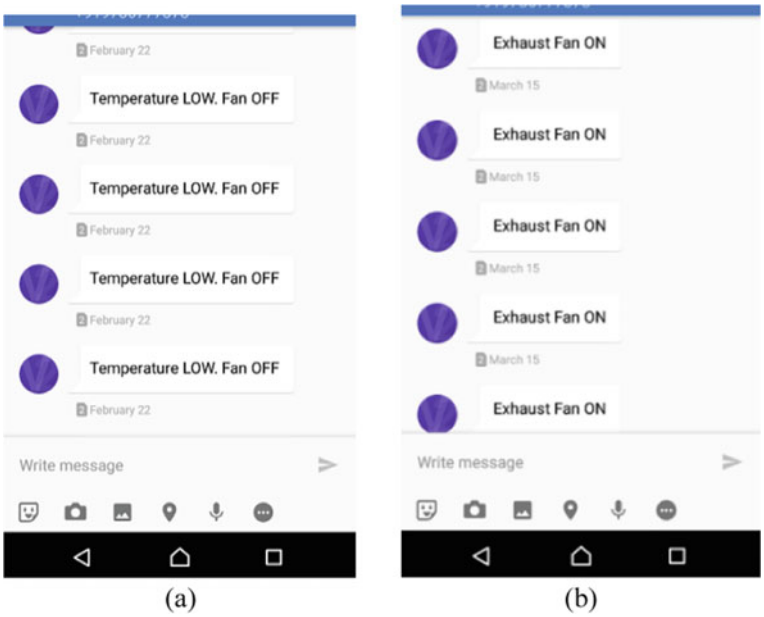



**Fig. 3** **a** Temperatures below 28 °C and **b** Exhaust fan is switched ON

### *4.3 Monitoring Gas Content*

To test the gas content, the MQ7 gas sensor is used. Increase in ammonia content will damage the respiratory track of the chicken and decrease in feed intake. Increase in ammonia also leads to mortality of the chicken. Ammonia is released in hot environment. The Ammonia will release excessively when a small amount of water falls in the fecal matter which release the excess amount of heat concentration. Ammonia concentration should not be over 25 pp. If ammonia gas exceeds above 25 ppm, the exhaust fan will be switched on and a message "Exhaust fan on" will be sent to a mobile. Ammonia gas content; the MQ7 gas sensor is used to sense the presence of ammonia gases. It will be operated on only 5 v. The output of gas sensor is given to the A2 pin of the Arduino board. The Arduino board is used to convey the output.

### *4.4 Monitoring Humidity*

HR202 humidity sensor is used in the proposed system which detects the ambient humidity in the poultry farm. The sensor will be operated on only 5 v. The output of humidity sensor is given the $A_1$ pin of the Arduino board. The Arduino board is used to convey the output. In the poultry farm humidity is monitored and controlled at 40 RH. If the humidity is above 40 RH, the Exhaust fan will be ON and a message "Exhaust fan on" will be sent to a mobile.

### *4.5 Artificial Lighting*

Chickens are often kept often in 6 h of darkness provided each night. During light, they consume more food. Light intensity should be kept below 10 lx to inhibit bird activity and increase feed efficiency. For artificial lighting, LDR sensor and white light is used in this system. If the intensity goes below 7 lx, lights will be switched on and a message "Lights on" will be sent to a mobile. LDR sensor operates only at 5 v. The output of LDR sensor is given to the $A_4$ pin of the Arduino board. The Arduino board is used to convey the output.

### *4.6 Monitoring Water Level*

Water level in the water tank is monitored with a magnetic floating switch. The float sensor is fixed in a tank at a certain level. If the water level goes below the threshold value, the motor is switched on and the water is pumped into the tank and a message

**Table 1** Comparison of temperature, LDR count, power consumption, and humidity for different designs

| LoRa designs | Temperature (in Celsius) | LDR count | Power consumption (uW) | Humidity (RH) |
|---|---|---|---|---|
| Asfaw et al. [1] | 35 | 45 | 258.63 | 44 |
| Zeng et al. [4] | 34 | 47 | 241.32 | 42 |
| Imtem et al. [5] | 39 | 63 | 247.3 | 59 |
| Liv et al.[12] | 42 | 98 | 221.4 | 62 |
| Proposed design | 29 | 31 | 178.25 | 37 |

"Water Pump Motor On" will be sent to a mobile. This water level monitoring will reduce the human effort. Float sensor operates only at 5 v. The output of float sensor is given to the $A_5$ pin of the Arduino board. The Arduino board is used to convey the output.

### *4.7 Comparative Analysis*

The Comparison of temperature, LDR count, power consumption, and humidity for different designs are given in Table 1. In proposed design, temperature, LDR count, power consumption, and humidity performances are improved than the conventional related works. When compared to the conventional designs, the proposed design worked effectively.

## 5 Conclusion

The manager of the poultry farm can obtain information on time and handle the poultry farm effectively. Here is a thorough account on how temperatures, humidity, gas content, water level, and artificial lightning of an industrial poultry farm is monitored and controlled automatically and the circuit is planned accordingly. The Global System for Mobile (GSM) communications modem is used to send a message to the mobile phone based on the environmental parameters such as temperature, humidity, gas content, water level, artificial lightning in the poultry, and it is monitored so that the better growth of the chicken is achieved and the mortality rate is reduced. The efficient management of the poultry, as a result, would contribute to good production and income. The statistical parameters are checked in the model, and the predicted value is within the predicted value range.

## References


1. Asfaw YT, Ameni G, Medhin G, Gumi B, Wieland B (2021) Poultry health services in Ethiopia: availability of diagnostic, clinical and vaccination Services. Poultry Sci 100:10102
2. Ananth Kumar T, John A, Ramesh Kumar C (2020) IoT technology and applications in Internet of Things, pp 43–62
3. Rajakumar G, Kumar TA, Samuel TA, Kumaran EM (2018) IoT based milk monitoring system for detection of milk adulteration. Int J Pure Appl Math 118(9):21–32
4. Zeng Z, Zeng F, Han X, Elkhouchlaa H, Yu Q, Lü E (2021) Real-time monitoring of environmental parameters in a commercial gestating sow house using a ZigBee-based wireless sensor network. Appl Sci 11(3):972
5. Imtem N, Sirisamphanwong C, Ketjoy N (2021) Development and performance testing of the automated building energy management system with IoT (ABEMS-IoT) case study: big-scale automobile factory. In: IT convergence and security. Springer, Singapore, pp 97–107
6. Tadesse T, Nichols MA, Hewett EW, Fisher KJ (2001) Relative humidity around the fruit influences the mineral composition and incidence of blossom-end rot in weet pepper fruit. J Hortic Sci Biotechnol 76(1):9–16
7. Lu J, Okada H, Itoh T, Maeda R, Harada T (2013) Towards the world smallest wireless sensor nodes with low power consumption for 'Green'sensor networks. In: Sensors. IEEE, pp 1–4
8. Dong F, Zhang N (2009) Wireless sensor networks applied on environmental monitoring in fowl farm. In: International conference on computer and computing technologies in agriculture. Springer, Heidelberg, pp 479–486
9. Ayaz M, Ammad-Uddin M, Sharif Z, Mansour A, Aggoune EHM (2019) Internet-of-Things (IoT)-based smart agriculture: toward making the fields talk. IEEE Access 7:129551–129583
10. Jindarat S, Wuttidittachotti P (2013) Smart farm monitoring using raspberry pi and Arduino. In: International conference on computer, communication, and control technology. https://doi.org/10.1109/I4CT.2015.7219582
11. Wathes CM, Kristensen HH, Aerts JM, Berckmans D (2008) Is precision livestock farming an engineer's daydream or nightmare, an animal's friend or foe, and a farmer's panacea or pitfall? Comput Electron Agric 64(1):2–10
12. Liu W, Wang X, Song Y, Cao R, Wang L, Yan Z, Shan G (2020) Self-powered forest fire alarm system based on impedance matching effect between triboelectric nanogenerator and thermosensitive sensor. Nano Energy 73:104843
13. Behr CJ, Kumar A, Hancke GP (2016) A smart helmet for air quality and hazardous event detection for the mining industry. In: 2016 IEEE International Conference on Industrial Technology (ICIT). IEEE, pp 2026–2031
14. Jiang JA, Tseng CL, Lu FM, Yang EC, Wu ZS, Chen CP, Liao CS (2008) A GSM-based remote wireless automatic monitoring system for field information: a case study for ecological monitoring of the oriental fruit fly, Bactrocera dorsalis (Hendel). Comput Electron Agric 62(2):243–259
15. Varghese VT, Sasidhar K, Rekha P (2015) A status quo of WSN systems for agriculture. In: 2015 international conference on advances in computing, communications and informatics (ICACCI). IEEE, pp 1775–1781

# Design of LORA Based Tracking System for Military Personnel


**G. P. Ramesh and Neha**

**Abstract** In this paper, an efficient tracking system is proposed that is used to track the location with the help of a LORA based embedded system connected to a real-time android application. The embedded module is used which incorporates GPS, LoRa-Wan and Cloud modules. On a Smartphone, an android based application is developed to plot the location on a Google map. LoRa is a wireless modulation technique derived from Chirp Spread Spectrum (CSS) technology. It encodes information on radio waves using chirp pulses - similar to the way dolphins and bats communicate! LoRa modulated transmission is robust against disturbances and can be received across great distances. The encoded data helps to secure the data in the communication channel. The tracking system provides the major application in the military personnel. Due to the usage of the effective communication channel, Experimental results have proved that the proposed tracking system is accurate and feasible. The spectrum range of the proposed work is 865 to 867 MHz which helps to coverage the maximum distance. With the help of AES encryption, the transmission of the date is progressed effectively with 10 mW energy consumption. Spread spectrum modulation helps to modulate the information by operating the AES design with 64-byte payload size.

**Keywords** AES encryption · Chirp Spread Spectrum · Cloud · Embedded System · Radio wave · LoRa WAN

## 1 Introduction

The unprecedented technological advancement has improved the lifestyles of human beings and devised the number of innovative technologies for object identifications. In 1960, Ernst F. W Alexanderson introduced the first RFID tag to the globe and incessantly the Sensorimatic and Checkpoint organization launched the first RFID anti-theft tracking application for high-value products in retail stores [1–3]. Later on, the RFID standards are being steadily raised to strengthen the use of RFID and make it

G. P. Ramesh (✉) · Neha
St. Peters Institute of Higher Education and Research, Chennai, India
e-mail: rameshgp@yahoo.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_32

a part of everyday life in the twenty-first century. RFID is currently leveraged in many industries such as production, health-care, transportation and tracking applications that captures human life in the direction of indoor smart living. Recent evidence shows that human security threats continue to rise in developing countries especially towards vulnerable members of society that causes great distress to them and to their families [4–6]. The mechanisms used in the above RFID applications to embed RFID tags with participants were not expedient and potential limitations were reported, including embedded-cloths tags, bracelets, student bags, cards, and keychains [7, 8]. They have been often feared to be miss-placed, used by prohibited users or get damaged. As such, model employed for this study is a sophisticated RFID-based human tracking system that utilizes a micro-flexible tag mechanism implanted into skins which is consistent and leaves no possibility for removal without the destruction of tag. The main goal of this study is to develop a safer commuting and monitoring system for the vulnerable members of society. The lab demo of this model has been programmed, and tested virtually as such tags were allocated to participants and tracked by RFID readers for monitoring purposes at the unified center. The error probability performance of the coded and uncoded LoRa system can be evaluated by the quasi orthogonal performances [9]. Moreover, the SDR implementation of LoRa is implemented based on the over Rayleigh channel [10].

## 2 Related Work

Zorbas et al. [11] presented Time slotted LoRa WAN for the industrial Internet of Things. In this work, the time slotted approach plays a vital role for the time synchronization. This device was performed in low latency and less cost which helped to improve the overall operation. But, the power consumption of the model was high for small module. Aftab et al. [12] proposed scalability analysis of multiple LoRa Gateways using stochastic geometry. The scalability of the performance is connected in the LPWAN technology while keeping the network structure. However, the outlet of the gate way was difficult to design. Hoang et al. [13] presented the slotted transmission with Collision avoidance for LoRa networks. The network performed for long range communication with low power consumption and less delay. But, it suffered from data Collision and signal suppression. Florita et al. [14] proposed LoRa based gateways for delay tolerant sensor data collection in urban settings. The choice of radio technology for the link between the sensor or end device and the mobile gateway was important: it would affect the number of mobile gateways needed, the required mobility pattern for them to provide a certain level of delay guarantees to the data, and the overall efficiency of the scheme. But, this proposed work delivered the more delay while utilizing the multiple gateways. Sciullo et al. [15] proposed LoRa based mobile emergency management system which used to identify the location of the users for further communication. In this work, Emergency Communication System (ECS) based on short-range Device-to-Device (D2D) solutions are proposed which was available on Commercial Off The Shelf (COTS) devices (e.g. Wi-Fi

Direct); however, the target of these solutions is constituted by small indoor areas, since the scalability on large-scale environments is often a problem.

## 3 Tracker Design

The overall implementation of the system consists of three parts- transmitter (a), receiver (b) and the real-time application used for tracking the location. The design and implementation of hardware and software architecture for each module are explained in the below. The transmitter (a) module configuration consists of GPS, microcontroller, external power source and a LoRa transmitter. The receiver (b) module consists of a LoRa receiver to receive signal transmitted, which is further connected to a microcontroller to transmit the data/location to the mobile application through the Wi-Fi. The receiver module is also connected to a external power source. The real-time application is connected to cloud for the GPS value and the application is integrated with google maps to provide the precise location. The transmitter and receiver modules are shown in Figs. 1 and 2 respectively. The internal mobile application development architecture is shown in Fig. 3 which contains the cloud and google map.

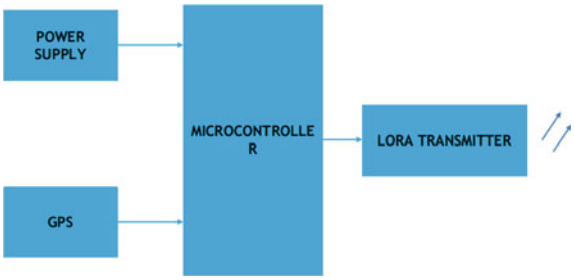



**Fig. 1** Transmitter module

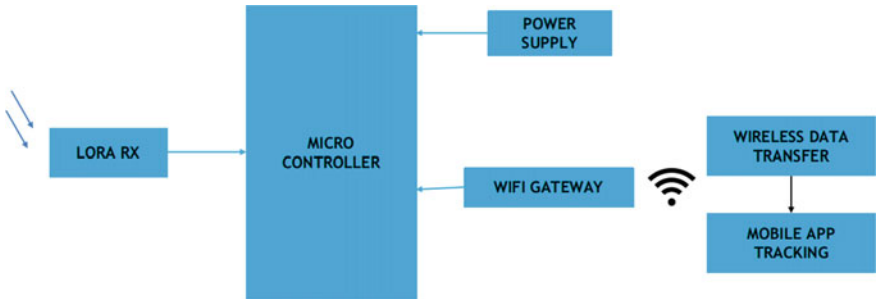



**Fig. 2** Receiver module

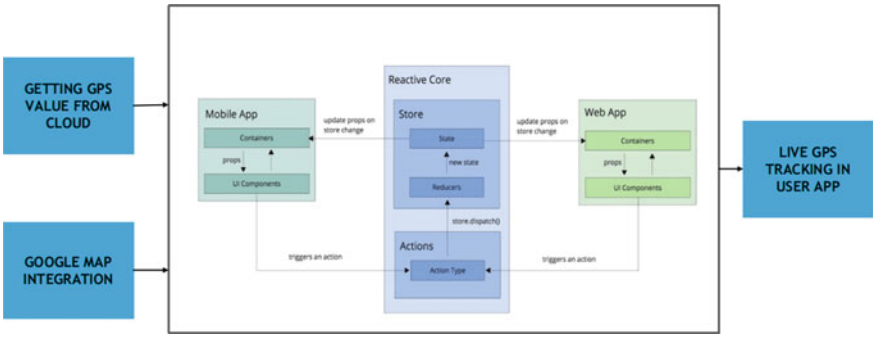



**Fig. 3** Internal mobile application development architecture

## 4 Experimental Results

The below image shows the results of the system design and developed for various modules.

### *4.1 Hardware and Simulation*

The Fig. 4 shows the complete hardware setup of the transmitter of the project:

The Fig. 5 shows the LoRa receiver connected to the USB port of the computer to receive the GPS co-ordinates on the terminal com software to check the transmission between the LoRa modules setup of the Receiver of the Research:

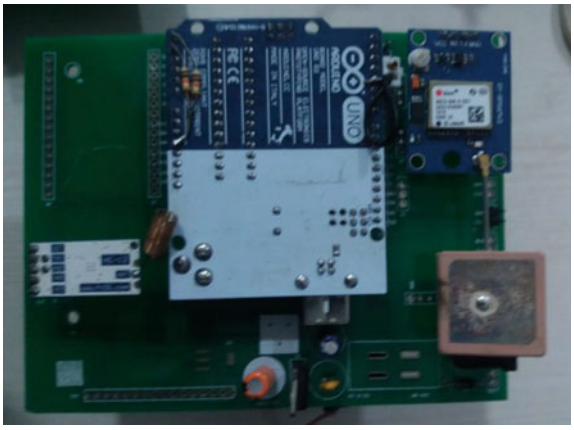


**Fig. 4** Hardware setup

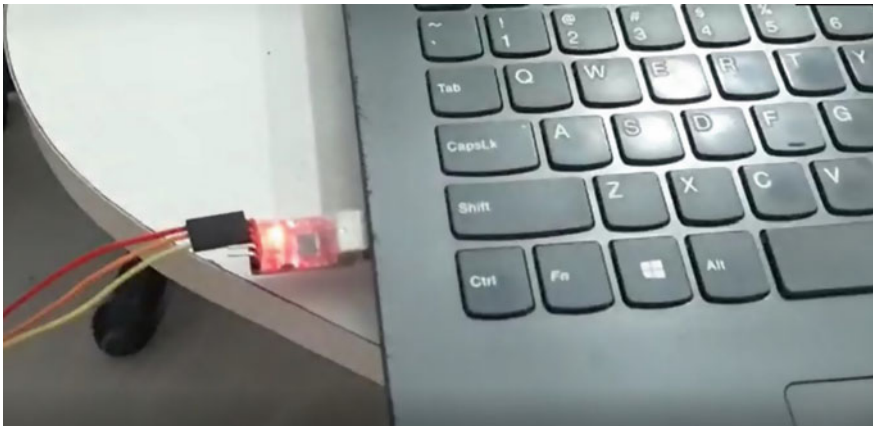



**Fig. 5** LoRa receiver connected to the USB port of the computer to receive the GPS co-ordinates

**Fig. 6** GPS co-ordinates of the LoRa transmitter module on the terminal port

The Fig. 6 shows the GPS co-ordinates of the LoRa transmitter module on the terminal port:

## *4.2 Application and User Interface*

The Fig. 7 shows the user interface of the Research and Fig. 8 shows the GPS identity.

The system was tested and it was found that the location was predicted precisely in most of the cases. However, the location shown in the app can have an error of approximately 10 m due to limitation in the hardware.

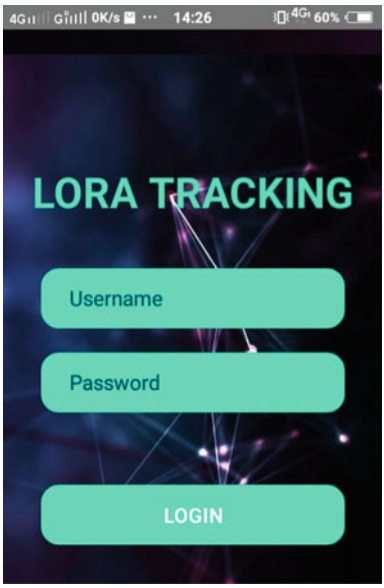



**Fig. 7** User interface

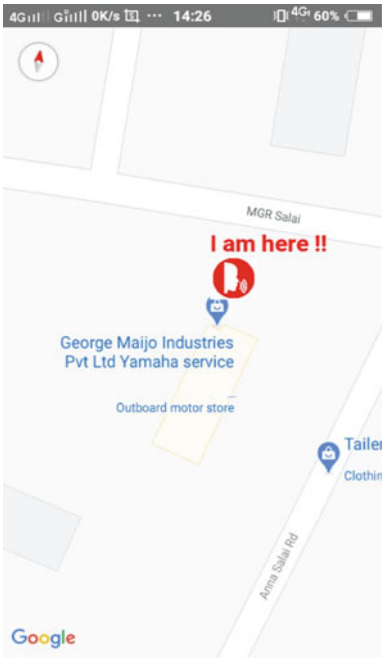



**Fig. 8** GPS identification

**Table 1** Comparison between LoRaWAN and other technologies

| Parameters | Zigbee | LoRaWAN | Bluetooth |
|---|---|---|---|
| Spectrum | 2.4 GHz | 865 to 867 MHz | 2402 MHz |
| Range | 10 to 100 m | 15 to 20 km | Less than 10 m |
| Modulation | Direct sequence spread spectrum (DSSS) | Spread spectrum modulation (SSM) | Gaussian frequency shift keying, (GFSK) |
| Payload size | 104 bytes | 64 bytes | 251 bytes |
| Security | 128-bit AES | AES | AES |
| Energy consumption | 10 to 100 milli watts (mW) | 10 milli watts | 5 milli watts |

**Table 2** Baud rate and serial port baud rate

| Serial port baud rate | 1,200bps | 2,400bps | 4,800bps | 9,600bps | 19,200bps | 38,400bps | 57,600bps | 115,200bps |
|---|---|---|---|---|---|---|---|---|
| Baud rate in the air | 5,000bps | | 15,000bps | | 58,000bps | | 236,000bps | |

## 5 Specification Analysis

### *5.1 The Module Used in the Project has the Following Specification*

In this mode, the module can automatically adjust the baud rate of wireless transmission in the air according to serial port baud rate, and the corresponding relationship is as shown in the table below: The performance comparison of the proposed work with other technology is given in Table 1. The baud rate information of the proposed LoRaWAN is given in Tables 2 and 3. Loopback time for different modes with remarks information is given in Table 4. The highlighted part is used in the project.

**Table 3** Baud rate in the air and sensitivity

| Baud rate in the air | 5,000bps | 15,000bps | 58,000bps | 236,000bps |
|---|---|---|---|---|
| Wireless receiving sensitivity | -117dBm | -112dBm | -107dBm | -100dBm |

**Table 4** Loopback time for different modes with remarks

| Mode | FU1 | FU2 | FU3 | Remark |
|---|---|---|---|---|
| Idle current | 3.6mA | 80μA | 16mA | Average value |
| Transmission time delay | 15-25mS | 500mS | 4-80mS | Sending one byte |
| Loopback test time delay 1 | 31mS | | | Serial port baud rate 9,600, sending one byte |
| Loopback test time delay 2 | 31mS | | | Serial port baud rate 9,600, sending ten bytes |

**Table 5** Comparison of spectrum, energy consumption, frequency, and bandwidth for different LoRa designs

| LoRa designs | Spectrum (MHz) | Energy consumption (mW) | Frequency (GHz) | Bandwidth (KHz) |
|---|---|---|---|---|
| Zorbas et al. [11] | 412 | 70 | 89.36 | 125 |
| Aftab et al. [12] | 514 | 68 | 147.9 | 145 |
| Hoang et al. [13] | 325 | 95 | 100.6 | 198 |
| Florita et al.[14] | 147 | 47 | 114.2 | 114 |
| Proposed LoRa | 865 | 10 | 159.4 | 269 |

The receiving sensitivity of module at different baud rates in the air is as shown in the table:

The following gives some characteristics reference values of various modes:

### *5.2 Comparative Analysis*

The Comparison of spectrum, energy consumption, frequency, and bandwidth for different LoRa designs are given in Table 5. In proposed LoRa, the energy consumption, bandwidth and all the performances are improved than the conventional related works [11–14]. When compared to the conventional LoRa designs, the proposed LoRa worked effectively to identify the GPS detection.

## 6 Conclusion

This project is used for safeguarding and monitoring the military personnel by effectively tracking them. This also help a person to view the current location of

the property through a mobile application developed using react native. So, this project reduces the number of cases of loss in property, thus saving a lot of money. The proposed work spectrum, energy consumption, frequency, and bandwidth are improved a lot compared to the conventional model. The baud rate and wireless connectivity helps to improve the system bandwidth and operating frequency. For secure the data, AES encryption algorithm is used in this research work. The hardware setup also implemented and verified with the mobile application.

## References


1. Su B, Qin Z, Ni Q (2020) Energy Efficient Uplink Transmissions in LoRa Networks. IEEE
2. Bernier C, Dehmas F, Deparis N (2020) Low complexity LoRa frame synchronization for ultra-low power software-defined radios. IEEE Trans Commun 68(5):3140–3152
3. Croce D, Gucciardo M, Mangione S, Santaromita G, Tinnirello I (2019) LoRa technology demystified: From link behavior to cell-level performance. IEEE Trans Wirel Commun 19(2):822–834
4. Gimenez F, Carlos Z, Guillermo R (2020) Extending SMS service coverage in rural areas by using LoRa communication technology. IEEE Latin Am Trans 18(02):214–222
5. Premsankar G, Ghaddar B, Slabicki M, Di Francesco M (2020) Optimal configuration of LoRa networks in smart cities. IEEE Trans Ind Inf 16(12):7243–7254
6. Colavolpe G, Foggi T, Ricciulli M, Zanettini Y, Mediano-Alameda J-P (2019) Reception of LoRa signals from LEO satellites. IEEE Trans Aeros Electron Syst 55(6):3587–3602
7. Sherazi HHR, Grieco LA, Imran MA, Boggia G (2020) Energy-efficient LoRaWAN for industry 4.0 applications. IEEE Trans Ind Inf 17(2):891–902
8. Lam Ka-Ho, Cheung Chi-Chung, Lee Wah-Ching (2019) RSSI-based LoRa localization systems for large-scale indoor and outdoor environments. IEEE Trans Veh Technol 68(12):11778–11791
9. Baruffa G, Rugini L, Germani L, Frescura F (2020) Error probability performance of chirp modulation in uncoded and coded LoRa systems. Dig Signal Process 106:102828
10. Marquet A, Montavont N, Papadopoulos GZ (2020) Towards an SDR implementation of LoRa: Reverse-engineering, demodulation strategies and assessment over Rayleigh channel. Comput Commun 153:595–605
11. Zorbas D, Abdelfadeel K, Kotzanikolaou P, Pesch D (2020) TS-LoRa: time-slotted LoRaWAN for the industrial Internet of Things. Comput Commun 153:1–10
12. Aftab N, Zaidi SAR, McLernon D (2020) Scalability analysis of multiple LoRa gateways using stochastic geometry. Internet of Things 9:100132
13. Hoang QL, Tran HP, Jung WS, Hoang SH, Oh H (2020) A slotted transmission with collision avoidance for LoRa networks. Procedia Comput Sci 177:94–101
14. Florita NJB, Senatin ANM, Zabala AMA, Tan WM (2020) Opportunistic LoRa-based gateways for delay-tolerant sensor data collection in urban settings. Comput Commun 154:410–432
15. Sciullo L, Trotta A, Di Felice M (2020) Design and performance evaluation of a LoRa-based mobile emergency management system (LOCATE). Ad Hoc Netw 96:101993

# Design of Quantum Encoders with Minimum Area Overhead


**M. Navaneetha Velammal, P. Hannah Blessy, J. Friska, and A. Rajeshwari**

**Abstract** Code converter is widely used for security purposes to protect sensitive data from spies. Code converter transforms data information into any coded form. Code converter is also useful for increasing the controllability and transportability of the data. There is a different way of implementing a code converter in a digital circuit. One way to implement a code converter using Complementary Metal Oxide Semiconductor (CMOS) technology. For reducing the area size, and increasing the power consumption, this paper presents the alternative technologies of CMOS that is Quantum dot Cellular Automata (QCA) nanotechnology. In recent times, QCA is an emerging technology that significantly overcomes the problems in CMOS technology. QCA is small size, low power consumption, and high speed compared to CMOS technology. In this research paper, logic gate-based efficient 4-bit QCA binary to gray conversion nanotechnology is proposed. The proposed technology is validated using QCA designer version 2.0.3. In the resulting phase, the performance is analyzed and verified for determining the capabilities of the quantum based converter.

**Keywords** Binary to gray · Cell count · Code converter · Complementary metal oxide semiconductor · Quantum dot cellular automata

## 1 Introduction

Code converter is used to prevent important data from third parties. Code converters are mostly used in the military for security purposes [4, 5]. It increases the data flexibility. In this paper logic gate based efficient 4-bit QCA based encoding is designed. When designing a code converter by using CMOS technologies, it increases the area size, power consumption and it cannot be able to extend its size up to nanoscale. Sometimes overlapping may occur due to the complex design of the circuit. This paper [10] present a data conversion B2G using CMOS technologies. Total 69.4 $\mu m^2$ area and 18 number of MOS devices are required for this method. To reduce the

M. N. Velammal (✉) · P. H. Blessy · J. Friska · A. Rajeshwari
Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli 627003, India
e-mail: navaneethavelammal.m@francisxavier.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_33

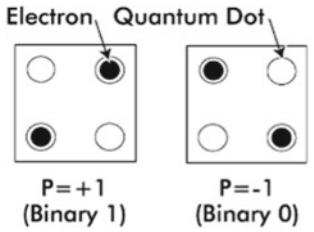



**Fig. 1** QCA charge configuration

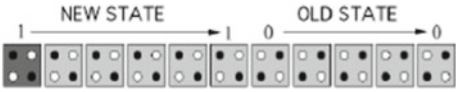



**Fig. 2** QCA wire

circuit complexity and overcome the limitation of CMOS, this paper presents a new latest technology that is used mostly in the recent time research field is called QCA nanotechnology. Area required for design proposed in QCA designer is 0.04 $\mu m^2$. So, area improvement in QCA is 1750 times that CMOS transistor design. QCA is a widely used method to design a digital circuit in nanotechnology [9]. This method is an alternative to CMOS with many additional features and helps to improve the efficiency of the circuit's design. The fundamental elements of QCA are QCA cell. In each corner of the QCA cell, quantum dots are located. Totally 4 quantum dots in each QCA cell. Every quantum cell contains added number of few electrons to represent their binary states in terms of charge configuration in QCA.QCA encodes binary information as a charge configuration. It has two polarization states to represent binary states. If $p = +1$, it represents binary state1. If $p = -1$, then it represents binary state 0, as stated in Fig. 1.

A large number of adjacent quantum cells are arranged in series is called QCA wire. Adjacent quantum cells which form a QCA wire and create a channel for transfer the data based on Columbic interaction. Four different clock phases are used in QCA. The QCA wire and clocking phase are depicted in the Fig. 2 and 3.

The majority voter and inverter are the two major tools for implementing a logic function in QCA technology. There are different types of majority voters based on their inputs. Inverters are used to invert the input of the signals. It is very useful for implementing a logic gate [11, 12].

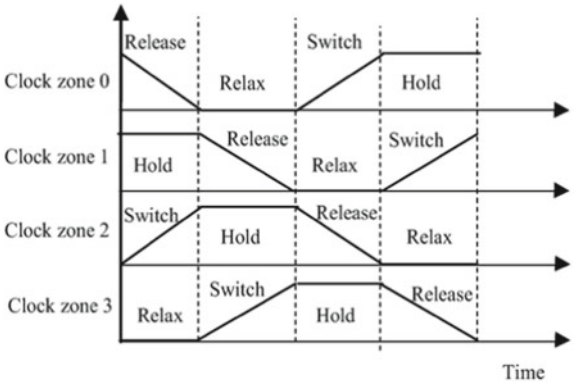



**Fig. 3** QCA Clocking phase

## 2 Related Works

Majeed et al. [7] developed an encoder by using major tools of logic function in QCA. In this paper majority voters and inverter are used to design. Maximum 65% of cells are used for implementing a 3-bit encoding. Similarly, 56% of cells are used for implementing a 4-bit encoding. Guleria et al. [6] developed a multilayer crossover-based 4-bit encoding in QCA technology. Tripathi et al. [1] shows that there are 9 majority voter are used to build OR gates, AND gates and an inverter in this design. For implementing a proposed design required area is 0.04 $\mu m^2$. Ramesh et al. [8] implemented a code converter in 2-bit, 3-bit, and 4-bits. In this study, logic gates based majority voter are used for implementation. There are 2 AND gates, 1 OR gate, and an inverter is used to implement a single EX-OR gate. Jeyalakshmi et al. [3] have proposed an enhanced both two type of code converter encoder. Khakpour et al. [2] presents the reversible logic is used to give a better solution for warmth dissipation in a digital circuit. This paper implements a 3-bit code converter up to a 5-bit code converter.

## 3 Methods

Code converters are combinational circuits that convert the data information and execute arithmetic processes. Logic XOR gates are major tools for designing a code converter. The Boolean equation for the encoding scheme is simplified by using Karnaugh maps and it shows that three gray outputs of B2G are the XOR gate combination of the present and previous binary inputs. In this paper, the 4-bit encoding scheme is proposed to use for coding techniques which will be used for communication applications [13]. Majority voter and inverter are not used for implementing a proposed XOR logic level gate, which is stated in Fig. 4.

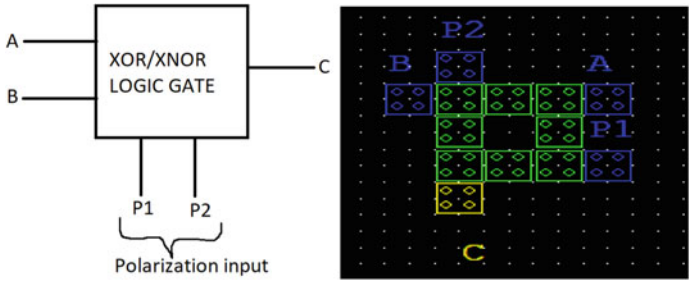



**Fig. 4** Schematic diagram and Layout diagram of special gates

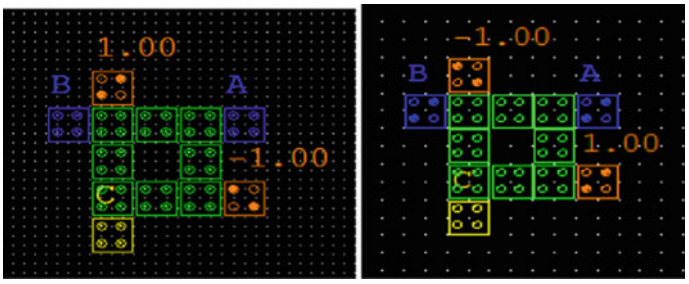



**Fig. 5** Proposed XOR and XNOR logic level gate in QCA

The existing XOR model is based on majority voter and inverter. Nearly three majority voter are used for the single XOR model. But in the proposed XOR model two polarization enable inputs are used instead of using majority voter. These two polarized inputs create a single structure or single layout for the XOR gates. A and B are the input of XOR/XNOR logic level gate [14, 15]. P1 and P2 are the polarized inputs of XOR/XNOR logic level, and C is the output. Based on the polarization value of two polarized inputs P21 and P2, proposed XOR logic level gates are designed. Assign the polarization value of polarized inputs P1 $= -1$ and P2 $= +1$ for operating the XOR logic operation in this proposed layout. Similarly assign the polarization value of polarized inputs P1 $= +1$ and P2 $= -1$ for operating the XNOR logic operation, as stated in Fig. 5.

There are a total of 3 numbers of proposed XOR logic level gates are used for designing a single code converter. Mathematical expression is used to represent 4-bit QCA encoding [16, 17].

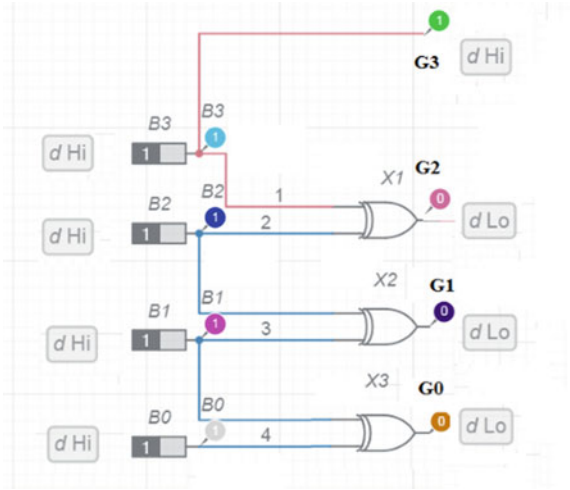



**Fig. 6** Schematic diagram of basic encoding circuitry

$$\text{G3 is expressed as B3} \tag{1}$$

$$G2 \text{ is equivalent to logical ex} - \text{or of B3 and B2} \tag{2}$$

$$\text{G1 is logically expressed as the ex} - \text{or operation of B2 and B1} \tag{3}$$

$$\text{G0 is equivalent to logical ex} - \text{or of B1 and B0} \tag{4}$$

Figure 6 shows that schematic diagram of basic encoding circuitry for optimized Boolean expression [18]. Boolean expression statement for 4-bit B2G encoding as shown in Eq. 1 to 4 and lookup table for 4-bit B2G encoding circuitry as follow.

Table 1 and Fig. 7 shows the layout of the QCA based encoding circuitry. The green colour quantum cells in this diagram are representing normal cells. Blue colour quantum cells are representing input cells. Yellow colour quantum cells are representing output cells [19, 20]. Finally, orange colour quantum cells are representing polarization cells.

**Table 1** Lookup Table for 4-bit B2G encoding circuitry

| Binary inputs | | | | Gray outputs | | | |
|---|---|---|---|---|---|---|---|
| B3 | B2 | B1 | B0 | G3 | G2 | G1 | G0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 |
| 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |

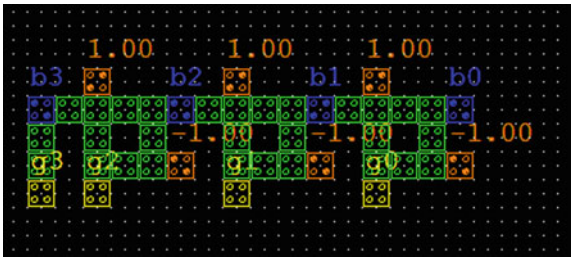



**Fig. 7** Layout of proposed encoding circuitry in QCA

## 4 Results and Discussion

The proposed QCA encoding circuitry is designed and simulated using QCAD. QCA Designer software is used for the design. Then assign the inputs and outputs for the design code converter. For assigning inputs and outputs click the select button in the left corner of the software. Figure 8 shows that inputs/outputs assigning. After assigning code converter in software then go to the left corner of the window click the simulation and select the simulation engine setup button. The bistable approximation window is open automatically. In this bistable approximation, the window selects the options button. Bistable options are open. In this option, there is so many default

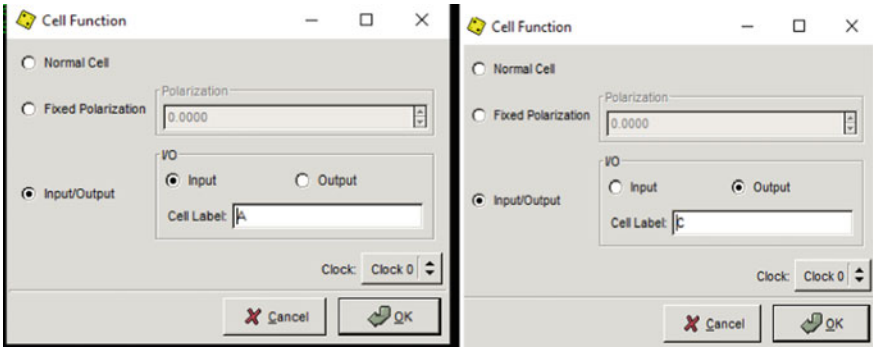



**Fig. 8** Assigning inputs and outputs of code converter

values present. In that select the animate option. The representation of bistable options window is shown in Fig. 9.

After the simulation engine setup is finished. Then select the simulation options and click the simulation vector setup. The input vector setup table is open automatically. Assign input value to the proposed design in this vector table setup based on encoding circuit truth table. The input vector table setup is shown in Fig. 10.

When binary input $B_0$ is alone high, then their corresponding gray output $G_0$ is high and remaining gray outputs are low. When binary input $B_1$ is alone high, then their corresponding gray outputs $G_0$ and $G_1$ is high and remaining gray outputs are low. When two binary inputs $B_0$ and $B_1$ are high, then their corresponding gray output $G_1$ is high and remaining gray outputs are low. The response for the given input is observed in Fig. 11.

The comparative Table 2 enumerate XOR gate. In existing XOR gate structure cell count is 51 cells and area is 0.08 $\mu m^2$. But in proposed structure cell count for

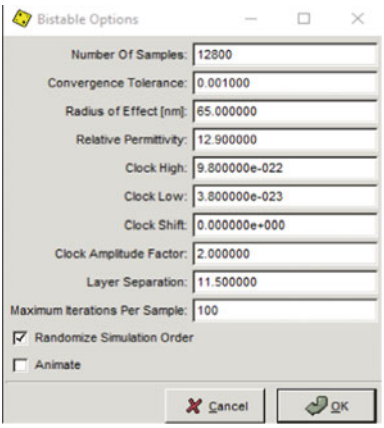



**Fig. 9** Bistable options window

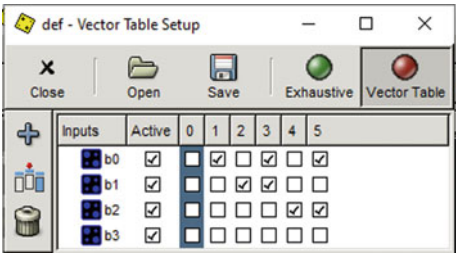



**Fig. 10** Input vector table setup

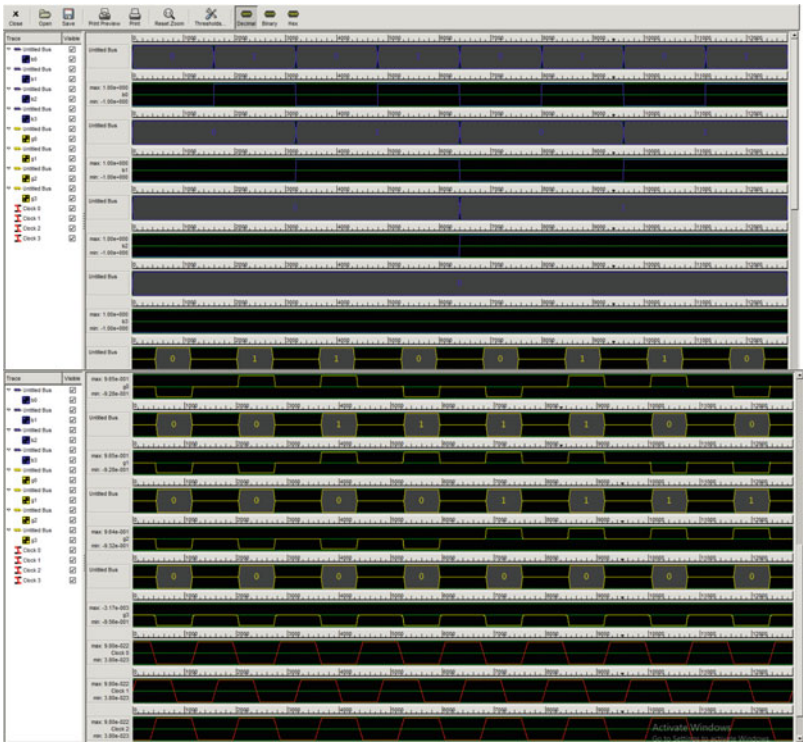


**Fig. 11** Encoding circuitry simulation response for the input pattern

**Table 2** Comparative table for XOR gate and encoding circuitry

| Ex-Or gate | Existing structure [1] | Proposed structure |
|---|---|---|
| Cell count | 51 cells | 13 cells |
| Area | 0.08 $\mu m^2$ | 0.02 $\mu m^2$ |
| **4-bit encoding circuitry** | **Existing structure** [2] | **Proposed structure** |
| Cell count | 225 cells | 43 cells |
| Area | 0.43 $\mu m^2$ | 0.04 $\mu m^2$ |

**Table 3** Comparison between XOR gate and Code Converter for proposed structure

| Proposed structure | | | |
|---|---|---|---|
| QCA structure | Cell count | Area $\mu m^2$ | Latency (Clocking cycle) |
| Ex-Or Gate | 13 cells | 0.02 $\mu m^2$ | 1 |
| 4-bit encoding circuitry | 43 cells | 0.04 $\mu m^2$ | 1 |

XOR gate is 13 cells and area is 0.02 $\mu m^2$. The Table 2 shows that the comparison between existing structure and proposed structure for code converter. In existing code converter structure cell count is 225 cells and area is 0.43 $\mu m^2$. But in proposed structure cell count for code converter is 43 cells and area is 0.04 $\mu m^2$. In addition to this, Table 3 indicates the comparison between XOR gate and code converter for proposed structure.

## 5 Conclusion

This paper first proposed XOR logic level gate and then 4-bit code converter is proposed based on XOR logic level gate in QCA technology. The proposed 4-bit QCA encoding circuitry has been occupied the space of 0.004 $\mu m^2$ and the latency of one clock cycle. It required 13 numbers of cell count to design. Proposed code converter is very smaller in size compare to existing method. Due to lower latency in proposed method cell count is less. This work has been proved efficient programmable circuit. Using this proposed XOR logic level gate, we will design more complex circuits in future.

## References

1. Tripathi D, Wairya S (2020) Energy efficient code converter for nanotechnology applications. J Crit Rev 7(13): 6006–6015. ISSN 2394–5125
2. Khakpour M, Gholami M, Naghizadeh S (2020) Parity generator and digital code converter in QCA nanotechnology. Int Nano Lett 10:49–59
3. Jeyalakshmi M, Muthulakshmi D, Shri VK, Shahana MM (2019) Enhanced design of binary to gray and gray to binary code converters using Quantum Dot Cellular Automata (QCA). Int J Sci Res Rev 8(13).ISSN NO: 2279–543X
4. Islam MD, Jahan GS, Bahar AN, Ahmed K, Abdullah-Al-Shafi M (2018) A new efficient non-reversible 4 bit binary to gray and 4 bit gray to binary converter in QCA. Nanosyst. Phys. Chem. Math. 9(4):473–483
5. Soni P, Agarwal VK (2018) A review of binary to gray and gray to binarycode conversion. Int J Innov Sci Res Technol 3(10). ISSN No: 2456–2165
6. Guleria N (2017) Binary to gray code converter implementation using QCA. 978-15090-6403-8/17/$31.00 © 2017 IEEE
7. Majeed AH (2017) A novel design "Binary to Gray Converter" with QCA nanotechnology. Int J Adv Eng Res Dev 4(9). p-ISSN (P): 2348–6406

8. Chaithanya P, Ramesh B, Srujana V, Vikram V (2017) Implementation of efficient binary to gray code converter using quantum dot cellular automata. Int J Res Appl Sci Eng Technol 5(6). ISSN: 2321–9653
9. Islam S, Abdullah-al-shafi M, Bahar AN (2015) Implementation of binary to gray code converters in quantum dot cellular automata. J Today's Ideas Tomorrow's Technol 3(2):145–160
10. Anand RK (2015) High performance binary to gray code converter using transmission gate. Int J Electr Electron Eng 2(1). p-ISSN: 1694–2426
11. Ahmad F, Bhat GM (2014) Novel code converters based on quantum-dot cellular automata (QCA). Int J Sci Res (IJSR) 3(5), ISSN (Online): 2319–7064, Paper ID: 020131715
12. Jahan WS, Ahmad PZ, Peer MA, Khan KA (2014) Circuit nanotechnology: QCA adder gate layout designs. IOSR J Comput Eng 16(2):70–78
13. Beigh MR, Mustafa M (2013) Design and implementations of quantum-dot cellular automatabase novel Parity generator and checker circuits with minimum cell complexity and cell count. Indian J Pure Appl Phys 51:60–66
14. Beigh MR, Mustafa M, Ahmad F (2013) Performance evaluation of efficient XOR structuresin quantum-dot cellular automata (QCA). Circ Syst 4:147–156
15. Cho H, Swartzlander EE (2007) Adder designs and analyses for quantum-dot cellular automata. IEEE Trans Nanotechnol 6(3):374–383
16. Askari M, Taghizadeh M, Fardad K, "Design and Analysis of a Sequential Ring Counter for QCA Implementation," International Conference on Computer and Communication Engineering, pp. 933–936, (2008).
17. Askari, M., M. Taghizadeh, Fardad K (2008) Digital design using quantum-dot cellular automata (a nanotechnology method). In: Proceedings of international conference on computer and communication engineering, pp 952–955
18. Tougaw PD, Lent CS (1994) Logical devices implemented using quantum cellar Automata. J Appl Phys 75(3):1818–1825
19. Lent CS, Tougaw PD (1993) Lines of interacting quantumdot cells: a binary wire. J Appl Phys 74:6227–6233
20. Walus K, Julien GA (2006) Design tools for an emerging SoC technology: quantum-dot cellular automata. Proc. IEEE 96:1225–1244

# Design of Smart Voice Enabled Walking Stick for Visually Impaired


**R. Santhana Krishnan, V. Nagaraju, N. Sasikala, K. Lakshmi Narayanan, S. Sundararajan, and Y. Harold Robinson**

**Abstract** India contributes to about 22% of the world visually impaired population. These people face many difficulties while traveling from one place to another. They can't able to detect the hindrance found in their way while walking. This may lead to increase in life threatening activity if they are not warned of the hindrances found on their path way. To alert the visually impaired people from this sort of life threatening activity we propose a "Smart Stick" for Visually Impaired People. This stick is capable of identifying the obstacles on their path way in advance and alerts the visually impaired people to travel with utmost precaution. In addition to it, the system also alerts the visually impaired person regarding the puddles and blaze in their path way. All these activities are also communicated to them with the help of voice commands to facilitate them while walking. One of the Family members of the visually impaired people will receive the location details through SMS whenever the help is requested from the visually impaired persons' end.

R. Santhana Krishnan (✉)
ECE, SCAD College of Engineering and Technology, Tirunelveli, Tamilnadu, India
e-mail: santhanakrishnan86@gmail.com

V. Nagaraju
CSE, Saveetha School of Engineering, Tamilnadu Chennai, India
e-mail: nagarajuv.sse@saveetha.com

N. Sasikala
New Horizon College of Engineering, Bengaluru, Karnataka, India
e-mail: mail.nsasi@gmail.com

K. Lakshmi Narayanan
Francis Xavier Engineering College, Tirunelveli, Tamilnadu, India
e-mail: klnarayanan@francisxavier.ac.in

S. Sundararajan
SCAD College of Engineering and Technology, Tirunelveli, Tamilnadu, India
e-mail: researchsundararajan@gmail.com

Y. Harold Robinson
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India
e-mail: yhrobinphd@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_34

**Keywords** Audio play back module · Global positioning system · Global system for mobile communications · Puddles · Smart stick · Visually impaired

## 1 Introduction

In India, nearly 40 million people are visually impaired or blind [1], out of which 1.6 million populations are children. India contributes to 21.9% of the World's visually impaired people. According to survey conducted by Ministry of Health and Family Welfare [14], in India the occurrence of blindness has declined to a great extent after implementing various strategies from the Indian Government. Even though the rate of occurrence of blindness has reduced, the number of people count seems to be in a huge number as far as India is concerned. As per the census taken at 2011 [15], nearly 74% of the disabled population in India is between 5 to 60 years. Out of that disabled populations in India, 19% of the population are having eye sight related disability issues. Since there is a huge number of visually impaired person are found in India, there is responsibility among the citizen to invent a new device which is capable of supporting those visually impaired people to perform their routine task. Hence a voice assisted smart stick has been proposed which aids the visually impaired person during their travel. The important merits of this system are listed below;

- To alert the visually impaired person about the puddles and blaze (fire) in their path way.
- To alert the family member whenever the visually impaired person lost his/her way during the travel
- To pass the alert information through voice commands during the visually impaired person's course of journey.

## 2 Related Works

Kumar et al. [2] proposed a smart stick which is capable of alerting the blind people with the help of Ultrasonic sensor. Similar works were carried out by Manikanta et al. [5], Loganathan et al. [11] and Dhanuja et al. [10]. Nowshin et al. [4] proposed an intelligent walking stick which is capable of detecting the obstacles on 3 sides (front side, left side and right side) of the user using 3 different ultrasonic sensors. This is helpful for the blind people to get the alert information regarding the obstacles from all the surrounding areas in which they travel. Similar work was carried out by Satam et al. [6]. Sathya et al. [3] proposed a Raspberry-pi controlled smart stick for the blind people which are capable of identifying the hurdles in their path way using ultrasonic sensor. This system can also detect the presence of water using rain sensor. Similar works were carried out by Nada et al. [8], Bele et al. [13] and Dada Emmanuel Gbenga et al. [12]. Sudhakar [9] designed a smart stick which detects the obstacles in the path way using ultrasonic sensor and wetland using the soil moisture

sensor. Pruthvi et al. [7] discovered a smart stick which is attached with a camera. Camera found in the smart stick captures the image in the pathway and sends it to the processor. Then the processor compares the images with the trained dataset which consists of images of all possible obstacles. Using these details, the processor alerts the blind people using a voice command if the camera detects any obstacles in their pathway.

## 3 Proposed Methodology

The overall architecture is depicted in Fig. 1, and the architecture comprises of Arduino, which gets the data from various sensors [16, 17]. The ultrasonic sensor present in this system is capable of detecting the hurdles present in the pathway and indicates it to the Arduino [18, 19]. In response to this, Arduino alerts the visually impaired person by activating the buzzer and the vibration motor. The puddles present in the pathway are identified using soil moisture sensor and similarly the

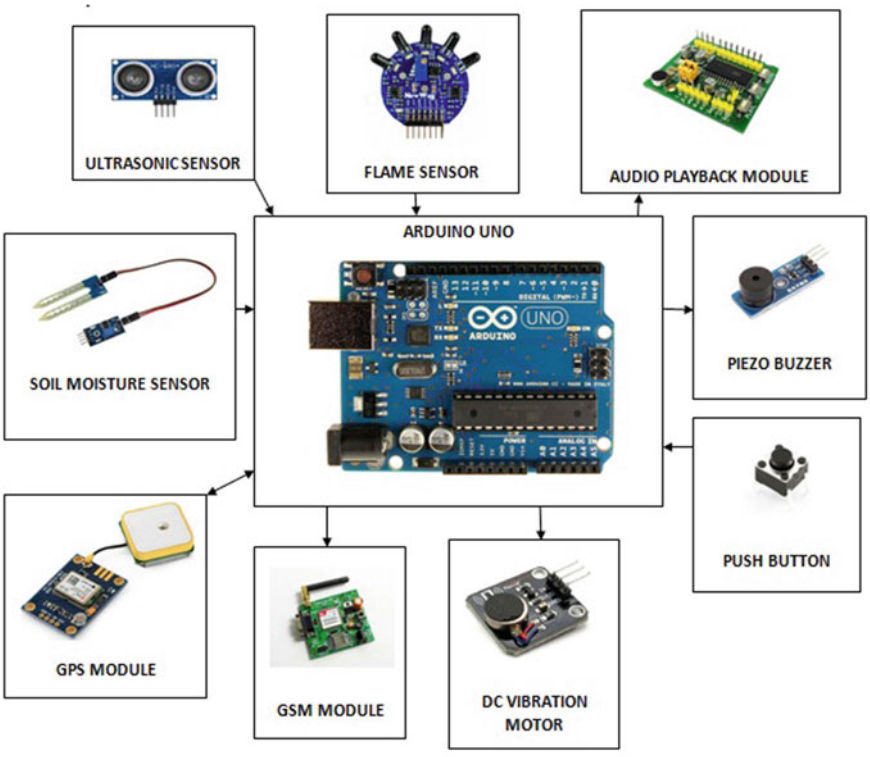



**Fig. 1** System architecture

flame particles present in the pathway are identified using flame sensor [20]. The information gathered by these sensors are sent to Arduino. After that Arduino alerts the visually impaired person by activating the buzzer and the vibration motor. Based upon the information received from the sensors, the voice command is activated by Arduino which helps the visually impaired person to travel precautious [21, 22]. Whenever the visually impaired person reaches an unknown place or if he forgets his way to home, then he is allowed to press an emergency push button. This will further trigger the Arduino to get the current location of the person from the GPS module and sends it to the family member through GSM to rescue them safely [23, 24]. Location of various sensors and other hardware components on the smart walking stick are shown in Fig. 2.

The overall flow of this system is explained using following steps, and Fig. 3 explains the overall flow diagram.

**Step 1:** The push button status is read by the Arduino when the system is turned ON.
**Step 2:** If the visually impaired person pressed the push button, then the Arduino gets the location information from GPS and sends the location details to family member through GSM module [25].
**Step 3:** Arduino reads the values of all the available sensors.
**Step 4:** Ultrasonic sensor is capable of detecting the hurdles in the footpath and send an alert information to the Arduino. Flame senor is capable of detecting the

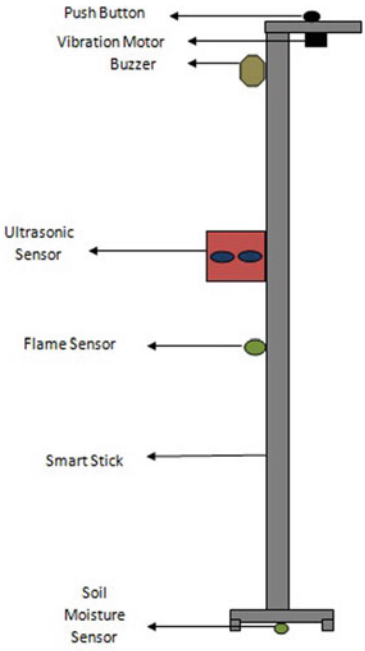



**Fig. 2** Sensor and other hardware locations on the smart walking stick

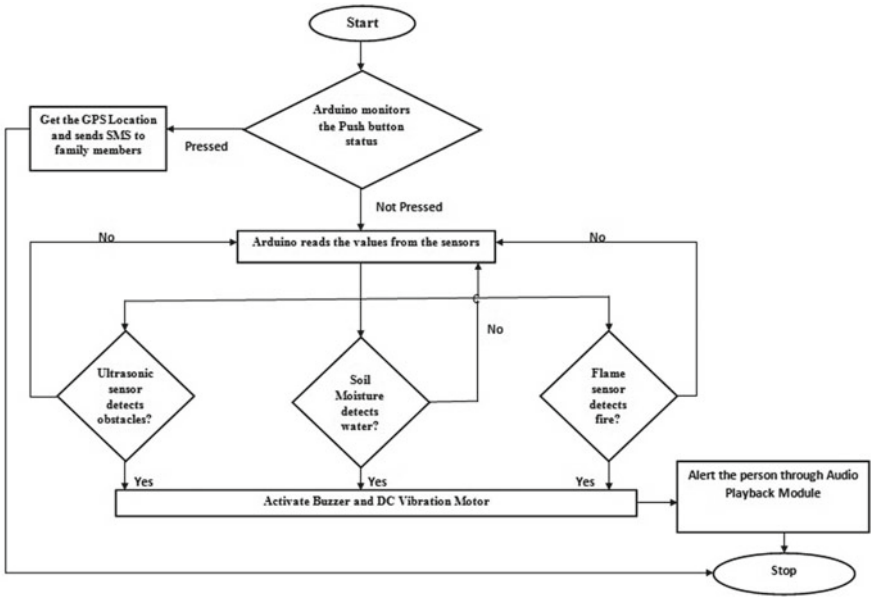



**Fig. 3** Overall flow diagram

flame particles present in the footpath and send an alert information to Arduino. Similarly, soil moisture sensor identifies the puddles in the footpath and sends an alert information to Arduino.

**Step 5:** Based on the inputs received from various sensors, Arduino sends the corresponding voice message. At the mean time it also activates the buzzer and vibration motor to make the visually impaired person physically feel the alert.

## 4 Result and Discussion

The proposed voice enabled smart stick for visually impaired system setup is explained using the Fig. 5. The major components of this system are GPS, Soil Moisture Sensor, Flame Sensor, GSM and Ultrasonic Sensor. The ultrasonic sensor present in this system detects the obstacle in the footpath and convey the information to Arduino. In turn Arduino activates the buzzer and the vibration motor to alert the visually impaired person regarding the hurdle. Flame sensor and soil moisture sensor present in this system are proficient in identifying the fire and puddles in the footway and sends an alert signal to Arduino. In turn Arduino activates the buzzer and the vibration motor to alert the visually impaired person regarding the fire and puddles present in the footpath. Based upon the information received from the sensors, the voice command is activated by Arduino which helps the visually impaired person to travel precautious. Whenever the visually impaired person reaches an unknown

**Table 1** Comparative analysis between the proposed and the existing methods

| Methods | MSE | RMSE |
|---|---|---|
| Existing model [12] | 0.89 | 0.97 |
| **Proposed model** | **0.62** | **0.65** |

place or if he forgets his way to home, then he is allowed to press an emergency push button. This will further trigger the Arduino to get the current location of the person from the GPS module and sends it to the family member through GSM to rescue them safely. This is explained using Fig. 6. Additionally, compared to existing model [12], the proposed model attained better performance in terms of MSE and RMSE, which is denoted in Table 1, and Fig. 4.

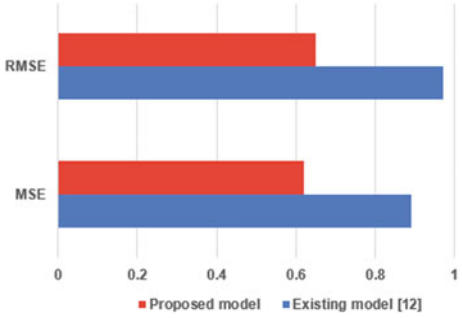



**Fig. 4** Graphical analysis of the proposed and the existing methods

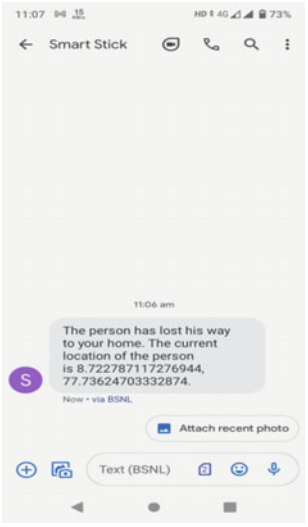



**Fig. 5** SMS intimation to family member

**Fig. 6** Overall system setup

## 5 Conclusion

The proposed smart stick will be very helpful for the visually impaired students in many ways. This stick identifies the obstacles on the path way in advance and it alerts the visually impaired people to travel with utmost precaution. In addition to it, the system also alerts the visually impaired person regarding the puddles and fire in their path way. This system also conveys the GPS location details to the family member whenever the help requested by the visually impaired person. All these activities are also communicated to them with the help of voice commands to facilitate them while walking. Hence, the proposed system facilitates the visually impaired people to travel anywhere without any hesitation.

## References

1. https://www.tribuneindia.com/news/archive/nation/india-home-to-20-per-cent-of-world-s-visually-impaired-738048
2. Kumar M, Verma R, Kumar M, Singh S, Singh ET (2017) Ultrasonic based smart blind stick for visually impaired persons. Int J Adv Res Electr Electron Instrument Eng6(3).
3. Sathya D, Nithyaroopa S, Betty P, Santhoshni G, Sabharinath S, Ahanaa MJ (2018) Smart walking stick for blind person. Int J Pure Appl Math 118(20):4531–4536
4. Nowshin N, Shadman S, Joy S, Aninda S, Minhajul IM (2017) An intelligent walking stick for the visually-impaired people. Int J Online Biomed Eng (iJOE) 13(11):94–101
5. Manikanta KS, Phani TSS, Pravin A (2018) Implementation and design of smart blind stick for obstacle detection and navigation system. Int J Eng Sci Comput 8(8):18785–18790

6. Satam IA, Al-Hamadani MN, Ahmed AH (2019) Design and implement smart blind stick. J Adv Res Dyn Control Syst – JARDCS 11(8):42–47
7. Pruthvi S, Nihal PS, Menon RR, Kumar SS, Tiwari S (2019) Smart blind stick using artificial intelligence. Int J Eng Adv Technol (IJEAT) 8(5S):19–22
8. Seddik AF, Nada A, Fakhr MA, Mashali S (2015) Effective fast response smart stick for blind people. In: Proceedings of the second international conference on advances in bio-informatics and environmental engineering - ICABEE 2015. ISBN: 978-1-63248-043-9, https://doi.org/10.15224/978-1-63248-043-9-29
9. Sudhakar S (2018) Smart cane for visually impaired. Int J Eng Sci Comput 8(8)
10. Dhanuja R, Farhana F, Savitha G (2018) Smart blind stick using arduino. Int Res J Eng Technol (IRJET) 05(03):2553–2555
11. Loganathan N, Lakshmi K, Chandrasekaran N, Cibisakaravarthi SR, Priyanga RH, Varthini KH (2020) Smart stick for blind people. In: 2020 6th international conference on advanced computing and communication systems (ICACCS), Coimbatore, India, 2020, pp 65–67. https://doi.org/10.1109/ICACCS48705.2020.9074374.
12. Gbenga DE, Shani AL, Adekunle AL (2017) Smart walking stick for visually impaired people using ultrasonic sensors and Arduino. Int J Eng Technol (IJET) 9(5). https://doi.org/10.21817/ijet/2017/v9i5/170905302
13. Bele S, Ghule S, Gunjal A, Anwat N (2020) Design and Implementation of Smart Blind Stick. In: 2nd International Conference on Communication & Information Processing (ICCIP). SSRN: https://ssrn.com/abstract=3645413, https://doi.org/10.2139/ssrn.3645413
14. https://npcbvi.gov.in/writeReadData/mainlinkFile/File341.pdf
15. https://censusindia.gov.in/census_and_you/disabled_population.aspx
16. Santhana Krishnan R, Golden Julie E, Harold Robinson Y, Raja S, Kumar R, Thong PH, Son LH (2020) Fuzzy logic based smart irrigation system using internet of things. J Cleaner Prod 252:119902
17. Ram CRS, Ravimaran S, Krishnan RS, Julie EG, Robinson YH, Kumar R, Son LH, Thong PH, Thanh NQ, Ismail M (2020) Internet of Green Things with autonomous wireless wheel robots against green houses and farms. Int J Distrib Sensor Netw 16(6)
18. Narayanan DKL, Ramesh GP (2018) Robust and brittle secured video for IoT. Int J Eng Technol 7(2.20):93–96
19. Robinson YH, Krishnan RS, Raja S (2020) A comprehensive study for security mechanisms in healthcare information systems using Internet of Things. In: Balas V, Solanki V, Kumar R (eds) Internet of Things and big data applications. Intelligent Systems Reference Library, vol 180. Springer, Cham. https://doi.org/10.1007/978-3-030-39119-5_15
20. Krishnan RS, Kannan A, Manikandan G, KB SS, Sankar VK, Narayanan KL (2021)Secured college bus management system using IoT for Covid-19 pandemic situation. In: 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV), pp 376–382. https://doi.org/10.1109/ICICV50876.2021.9388378
21. Krishnan RS, Sangeetha A, Kumar A, Narayanan KL, Robinson YH (2021) IoT based smart rationing system. In: 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV), pp 300–305. https://doi.org/10.1109/ICICV50876.2021.9388451
22. Narayanan KL, Ram CRS, Subramanian M, Krishnan RS, Robinson YH (2021) IoT based smart accident detection & insurance claiming system. In: 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV), pp 306–311. https://doi.org/10.1109/ICICV50876.2021.9388430
23. Thirupathieswaran R, Suria Prakash CRT, Krishnan RS, Narayanan KL, Kumar MA, Robinson YH (2021) Zero queue maintenance system using smart medi care application for covid-19 pandemic situation. In: 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV), pp 1068–1075. https://doi.org/10.1109/ICICV50876.2021.9388454.

24. Krishnan RS, Narayanan KL, Murali SM, Sangeetha A, Sankar Ram CR, Robinson YH (2021) IoT based blind people monitoring system for visually impaired care homes. In: 2021 5th international conference on trends in electronics and informatics (ICOEI), pp 505–509. https://doi.org/10.1109/ICOEI51242.2021.9452924
25. Niranjana R, Darney PE, Narayanan KL, Krishnan RS, Fernando AV, Robinson YH (2021) Prolific sensor glove based communication device for the disabled. In: 2021 5th international conference on trends in electronics and informatics (ICOEI), pp 636–640. https://doi.org/10.1109/ICOEI51242.2021.9452966.

# Diagnosis of Parkinson's Disease Using Optimized Neural Network Model


**M. Anila and G. Pradeepini**

**Abstract** Parkinson's disease being neuro-degenerative syndrome that effects central nervous system and is observed in many people worldwide. PD diagnosis is complex for the clinicians as it requires meticulous analysis of the patient. Though there are many characteristics and symptoms that indicate the disease, voice characteristics play a major role among the predictive characteristics. Person with PD experiences several vocal degradations like shaky and low speech. Voice analysis offers the additional benefit of being non-invasive, low cost and simple to diagnose. Many enthusiastic and great researchers have created new models and improved existing models in this area, and there is a vast amount of research in this field all over the world. The proposed optimized neural network (Opt-NN) model is implemented and compared it to various algorithms such as random forests, SVM, XG Boost, and KNN. Among all the algorithms used, the proposed model turned up to be the best algorithm with accuracy 95.14%

**Keywords** Accuracy · K-Nearest neighbor principal component analysis Parkinson's disease · Neural network · Random forest · SVM · XG boost

## 1 Introduction

Parkinson's disease is a progressive neurological disease that primarily affects the elderly. The cause of Parkinson's disease is unclear, but if the disease is diagnosed at early stage, the symptoms can be alleviated. Majority of the studies revealed that people with Parkinson's disease had voice problems. As a result, voice data can be used to diagnose Parkinson's disease [1–3].

Millions of people are affected with Parkinson's disease, according to the American Parkinson Disease Association, and causes severe health problems. Even though people with Parkinson's disease show a wide range of symptoms, determining the root cause remains difficult. People over the age of 60 are more likely to develop

M. Anila (✉) · G. Pradeepini
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India
e-mail: anilarao.m@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_35

this disease early because their bodies are more susceptible to degenerative diseases. PD is caused by a lack of dopamine or a decrease in dopamine levels, which makes motor movements difficult. There are two types of PD symptoms: motor symptoms and non-motor symptoms, clinical tests of motor symptoms are used to diagnose the disorder. Most patients with Parkinson's disease have vocal impairments, which are referred to as dysphonia. The key characteristic used to diagnose the presence of PD is dysphonia [4–6].

The diagnosis of Parkinson's disease at an early stage is a difficult challenge for doctors because the symptoms intensify and affect the individual day by day. Many researchers in this area have conducted comprehensive surveys and developed numerous models for detecting Parkinson's disease. [7] created a model that used a combination of SVM and a Gaussian Radial basis kernel function to predict PD with a 91.4 percent accuracy. [8] have done a comparison of regression tree, decision tree, and ANN and found that ANN produces better results. [9] proposed a multi-class classifier with an accuracy of 89.47 percent, as well as a new collection of measures and a different strategy for selecting features. [10] presented a fuzzy-based transformation methodology that was mixed with an SVM classifier to achieve a 93.47 percent accuracy. For successful classification, it is critical to practice feature selection to select the most important attributes [11].

Aim of this paper is to develop an optimized Neural Network Model for classification of PD. We proposed, an optimal NN-model including PCA for attribute selection. Remaining content of the paper is organized as: various aspects and results achieved by other authors is discussed in Sect. 2. Then we presented our Proposed methodology in Sect. 3 along with metrics and Sect. 4 deals with experimental set-up, results, and discussions. The next section includes future scope and conclusions.

## 2 Literature Survey

The analysis for PD diagnosis using voice dataset is discussed in [1]. The speech dataset is analyzed using a variety of machine learning algorithms. The speech dataset includes the voice frequencies of 31 Parkinson's disease patients. NN shows highest accuracy of all algorithms, while random forest has a decent accuracy and Naive Bayes has the lowest accuracy for disease detection.

The author proposed a hybrid intelligent framework for predicting disease progression in the paper [2], which used unique methods to eliminate noise, a clustering method to define class labels, and prediction methods to predict disease progression. PCA is used to determine which dimensions are the most important. Later, SVR approaches, and neuro-fuzzy interface systems are used. This hybrid intelligent system significantly improved the accuracy of PD prediction. Using deep neural networks on speech datasets, the severity of the disease can be predicted [3]. Tensor flow is a deep learning library that is used to implement artificial neural networks to predict the state of Parkinson's disease.

The experiment was evaluated using standard methods for separating a healthy person from a person with Parkinson's disease by detecting dysphonia in this paper [4]. PPE (Pitch Period Entropy), a new measure of dysphonia, is added. This procedure has been found to be reliable and has revealed only a few perplexing results. The ability to distinguish healthy people from people with Parkinson's disease has been demonstrated with a kernel SVM approach in conjunction with conventional methods. The use of dysphonia measurements for early diagnosis of Parkinson's disease is discussed [5]. The study's aim is to find the smallest subset of features that are important to PD score. This method yielded two scores to denote the test sample, i.e., an individual with Parkinson's disease or a healthy person. To optimize prediction generalization, a model with minimal bias was also used. To construct learning functions based on characteristics, a combination of the expectation maximization algorithm and genetic programming (GP-EM) is used in this paper [6]. To fit the data as a modular structure, the transformed data is modelled as a Gaussian mixture using EM. This allows us to determine whether or not an individual has Parkinson's disease.

Genetic Algorithm, Extreme Learning Machines and Wavelet Kernel are used in the proposed solution. For classification, an ELM learning approach is used to train a single layer neural network. In the WK-ELM structure, WK has three customizable parameters. The number of hidden neurons chosen in the architecture had a significant impact on ELM output. To find the best values for these parameters and the number of hidden neurons in ELM, a genetic algorithm was used. A few metrics, such as sensitivity, accuracy, and ROC curves, are used to assess the efficacy of proposed methodology.

One of the most cost-effective strategies for diagnosing people is to use a Deep Belief Network [8]. The DBN is capable of diagnosing Parkinson's disease by separating and analyzing patients' speech data. After the most important features have been detected, the DBN is used to establish a pattern and fit voice samples. DBN uses two stacked Restricted Boltzmann Machines (RBMs) and an output layer to classify PD. Since the initial weights are random, the RBM technique is used in unsupervised learning. The fine-tuning is done with Back Propagation. The proposed system's efficacy is assessed by contrasting experimental findings with different methodologies.

The optimal solution to the problem of Parkinson's disease is identified in this paper [9], which involves three major steps and a novel Multiple Feature Evaluation method. The proposed approach uses a multi-agent scheme with five classification techniques: Naive Bayes, Decision Tree, Neural Network, Random Forest, and SVM. The tenfold-CV is used to measure process learning and monitor performance differences. The results show that MFEA identifies the best set of attributes and thus improves classifier efficiency.

They used an incremental support vector machine to estimate the Unified Parkinson's Disease Rating Scale in this paper [10] (UPDRS). The aim of this approach is to minimize the amount of time needed to diagnose Parkinson's disease and improve the performance. The study's key findings are that the approach

**Table 1** Hyper parameter setting

| | |
|---|---|
| Number of hidden layers | 3 |
| Number of neurons in hidden layer | 50, 180, 546 |
| Epochs | 100 |
| Batch sizes | 15 |
| Activation function | ReLu (Hidden Layers), sigmoid (Output Layer) |
| Optimizers | SGD |
| Learning rates | 0.01 |
| Performance metrics | Accuracy, MSE, MAPE, F1, Recall, Precision |
| Cross-validation | 10 |

incorporates reducing dimensions/features, clustering techniques, and improved PD prediction accuracy while reducing computation time.

The author of this paper [11] used speech impairments to predict Parkinson's disease. The earliest symptom of this condition is a loss of voice. As a result, the author proposed a DNN classifier with stacked auto encoder as well as a SoftMax classifier to predict it. They used two separate datasets for this study. The dataset contains the "Parkinson Speech Dataset with Several Types of Sound Recordings (PSD)" and "Oxford Parkinson's Disease Detection (OPD)" datasets to differentiate between individuals with PD and those who are not. Deep neural networks achieve a 94% accuracy rate. To predict Parkinson's disease, various machine learning methods such as SVM, LDA, regression trees (RT), KNN, naive Bayes, RBF-NN, and Mahalanobis distance classifier are used. The SVM classifier had the best overall results. SVM achieves a 92.7% accuracy rate [12]. Feasibility study is given in the Table 1 containing information of authors and their methodologies.

## 3 Proposed Methodology

There are two important steps in our proposed method. The following sub sections clearly describe the process. First sub-section is made to explain importance of data visualization and the later sub-section presents the proposed approach of diagnosing PD.

### *3.1 Data Visualization*

Data visualizations will reveal new perspectives to those who are unfamiliar with the data, as well as communicate conclusions to those who may not see the raw data.

We used a dataset from the University of California, Irvine (https://archive.ics.uci.edu/ml/index.php). with a total of 31 records of people, 23 of whom are affected and 8 are not. The participants were between the ages of 46 and 85. Everyone gave an average of six continuous vowels "ahh…" to speak, ranging in length from one to 36 s [13], for a total of 195 samples. Different measures were applied to each recording, resulting in 22 real-value features. The last attribute (23rd feature) in the dataset shows the patient's condition, i.e., 0-represents a stable patient and 1-represents a patient with Parkinson's disease. Representation of boxplots, for every feature for visualization of data can be seen in Fig. 1. Boxplot also known as whisker-plot, is a five-point summary of data distribution ("minimum", first quartile (Q1), median, third quartile (Q3), and "maximum"). It provides information regarding the outliers and their values. It also gives us whether our data is symmetrical, and whether data is skewed [15].


**Fig. 1** Boxplots for Features in the PD voice Dataset

### *3.2 Methodology*

The basic step is to choose a desired model by the construction of a Neural Network Architecture. To do so, the following steps are performed:

a) Load dataset, the features extracted from voice recordings of the patients.
b) Perform pre-processing to remove noisy data (if any).
c) Apply PCA to extract most-relevant features. PCA is relied on the assumption that only the features with the most variance contain the information about certain classes.
d) Implement ML algorithms like KNN, Random Forest, SVM and XGBoost with proposedalgorithm.
e) Compare accuracies of all algorithms
f) Assessment of Model is done with various metrics.

Construction of Neural Network architecture for the Binary classification of PD is done as below.

a) Design the Neural Network with the criteria: Number of input features, Number of Hidden Layers, Number of neurons in each Hidden Layer, Class labels.
b) Execute the NN with hyper parameter tuning, record the accuracy, freeze the architecture for further implementation.

As a part of Data pre-processing, each attribute/feature is checked with presence of missing or null values and when it is done while implementation of our code, it is observed that no such values are existing. So, the methodology followed by the design of Neural Network with hyper parameter tuning. Following Fig. 2 depicts flow of proposed methodology, Opt-NN.

To ensure that most relevant features are extracted, Principal Component Analysis is used. The purpose, to make sure the most important features are identified so that we get better accuracy than existing models. Out of 22 features (23rd feature being representing status/label), 17 features are extracted.

Implementation of algorithms is performed using python in Google colab notes (https://colab.research.google.com). For hyper parameter tuning and setting up the architecture of Optimized Neural Network, GridSearchCv method for selecting hyper parameters and the setting of those values returned by the method as given in the Table 1, which when considered, achieved better accuracy than other existing models. As part of optimizer, SGD is used. In each iteration, SGD selects one data point at random from the whole data collection, thereby reducing computations. Neural Network with hyper parameter tuning is done as follows:

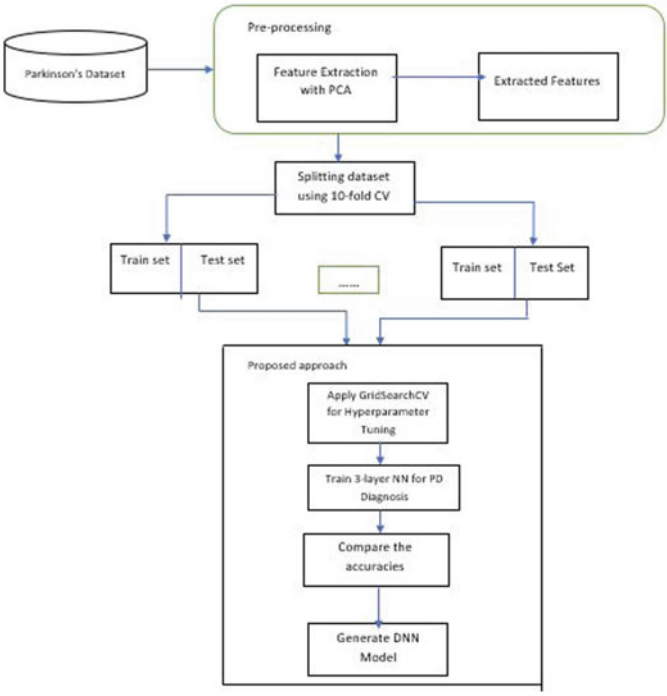



**Fig. 2** Proposed Opt-NN approach

## 4 Results and Discussions

We have used Python programming in Google colab for implementation of Opt-NN. It includes Artificial Neural Networks, Random Forest, SVM, XG-Boost, and KNN. To assess the performance of Opt-NN various metrics like Accuracy(ACC), precision, recall, f1-score MSE and MAC are recorded and is presented in Fig. 3 below.

The performance measure of the proposed model is given in Table 2. In this research work, mean square error, mean absolute classifier, recall, accuracy, F1 score, and precision values are evaluated.

### *4.1 Comparative Analysis*

The Comparison of mean square error, mean absolute classifier, recall, accuracy, F1 score, and precision values are given in Table 3. The proposed design achieved 0.05 value of mean absolute classifier, 0.95 recall value, 0.95 F1- score, and accuracy in 95.17%. In proposed design, mean square error, mean absolute classifier, recall, accuracy, F1 score, and precision values are improved than the conventional methods [11–14].

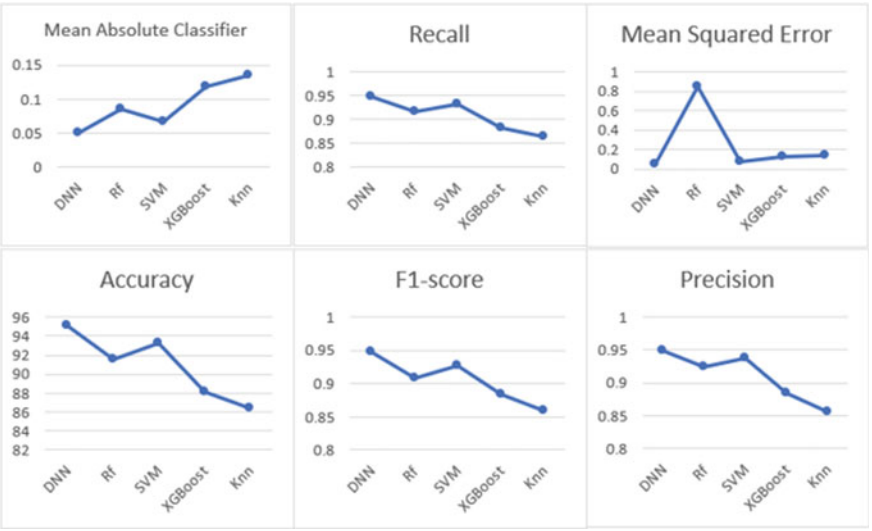



**Fig. 3** Comparison of metrics

**Table 2** Performance measure

| Performance parameters | Mean absolute classifier | Recall | Mean squared error | Accuracy | F1-score | Precision |
|---|---|---|---|---|---|---|
| Values | 0.05 | 0.95 | 0.18 | 95.14 | 0.95 | 0.95 |

**Table 3** Comparison of performance parameters

| Designs | Mean absolute classifier | Recall | Mean squared error | Accuracy | F1-score | Precision |
|---|---|---|---|---|---|---|
| RF [9] | 0.1 | 0.45 | 0.9 | 94.2 | 0.45 | 0.45 |
| SVM [10] | 0.12 | 0.5 | 0.8 | 92.21 | 0.5 | 0.5 |
| XGboost [11] | 0.08 | 0.65 | 0.7 | 91.4 | 0.65 | 0.65 |
| KNN [12] | 0.15 | 0.7 | 0.4 | 86.21 | 0.7 | 0.7 |
| Proposed design | 0.05 | 0.95 | 0.18 | 95.14 | 0.95 | 0.95 |

## 5 Conclusion

Parkinson's disease (PD) is a neurodegenerative ailment that affects the neurological system of the aged, with symptoms that progressively worsen. The condition is detected in this study by analyzing the voice data of persons with Parkinson's disease. Random Forest, KNN, SVM, and XG Boost are some of the machine learning algorithms utilized for this. For all the models employed, error rates are determined, and performance measures are analyzed to classify the best model. Accuracy, f1-score,

recall score, precision score, confusion matrix, and classification report are among the measures. With an accuracy of 95.14%, precision of 0.9484, and a f1-score of 0.9484, Opt-NN emerges as the best model among all the other ML approaches.

## References


1. Sakar CO, Kursun O Telediagnosis of Parkinson's Disease Using Measurements of Dysphonia. Springer (2009)
2. Nilashi M, Ibrahim O, Ahani A (2016) Accuracy improvement for predicting Parkinson's disease progression. Sci Rep
3. Grover S, Akshama SB, Yadav A, Seeja KR (2018) Predicting severity of Parkinson's disease using deep learning. Elsevier
4. Little MA, McSharry PE, Hunter EJ, Spielman J, Ramig LO (2009) Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE
5. Marar S, Swain D, Hiwarkar V, Motwani N, Awari A (2018) Predicting the occurrence of Parkinson's disease using various classification models. IEEE
6. Guo PF, Bhattacharya P, Kharma N (2010) Advances in detecting Parkinson's disease. Springer
7. Avci D, Dogantekin A (2016) An expert diagnosis system for parkinson disease based on genetic algorithm wavelet kernel-extreme learning machine. Hindawi
8. Al-Fatlawi AH, Jabardi MH, Ling SH (2016) Efficient diagnosis system for Parkinson's disease using deep belief network. Elsevier https://doi.org/10.1109/CEC.2016.7743941
9. Gokten ES, Uyulan C (2021) Prediction of the development of depression and post-traumatic stress disorder in sexually abused children using a random forest classifier. J Affect Disord 279:256–265
10. Kashef R (2021) A boosted SVM classifier trained by incremental learning and decremental unlearning approach. Expert Syst Appl 167:114154
11. Ji C, Zou X, Hu Y, Liu S, Lyu L, Zheng X (2019) XG-SF: An XGBoost classifier based on shapelet features for time series classification. Procedia Comput Sci 147:24–28
12. Narayan Y (2021) SEMG signal classification using KNN classifier with FD and TFD features. Mater Today: Proc 37:3219–3225
13. Sheela KG, Deepa SN (2013) Review on methods to fix number of hidden neurons in neural networks. Hindawi Publishing Corporation, Mathematical Problems in Engineering
14. Postuma R, Montplaisir J (2009) Predicting Parkinson's disease-why, when, and how? Parkinsonism Relat Disord 15:S105–S109
15. Das R (2010) A comparison of multiple classification methods for diagnosis of Parkinson disease. Expert Syst Appl 37:1568–1572

# Drug-Drug Interactions and Side Effects Prediction Using Shallow Ensemble Deep Neural Networks


**Alpha Vijayan and B. S. Chandrasekar**

**Abstract** Drug- Drug Interactions (DDI) is one of the key challenges faced in the drug development process. An average serious disease patient can take up to 8–10 drugs. The interaction between these drugs and likely hood of creating the side effects is very high, which includes total organ damages and leads to death as well. The DDI study is the need of the hour as most patients take multiple drugs at a time. The identification of DDI and it side effects requires millions of different combinations to be analyzed and predict the side effects upfront, to reduce the damage due to multiple drugs consumption. The time taken for drug development is usually is ten years for a single drug, which involves large cost. Study of interactions between multiple drugs itself takes 3–4 years. This paper proposes a custom deep shallow network with 2 hidden layers and ensemble multiple deep shallow networks to predict the side effects of Drug-Drug Interactions.

**Keywords** Activation functions · Deep neural network · Deep shallow neural networks · Drug-Drug interactions · Ensemble machine learning models

## Abbreviations

| | |
|---|---|
| CAS | Chemical Abstracts Service |
| DSSTox | Distributed Structure-Searchable Toxicity |
| InChI | International Chemical Identifier |
| IUPAC | International Union of Pure and Applied Chemistry |
| HPC | High Performance Computing |
| ODSF | Optimized Drug Similarity Framework |

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



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_36

| | |
|---|---|
| OTC drugs | Over the Counter drugs |
| LASSO | Least Absolute Shrinkage & Selection Operator |
| MeSH | Medical Subject Headings |
| MHRA | Medicines & Healthcare products Regulatory Agency |
| PRR | Proportional Reporting Ratio |
| RSS | Residual Sum of Squares |
| RTECS | Registry of Toxic Effects of Chemical Substances |
| SMILES | Simplified Molecular-Input Line-Entry System |
| SRS | Spontaneous Reporting System |
| UNII | Unique Ingredient Identifier |

## 1 Introduction

A Drug-Drug Interaction can be defined as interaction between multiple (2 or more drugs) when consumed at the same time period. DDI side effects occur when a particular is drug consumed while another drug is still active in metabolism. The drugs consumed can be different either a prescription medication or over the counter drugs or regular herbal or vitamins related drug. The consumption of both drugs can alter the way each drug functions and cause unexpected side effects. When multiple drugs are being taken; the interaction between the drugs happens, which can potentially increase or decrease the effectiveness of the drug. Hence, people who takes multiple medications are at greater risk for interactions. In summary, the likelihood of side effects increases with the number of drugs consumed as a part of the medication treatment.

## 2 Literature Review

An extensive literature survey to understand the existing work has been carried out and summarized.

Qiu et al. [1] performed a detailed analysis on all the drug- drug interaction research studies under three categories, literature-based extraction methods, machine learning-based prediction methods and pharmacovigilance-based data mining methods. Each paper was compared with different other researches in terms of accuracy and other model performance metrics. The proposed model falls under literature-based extraction methods and this paper was used as point of reference for comparison of proposed solution with historical research work.

Zylich et al. [2] proposed a visualization tool that can combine the existing algorithms and made the application available to users via web platform. There was no novel algorithm was proposed, they have combined 2 different algorithms and generated predictions based on these two algorithms.

Xu et al. [3] Proposed full attention mechanism models that can be used to predict drug- drug interactions. The researchers used User Generated Content for Drug- Drug Interaction Extraction content for predicting drug- drug interaction. An effective model should consider the properties of drugs and their structural chemical details to predict the drug- drug interactions. The current proposed model fails short for using as an input to health care industries due to lack drug structural composition considerations.

Zheng et al. [4] proposed an Optimized Drug Similarity Framework (ODSF) for result forecast. A clustering-based strategy was then embraced to upgrade OCDS and aberrant medication closeness which could support the immediate medication similitude was registered. The clustering-based strategy is considered to be effective when similar drugs are present in the sample database considered for the analysis.

Ngufor [5] research work manages methodical and organized prescient model for ADRs created from pre-clinical qualities of medications and unconstrained reports of ADRs in an appropriated superior processing structure. Prescient exactness was improved by utilizing Variational Bayesian outfit learning strategy. Carried out on HPC cloud machines, the Bayesian graphical outfit technique beat different strategies considered for correlation. Measurements for correlation were AUC and G-mean over an aggregate of 800 known results. The instinct behind the grouping approach utilized in this examination depends on the perception that comparative medications show comparative results, and prescient models built inside each bunch can be more exact than worldwide models developed on the total training set. The research works include the patient characters to be included in the side effects prediction. The spontaneous reporting system is traditional recommendation system that can suggest the side effects of the drug- drug interactions. The study is depending on the correlation between the variables under consideration. Correlation studies are subject to outlies and can create high error results.

Lu [6] research work point Adverse drug reactions (ADRs) are the major source of morbidity and mortality. An aggregate of 1000 assortments of medications were examined under 3 distinct classes (Rx, OTC-A, OTC-B). It is to be observed that Risk level of Rx drugs interaction is very high when compared to OTC- A, OTC-B.

A combined multi class order model was built to predict the drug- drug interactions. Frequentist/non-Bayesian Signal discovery techniques, for example, Proportional Reporting Ratio (PRR), Reporting Odds Ratio (ROR), Medicines and Healthcare items Regulatory Agency (MHRA) and Bayesian strategy like Information Component (IC) were utilized. A combination of different machine learning algorithms like Random Forest (Bagging algorithm), Gradient Boost (Boosting algorithm), Ada Boost (Boosting calculation) and Logistic Regression (individual algorithm). Destroyed procedure was utilized to adjust the imbalanced dataset.

Lee and Huang [7] presented drug results (SEs) are a significant bottleneck for drug advancement. The interaction of post-market drug withdrawal is exorbitant. Thus, the capacity to assess the likely symptoms of medications however ahead of schedule as conceivable seems to be basic during the medication plan and advancement measures. The aftereffects of these assessments can be utilized as direction in the work to diminish results and give safe treatments in the clinical setting.

**Table 1** Variables and dataset used for analysis-dependent variable: Side Effects (1–10)

| Category | Independent variables |
|---|---|
| Structures | Crystal Structures, D Conformer, D Structure |
| Names and identifiers | Canonical SMILES, InChI Key, InChI, IUPAC Name, Computed Descriptors, |
| Molecular formula | Depositor-Supplied Synonyms, MeSH Entry Terms, DSSTox Substance ID, UNII, UN Number, RTECS Number, NSC Number, ICSC Number, European Community (EC) Number, Deprecated CAS, CAS, |
| Chemical and physical properties | Springer Materials Properties, Other Experimental Properties, Kovats Retention Index, Collision Cross Section, Dissociation Constants, pKa, Decomposition, Stability/Shelf Life, LogP, Vapor Pressure, Density, Solubility, Flash Point, Melting Point, Boiling Point, Odor, Color/Form, Physical Description, Experimental Properties, Computed Properties, |

# 3 Proposed Methodology

## *3.1 Data Collection*

For harvesting drug characteristics from the PUBCHEM database, a custom script was built. It scrapes the InChI, Drug Canonical SMILES, Molecular Formula, Deprecated CAS, and DSSTox Substance ID after the user enters the relevant drug name. The molecular formula will reveal the medicine's components and will be utilized to construct the specific molecules that will be used to make the drug. The data from PUB CHEM was scraped using a combination of traditional API and effective scraping techniques. The independent and dependent variables utilized in the analysis are as follows (Table 1):

## *3.2 Pre Processing*

The data is scraped from PubChem database and processed through a series of steps like null value treatment, outlier treatment, standardization, normalization, non-character value remove. By doing these series of steps we have created a clean dataset which can used as an input for neural networks.

# 4 Model Network-Ensemble Algorithms

A deep shallow network of 5 layers and each layer has 8 nodes have been created to validate multiple regression algorithms. We have considered different properties of

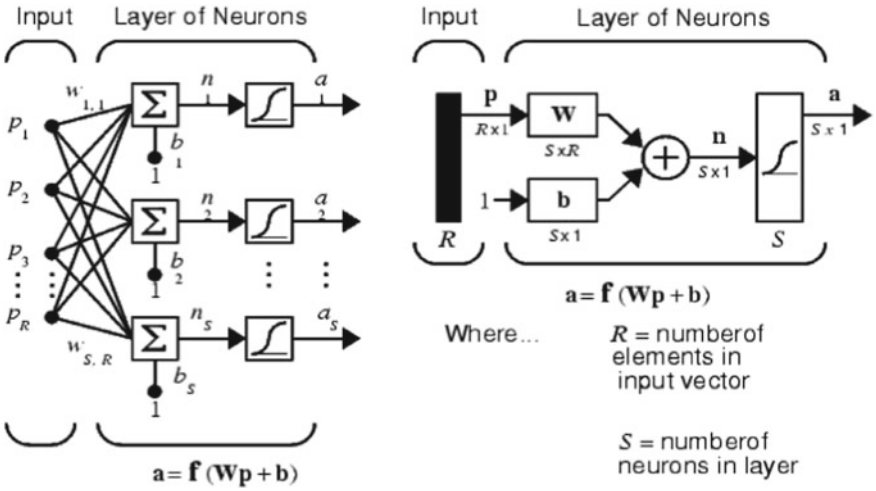



**Fig. 1** Shallow neural network architecture

drugs as dependent variables and side effects as y variable (Index of seniority of the drug). We have used 3 different algorithms and variables used for prediction of the final outputs.

**Steps in the Algorithm**

1. Create a deep shallow neural network with 2 hidden layers of 8 nodes each.
2. The activation functions that process the information are found in each layer.
3. Reduce the loss in each cycle through backpropagation.
4. To complete the final outputs, arrive at a minimum RMSE (Figs. 1 and 2).

## *4.1 Multivariate Regression*

The most popular prediction model across multiple areas is an algorithm that analyses the relationship between the variables are expressed from Eq. (1) to Eq. (5).

$$RSS(\beta) = (Y - X\beta)T(Y - X\beta) \tag{1}$$

We can use the least-square estimator to evaluate the boundaries:

$$\beta^{\wedge} = argmin\, \beta(Y - X\beta) \tag{2}$$

We can infer a closed structure for the estimator by calculating derivatives with respect to the parameters and setting the parameters. it equal to 0:

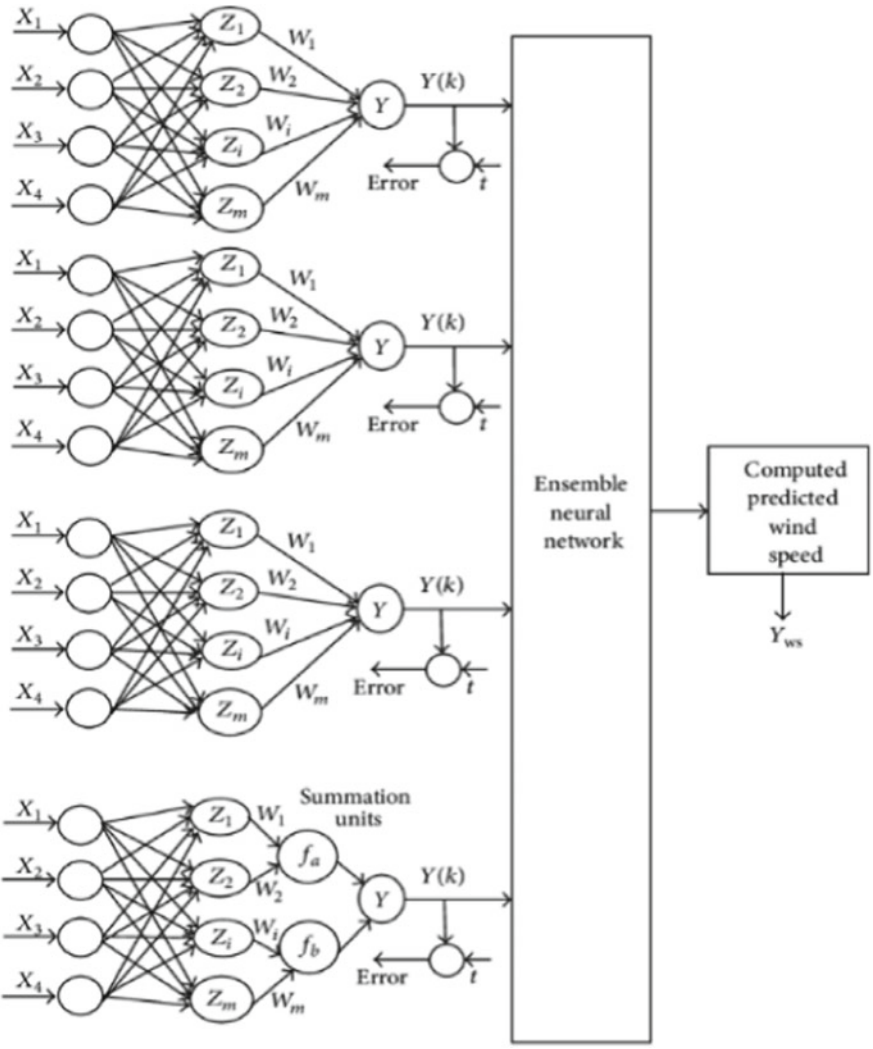



**Fig. 2** Ensemble shallow neural network models

$$\partial RSS\partial\beta\partial RSS\partial\beta = -2XT\left(Y - X\beta^{\wedge}\right) = -2XTY + 2XTX\beta^{\wedge} = 0 \quad (3)$$

solving for β^ we have:

$$XTX\beta^{\wedge} = XTY \quad (4)$$

$$\beta^{\wedge} = (XTX) - 1XTY \quad (5)$$

With labels and it is use that to build the model and the output variables that can be predicted is only continuous variables. It is most common predictive model across various domains.

### 4.2 Lasso Regression

Lasso Regression Required iteration to get the final solution. It makes Lasso computationally more intensive and it takes the sum of absolute values of coefficients ($\lambda \Sigma |\alpha|$). We can change the first streamlining issue with the limitation characterized by least absolute shrinkage and selection operator, LASSO is expressed in Eq. (6) and (7).

$$\beta^{\wedge} lasso = argmin\, B\{(Y — XB)T(Y — XB) + AYj = 1p|\beta jl\} \tag{6}$$

or, in an equivalent form:

$$\beta^{\wedge} lasso = argmin\, \beta(Y — XB)T(Y — X\beta),\ subject\ to \sum j = 1\ p|\beta j| \leq t \tag{7}$$

The parameter $\lambda \geq 0$ controls how much shrinkage you want. If you set $\lambda = 0$ we are back to the standard linear model.

### 4.3 Ridge Regression

$$\beta^{\wedge} ridge = argmin\beta\{(Y - X\beta)T(Y - X\beta) + AYj = 1B2j\} \tag{8}$$

or, in an equivalent form which are expressed in Eq. (8) and (9):

$$\beta^{\wedge} ridge = argmin\, B(Y — X\beta)T(Y — X\beta),\ subject\ toy\ j = 1\ \beta 2j < t, \tag{9}$$

Again, if you set $\lambda = 0$ we are back to the standard linear model.

### 4.4 Elastic Net Regression

This is one of Regularized Linear Regression which is a combination of both ($L1\&L2$) Ridge and lasso (they are penalty functions) forms an ensemble technique which is expressed in Eq. (10). It overcomes

$$\beta^{\wedge} relastic\ net = argmin\, \beta\{(Y — X\beta)''(Y — X\beta) + ADj = 1(@\beta 2j$$

$$+ (1—a)|\beta jl)\} \tag{10}$$

The drawbacks of Lasso Regression, for $q \geq 0$.

We set several constraints for the optimization problem by changing the value of $q$. We make a tradeoff between tether and edge relapse while selecting estimations of $q(1, 2)$. One problem with this methodology is that, except from $q = 1$, the capacity is differentiable at 0 and so LASSO's feature selection capacity is lost. The elastic net is offered in this context as an alternative regularization strategy that preserves the lasso's capabilities for feature selection while shrinking coefficients of correlated characteristics like ridge (0, 1). The elastic net becomes a lasso when the value is 0, and a ridge when the value is 1.

## 5 Results and Inferences

### *5.1 Algorithm's Comparisons*

A set of 1000 drugs has been considered into for the analysis and for the verification of proposed methodology. Below results are the summary of outcomes (Table 2).

To predict side effects and no side effects, four different models were used as the basic classifier. These are high-accuracy individual classifiers. To anticipate the model's eventual outcome, an ensemble model based on these individual classifiers was developed. Each model tested had an accuracy of at least 80%, and a combination of these four models resulted in a 91% accuracy (Table 3).

**Table 2** Comparative analysis-proposed algorithms

| | Multivariate regression | | Lasso regression | | Ridge regression | | Elastic net regression | |
|---|---|---|---|---|---|---|---|---|
| | Side effects | No side effects | Side effects | No side effects | Side effects | No side effects | Side effects | No side effects |
| Side effects | 420 | 66 | 414 | 66 | 402 | 62 | 426 | 24 |
| No Side effects | 80 | 434 | 86 | 444 | 98 | 438 | 84 | 486 |

**Table 3** Comparative analysis-historical research

| State-of-the-art model | Year of publication | Accuracy (%) | Recall (%) | Specificity (%) |
|---|---|---|---|---|
| Proposed model | 2021 | 91 | 84 | 85 |
| Qiu et al. [1] | 2021 | 78 | 82 | 84 |
| Zylich et al. [2] | 2020 | 85 | 81 | 76 |
| Bo et al. [3] | 2019 | 88 | 79 | 77 |

**Table 4** Model key performance indicator comparison-values

| Model | Accuracy | Recall | Specificity |
|---|---|---|---|
| Regression | 85 | 0.84 | 0.86 |
| Lasso | 84 | 0.828 | 0.87 |
| Ridge | 84 | 0.8 | 0.8 |
| Ensemble deep neural network | 91 | 0.84 | 0.85 |

## *5.2 Historical Research Comparison*

Shallow neural networks, according to the statistics, outperform the existing model in terms of accuracy by a significant margin. According to a recent study that used a variety of methods, up to 88% of predictions were true. Using a unique ensemble deep shallow neural network, we were able to achieve a final accuracy of around 91%. The predictions are double-checked using k fold validation methods to ensure that they are right. Other metrics such as false positive rate, true positive rate, precision, recall, and f1 score have been calculated and compared for different models. The suggested network outperforms on all of the aforementioned critical performance parameters (Table 4).

## *5.3 Comparison Charts*

Many independent researchers' results compared to the over model during the last three years showed that we outperformed all other research papers in terms of accuracy, recall, specificity, and F1 Score. We notice a 5% gain in accuracy when comparing older tests to the current proposed model. This is a significant rise when compared to earlier study. Both independent and ensemble neural networks were used in the proposed model. The ensemble model outperformed individual models on all KPIs, including accuracy score, precision, and recall (Fig. 3).

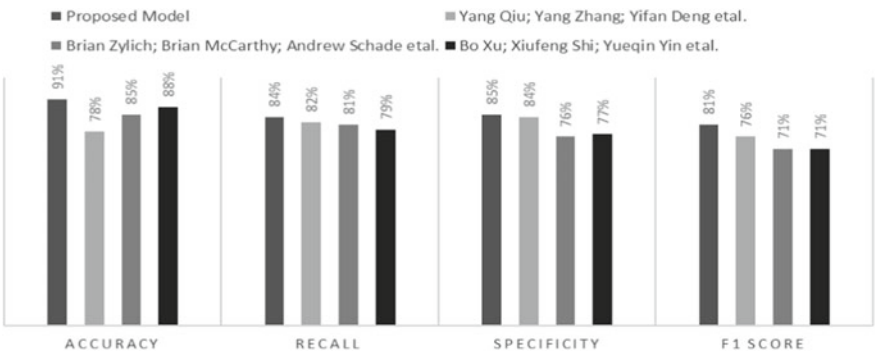



**Fig. 3** Graph-model key performance indicator comparison

## 6 Conclusion

Deep networks are a type of network that is used to tackle problems of greater complexity. We can answer some of the main breakthrough research problems in the health care industry using an efficient quantum computing and algorithms. Because there are hundreds of characteristics to consider throughout the calculations, determining the best combination of pharmaceuticals that can be eaten together is a huge computational problem. The proposed model is based on recognizing these interactions at a granular level and determining how drug-drug interactions affect side effects. The algorithms were created to work in a variety of ways. The tests were carried out using deep neural network simulators and a variety of drug interactions. The outcomes were superior to the traditional solutions.

## References

1. Qiu Y, Zhang Y, Deng Y, Liu S, Zhang W (2021) A comprehensive review of computational methods for drug-drug interaction detection. IEEE/ACM Trans Computat Biol Bioinform 1. https://doi.org/10.1109/TCBB.2021.3081268
2. Zylich B et al (2018) Multiple-drug interaction analytics platform: developing a minimum viable web application for visualizing multi-drug interactions. Worcester Polytechnic Institute
3. Xu B et al (2019) Incorporating user generated content for drug drug interaction extraction based on full attention mechanism. IEEE Trans NanoBiosci 1. https://doi.org/10.1109/TNB.2019.2919188
4. Zheng Y, Ghosh S, Li J (2017) An optimized drug similarity framework for side-effect prediction. https://doi.org/10.22489/CinC.2017.128-068
5. Ngufor C, Wojtusiak J, Pathak J (2015) A systematic prediction of adverse drug reactions using pre-clinical drug characteristics and spontaneous reports, pp 76–81. https://doi.org/10.1109/ICHI.2015.16
6. Wei J, Lu Z, Qiu K, Li P, Sun H (2020) Predicting drug risk level from adverse drug reactions using SMOTE and machine learning approaches. IEEE Access Digital Object Identifier, pp 185761–185775, 21 October 2020. https://doi.org/10.1109/ACCESS.2020.3029446

7. Lee W-P, Huang J-Y, Chang H-H, Lee K-T, Lai C-T (2017) Predicting drug side effects using data analytics and the integration of multiple data sources. IEEE Access 5:20449–20462
8. Liu M, et al (2012) Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. J Am Med Inform Assoc 19(e1):e28–e35
9. Sarkar B, Karis G, Konstanz J, Riel J (2001) Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th International conference on World Wide Web, pp 285–295. ACM
10. Celebes R et al (2014) Link prediction for drug-drug interaction network. In: Proceedings of the International Conference on Applied Informatics for Health and Life Sciences, pp 99–102
11. Niu Y, Zhang W (2017) Quantitative prediction of drug side effects based on drug-related features. Interdiscip Sci. 9(3):434–444. https://doi.org/10.1007/s12539-017-0236-5. Epub 2017 May 17. PMID: 28516319
12. Zhou Z-W et al (2015) Clinical association between pharmacogenomics and adverse drug reactions. Drugs 75(6):589–631
13. Lee WP, Huang JY, Chang HH, Kee K-T, Lai C-T (2017) Predicting drug side effects using data analytics and the integration of multiple data sources. IEEE Access 5:20449–20462

# Dual-Band Microstrip Patch Antenna for 5G-NR Applications


**R. Manu, C. M. Bhoomika, Abhinandan Ajit Jugale, and Mohammed Riyaz Ahmed**

**Abstract** 5G has become the next hope in the technology field. Though generations of communication technologies have improved a lot, they also have become complex in nature. The implementation of 5G-NR has already been tried at some level. This New Radio (NR) which works on the current infrastructure of the 4G study was started in 2015 within 3GPP and was expected to be commercially launched by the end of 2019. Massive MIMO system is one of the enabling technologies which has got huge attention recently. This paper proposes a dual-band microstrip patch antenna operating at 30 and 72.6 GHz. The proposed antenna has been designed to work under the application of 5G-NR range of frequency. The substrate material of proposed antenna is Rogers RO3003(TM), whose dimensions are $5.9861 \times 7.3880 \times 0.6$ mm. The antenna has been analyzed for its return loss, VSWR, gain, radiation pattern and current distribution. The obtained results are varied to suit the requirements and are discussed for 5G application.

**Keywords** 5G-NR · Dual band frequency · HFSS v15.0 · Massive MIMO · Microstrip patch antenna · mmWave

## 1 Introduction

The race towards 5G is moving forward at full speed. With tech giants such as Apple, Samsung, etc. already unavailing their 5G products, it becomes crucial to spotlight the important factors of 5G. The technology, scientifically called 5G New Radio (5G-NR) uses sub 6 GHz of frequency to operate. Usually occupying a bandwidth from 28 to 100 GHz, 5G-NR gives up to ten times faster data transfer rate than present available 4G networks [1, 2]. For the systems to be connected in 5G, high data rates across a wide range of area has to be achieved. As the technology empowers the user with extended capabilities, the data consumption per device will be increased with the new speed [3]. When the number of devices connected with a good data rate in

R. Manu (✉) · C. M. Bhoomika · A. A. Jugale · M. R. Ahmed
REVA University, Bangalore 560064, India
e-mail: manukushal18@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_37

a particular area of connection increases, the requirement towards high reliability and low latency of the entire network also increases [3]. With a huge boom towards the IoT devices, 5G also covers complex and critical operations to emerge with new dimensions and implementations. Applications such as Industrial IoT, Virtual Reality (VR), Augmented Reality (AR), and autonomous vehicles are some examples which can be extensively realized by 5G [4]. Now when IoT devices are connected in a network, it has to be realized that IoT devices send data in irregular intervals of time and 5G should be able to handle these irregular and spontaneous IoT communications. The fifth generation of the communication network uses millimeter waves. As shown in Fig. 1 supporting a large number of devices across a wide range of area is the main challenge of mm waves [5]. The concept of enhanced Mobile Broadband (eMBB) solves this issue of connectivity [6, 7]. The basic application of providing 5G to consumer electronic devices such as smartphones, computers, etc. is taken care of by this protocol. The requirement towards the network is highly reliable and least latent is achieved by the second protocol or case of the 5G-NR, Ultra-Reliable Low Latency Communications (URLLC) [8, 9]. Mission Critical controls or operations such as Autonomous vehicles, Medical surgeries through AR, Outer space missions, etc. are supported by this protocol. The IoT type of communications or generally machine type of communications which are highly irregular and spontaneous are taken care by the third protocol, massive Machine Type Communication (mMTC). mMTC handles all machine types of communication thus paving a smooth path towards Industrial

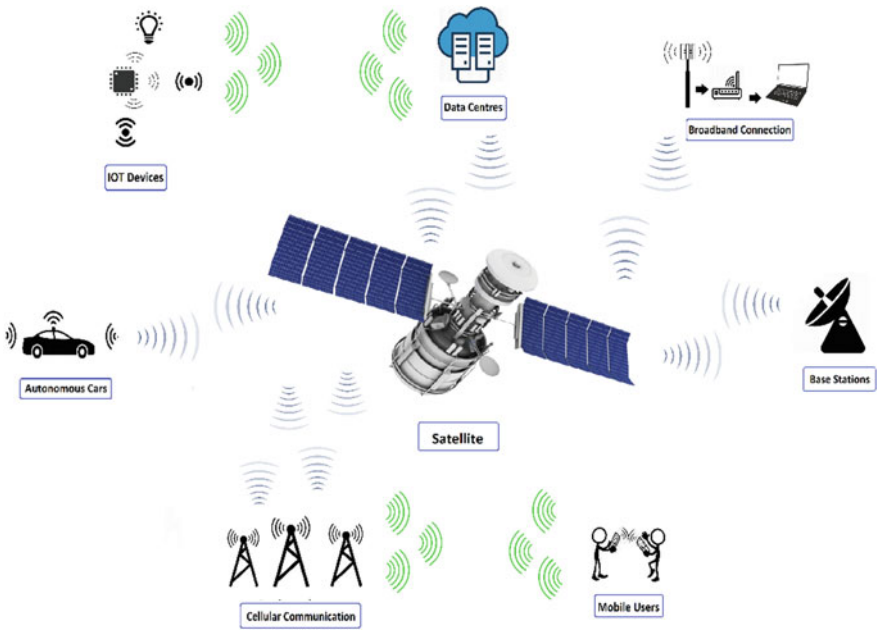



**Fig. 1** Pictorial representation of 5G-NR use cases.

**Table 1** Comparison with related work

| Reference paper | Patch size (mm) | Radiation frequency (GHz) | Return Loss (dB) | Gain (dB) |
|---|---|---|---|---|
| [16] | 1.3 × 1.2 | 28, 38 | −43, −18 | 3.75, 5.06 |
| [17] | 3 × 3 | 30, 50 | −12.5, −14.85 | 10.2, 11.6 |
| [14] | 1 × 1 | 38, 54 | −15.5, −12 | 6.9, 7.4 |
| Proposed work | 2.386 × 3.788 | 30, 72.6 | −22.07, −19.29 | 5.63, 7.63 |

IoT (IoT) [10]. It is now doubt that 5G will offer great number of services and will provide huge quantum leap for the technology dependent industries. To cope up with the current interest in the field, a great amount of research has to be done within the domain.

## 2 Related Works

In [11], Swain et al. have used a quadrilateral feed patch which has allowed them to achieve high gain, dual-band frequency and the use of quadrilateral feed patch has improved their impedance matching. In [12], the authors have designed a patch antenna which can be used in multiple-input-multiple-output (MIMO) applications, here the authors have used a Rogers 4003 substrate to achieve millimeter wave. In [13] to achieve an increase in the directivity and bandwidth of antenna the authors have used more dielectric substrate and proper air gap spacing, the authors have proposed a microstrip patch antenna with beam-steering capability. In [14], the authors have compared the characteristics between a 5G feed line antenna and a 5G antenna array with same dimensional values and same substrate material of Rogers 5880, and they have concluded that they have found high gain by using an array. In [15], the authors have designed a dual band array antenna which is of 8 elements, and the proposed array antenna ensures lower mutual coupling with radiation and acceptable gain characteristics.

The proposed antenna in this paper is a dual-band frequency with high radiation frequency of 30 and 72.6 GHz, high return loss of −22.07 and −19.29, and high gain 5.63 and 7.63 dB. Some comparisons between the proposed antenna and work related is given in Table 1.

## 3 Antenna Design

While designing an antenna, parameters such as gain, beam width, polarization and impedance must be considered [18], Fig. 2 depicts the antenna designed by the authors. Based on their application and their frequency range, all the parameters

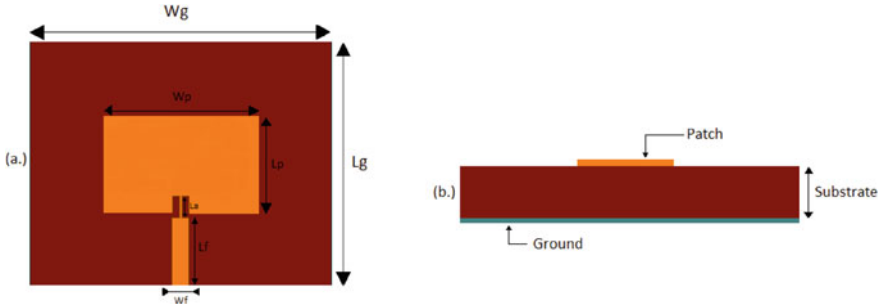



**Fig. 2** **a** Top view of designed antenna; **b** Side view of designed antenna

are designed and the substrate material is decided. The authors use Ansys High-Frequency Function Simulator (HFFS) v15.0 to design the antenna and simulate the results. The first step of designing any antenna is by placing the ground at the center of the axis, this forms the base which provides physical stability to the antenna. The Second layer present above the ground is the substrate material whose relative permittivity decides the behavior of the antenna. The substrate is sandwiched in-between the ground and the patch. The height of the substrate decides the bandwidth of the antenna wherein they are directly proportional to each other. Upon the substrate, the radiating and predominant part i.e., the patch is placed [19, 20]. The current to the patch can be provided through feeding techniques and some of the feeding techniques are namely line feed, coaxial feed and coupling feed. A rectangular patch with line feed is designed. To match the impedance of the antenna, a microstrip line feed is used. It is the simplest and easiest technique for feeding an antenna. Conducting strip supplies current through one end of the patch. The conducting material is small when compared to the size of the antenna [21, 22]. Patch and feed are placed on the same plane above the substrate. Radiating power determines a few of the important parameters of an antenna such as gain, bandwidth and efficiency. It is directly proportional to the fringing field, which can be increased by increasing the antenna width, substrate height and decreasing the dielectric constant. By creating a certain slot on the patch, the fringing field can also be increased. The ground dimension of the designed antenna is 5.9861 $\times$ 7.3880 mm. The substrate material used by the authors is Rogers RO3003(TM) which has a dielectric constant of 3. Its height is 0.6 mm. Further, the patch has a dimension of 2.3861 $\times$ 3.7880 mm, it is placed at the center of the axis. The dimensions of the designed antenna are shown in Table 2.

From one side of the substrate, the patch is fed with a line feed of dimensions 1.6653 $\times$ 0.4163 mm. All the dimensions are varied and calculated to obtain and observe a desirable gain, return loss, current distribution, VSWR and radiation pattern.

**Table 2** Parametric dimensions of the designed antenna

| Parameter | Dimension in mm | Parameter | Dimension in mm |
|---|---|---|---|
| Lg | 5.986 | Wg | 7.388 |
| Ls | 5.986 | Ws | 7.388 |
| Lp | 2.386 | Wp | 3.788 |
| Lf | 1.6533 | Wf | 0.4163 |
| H | 0.6 | - | - |

## 4 Results

For determining the design parameter and the performance of antenna design, radiation pattern, resonance frequency, VSWR, and its current distribution we have used Ansys High-Frequency Structure stimulator (HFSS) v 0.15.0 tool. Satisfactory results for the antenna was achieved which satisfied all the necessary parameters for a stable and working antenna. The return loss of an antenna is appreciated if it is below −10 db. Greater the result loss, higher is the signal strength and transmission efficiency [23, 24]. As shown in Fig. 3, the return loss obtained is −22.0734 at 30 Ghz and −19.2946 at 72.6 Ghz. Voltage Standing Wave Ratio (VSWR) of an antenna and the feed must be below 2 and above 1 for the practical applications, lesser power is reflected from the antenna and so that a large amount of reflected power would not damage the antenna [25, 26]. The designed antenna has a VSWR of 1.3711 and 1.8916 dB at the resonant frequency as shown in the Fig. 4. By plotting the radiation being emitted from the antenna, radiation pattern was obtained in E and H plane, whose obtained results were satisfactory as shown in Fig. 5(a) and Fig. 5(b). The 3D plots in Fig. 6 depicts the gain of the frequencies obtained by the antenna. The gain obtained by frequencies 30 and 72.6 GHz are 5.63 and 7.63 dB respectively.

Further moving on to the next parameter, the gain obtained should be of positive value as it indicates how the antenna converts the input power to waves. A higher gain at a particular direction would mean that the signal is received or transmitted without

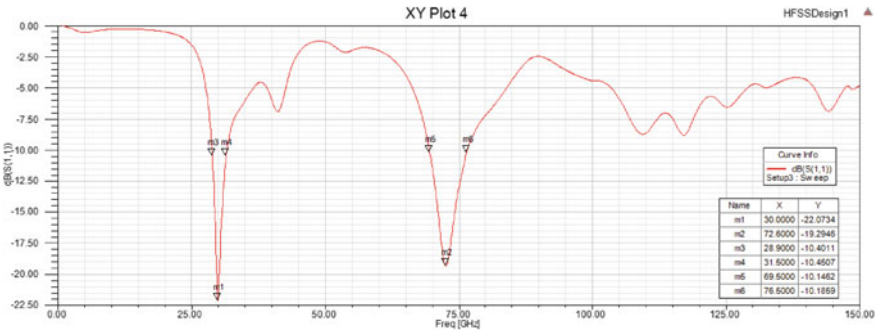



**Fig. 3** Return loss versus frequency. Antenna loss should always be less than –10 dB

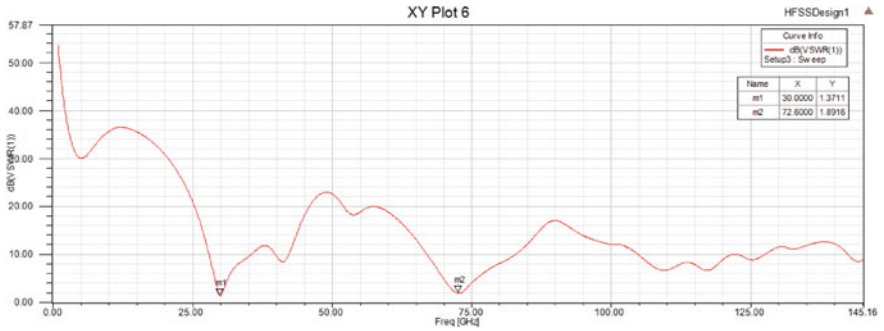



**Fig. 4** VSWR of antenna must practically be between values of 1 and 2

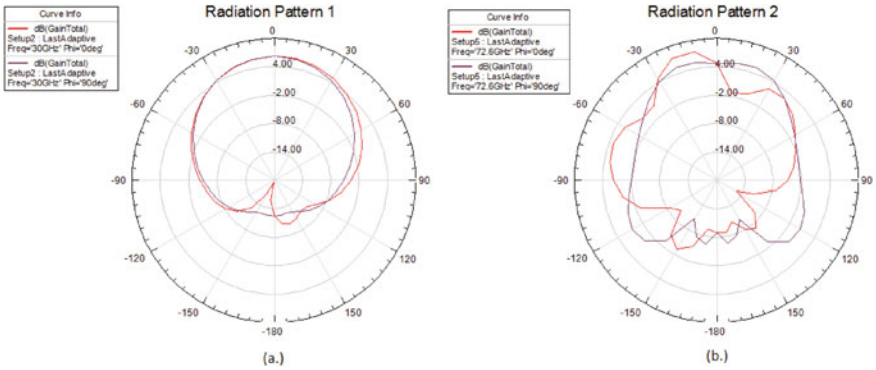



**Fig. 5** Radiation patterns at 30 and 72.6 GHz respectively of the designed antenna

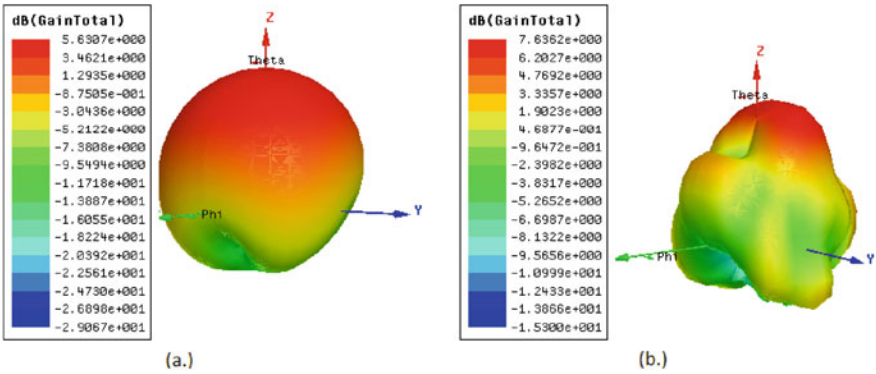



**Fig. 6** 3-D Gain plot of the designed antenna at 30 and 72.6 GHz

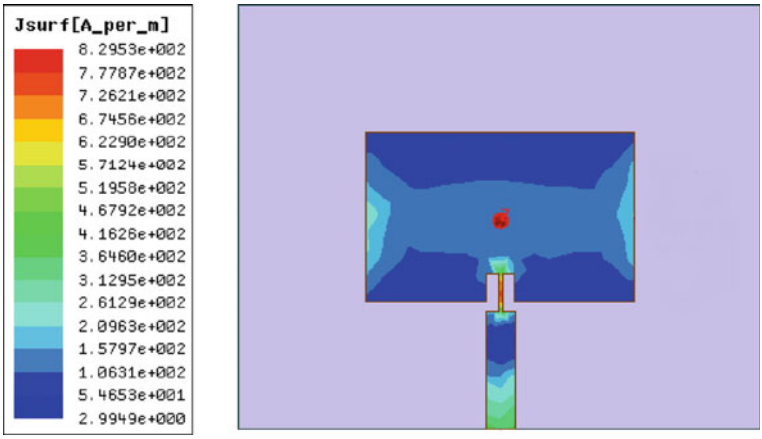



**Fig. 7** Current distribution is a qualitative measure of how current is flowing in the antenna

any hassle [27, 28]. The communications of voice and data service commercially are seen to happen through two stations, one on earth and one in the space. Cell phone transceivers are used in mobile which broadcasts and receives all cell phone communications [29]. Most of this mobile satellite behaves like a mobile tower and forwards all the calls to the nearest tower [30]. The mobile station creates high frequencies that can be transmitted to great distances by which calls can be received from any place in the world. Military and airborne personnel and even people who travel abroad often use mobile satellite systems.

To maintain communication within a warzone or a remote village, these systems can be mounted on aircraft and then the communication becomes much easier [31, 32] and satellite can be used at any part of the world and still maintain constant cell phone function for several cell phones in an area. These systems are comparatively cheap and no purchase of license is required [33]. Similarly, the xed communication uses geosynchronous satellite for TV, radio broadcasting purposes. The xed satellite has relatively low power output, they generally require large dish antennas for the reception/function. The power used by xed satellite is lower than the power requirement for satellites which broadcast directly. C-band, Ku- band and Ka-band are the three bands that a xed satellite can work on [34]. They can provide the best performance result in military organizations, small-large enterprises and other end-to-end users. These systems can provide resourceful information to national intelligence and security forces [35] (Fig. 7).

## 5 Discussion and Conclusion

In this paper, the authors have designed a microstrip patch antenna for 5G mobile communication. The dual-band antenna works at dual frequencies of 30 and

72.6 GHz. The substrate material of Rogers RO3003(TM) with dielectric constant 3 is used. An insertion slot at feed line has been adapted to obtain good impedance matching criteria. The major design application would be 5th generation New Radio (5G-NR). The antenna has good return loss. The VSWR is in the practical range of 1 to 2. Positive gain and justified radiation pattern are the appreciable features of the antenna. The future work of the antenna would include the fabrication and verification of the obtained results in a practical environment.

**Acknowledgements** We would like to thank REVA University for providing all the necessary research infrastructure to carry out our research. We also would like to thank School of ECE, REVA University for the enduring support.

## References

1. Akpakwu GA et al (2017) A survey on 5G networks for the Internet of Things: communication technologies and challenges. IEEE Access 6:3619–3647
2. Rost P et al (2017) Network slicing to enable scalability and flexibility in 5G mobile networks. IEEE Commun Mag 55(5):72–79
3. Mavromoustakis CX, Mastorakis G, Batalla JM (eds) (2016) Internet of Things (IoT) in 5G mobile technologies. vol 8. Springer, Cham. https://doi.org/10.1007/978-3-319-30913-2
4. Giordani M, Mezzavilla M, Zorzi M (2016) Initial access in 5G mmWave cellular networks. IEEE Commun Mag 54(11):40–47
5. Khan I, Ali T, Devanagavi GD, Sudhindra KR, Biradar RC (2018) A multiband slot antenna loaded with stubs for WLAN/WiMAX/Satellite TV applications. Adv Electromagn 7(5):74–81
6. Ali T, Aw MS, Biradar RC (2018) AA compact bandwidth enhanced antenna loaded with SRR for WLAN/WiMAX/satellite applications. Adv Electromagn 7(4):78–84
7. Khan I, Ali T, Devanagavi GD, Sudhindra KR, Biradar RC (2019) A compact multiband band slot antenna for wireless applications. Internet Technol Lett 2(6):e94. Feb 2019
8. Punith S, Praveenkumar SK, Jugale AA, Ahmed MR (2020) A novel multiband microstrip patch antenna for 5G communications. Procedia Comput Sci 171:2080–2086. https://doi.org/10.1016/j.procs.2020.04.224
9. Monti G, Corchia L, De Benedetto E, Tarricone L (2016) Wearable logo-antenna for GPSGSM-based tracking systems. http://dx.doi.org/10.1049/iet-map.2015.0774. Sept 2016
10. Shikder K, Arifin F (2016) Extended UWB wearable logo textile antenna for body area network applications. In: 2016 5th international conference on informatics, electronics and vision (ICIEV), pp 484–489. IEEE
11. Swain BR, Sharma AK (2019) An investigation of dual-band dual-squarering (DSR) based microstrip antenna for WiFi /WLAN and 5G-NR wireless applications. Progress Electromagn Res M 86:17–26. https://doi.org/10.2528/PIERM19060501
12. Shen X, Liu Y, Zhao L, Huang G-L, Shi X, Huang Q (2019) A miniaturized microstrip antenna array at 5G millimeter-wave band. IEEE Antennas Wirel Propag Lett 18(8):1671–1675. https://doi.org/10.1109/LAWP.2019.2927460
13. Mohammed ASB et al (2019) A review of microstrip patch antenna design at 28 Ghz for 5G applications system. Int J Sci Technol Res 8:341–352
14. Imran D et al (2018) Millimeter wave microstrip patch antenna for 5G mobile communication. In: 2018 international conference on engineering and emerging technologies (ICEET), pp 1–6. https://doi.org/10.1109/ICEET1.2018.8338623
15. Rafique U, Khalil H, Saif-Ur-Rehman (2017) Dual-band microstrip patch antenna array for 5G mobile communications. In: 2017 Progress in electromagnetics research symposium - fall (PIERS - FALL), pp 55–59. https://doi.org/10.1109/PIERS-FALL.2017.8293110

16. Ahmad W, Khan WT (2017) Small form factor dual band (28/38 GHz) PIFA antenna for 5G applications. In: 2017 IEEE MTT-S international conference on microwaves for intelligent mobility (ICMIM), pp 21–24. https://doi.org/10.1109/ICMIM.2017.7918846
17. Misbah Un Noor ES Khattak MI (2019) Design and analysis of dual-band microstrip patch antenna array for 5G cellular communication networks with improved radiation characteristics. In: 2019 second international conference on latest trends in electrical engineering and computing technologies (INTELLECT), pp 1–6. https://doi.org/10.1109/INTELLECT47034.2019.8954981
18. Sanz-Izquierdo B, Huang F, Batchelor JC (2006) Dual-band button antennas for wearable applications. In: 2006 IEEE international workshop on antenna technology: small antennas and novel meta materials (IWAT), pp 132–135
19. Bist S, Saini S, Prakash V, Nautiyal B (2014) Study the various feeding techniques of microstrip antenna using design and simulation using CST microwave studio. Int J Emerg Technol Adv Eng 4(9):318–324
20. Mahmud MdS, Shuvashis D (2012) Design, performance and implementation of UWB wearable logo textile antenna. In: 2012 15 international symposium on antenna technology and applied electromagnetics. IEEE
21. Monti G, Corchia L, Tarricone L (2013) Fabrication techniques for wearable antennas. In: 2013 European radar conference. IEEE
22. Kumar V (2017) Logo based dipole antenna for RFID applications. In: 2017 international conference on energy, communication, data analytics and soft computing (ICECDS), pp 3889–3891. IEEE, August 2017
23. Jyosthna R, Akshay Sunny R, Jugale AA, Ahmed MR (2020) Microstrip Patch Antenna Design for Space Applications, pp 406–410. https://doi.org/10.1109/ICCSP48568.2020.9182250
24. Saha P, Mandal B, Chatterjee A, Parui SK (2016) Harmes Paris logo shaped wearable antenna for multiband applications. In: 2016 Asia-pacific microwave conference (APMC), pp 1–3. IEEE
25. Tak J, Choi J (2015) An all-textile Louis Vuitton logo antenna. IEEE Antennas Wirel Propag Lett 14:1211–1214
26. Rashmitha R, Niran N, Jugale AA, Ahmed MR (2020) Microstrip patch antenna design for fixed mobile and satellite 5G communications. Procedia Comput Sci 171:2073–2079. https://doi.org/10.1016/j.procs.2020.04.223
27. Lee KF, Luk KM, Lai HW (2017) Microstrip patch antennas. World Scientific
28. Ghosh T et al (2016) Mutual coupling reduction between closely placed microstrip patch antenna using meander line resonator. Progress Electromagn Res 59:115–122
29. Zhang X, Zhu L (2016) High-gain circularly polarized microstrip patch antenna with loading of shorting pins. IEEE Trans Antennas Propag 64(6):2172–2178
30. Smyth BP, Barth S, Iyer AK (2016) Dual-band microstrip patch antenna using integrated uniplanar metamaterial-based EBGs. IEEE Trans Antennas Propag 64(12):5046–5053
31. Sanil N, Venkat PAN, Ahmed MR (2018) Design and performance analysis of multiband microstrip antennas for IoT applications via satellite communication. In: 2018 second international conference on green computing and Internet of Things (ICGCIoT), Bangalore, India, pp 60–63
32. Prahlad PM, Rakesh V, Ahmed MR (2018) Design of dual-band microstrip antenna for WiMax and X band applications. In: 2018 second international conference on green computing and Internet of Things (ICGCIoT), Bangalore, India, pp 598–602
33. Prahlad R, Kandakatla A, Ahmed R (2018) Design and performance analysis of dual-band microstrip patch antennas for smart apparels. In: 2018 Second international conference on green computing and Internet of Things (ICGCIoT), Bangalore, India, pp 573–576
34. Prahlad N, Sanil P, Naga Venkat A, Ahmed MR (2018) Design of an U shaped slotted patch antenna for RFID vehicle identification. In: 2018 second international conference on green computing and Internet of Things (ICGCIoT), Bangalore, India, pp 300–304
35. Prasanna M, Rakesh V, Ahmed MR (2018) Multiband circularly polarized microstrip reader antenna for RFID applications. In: 2018 second international conference on green computing and Internet of Things (ICGCIoT), Bangalore, India, pp 64–67

# Dynamic Gradient Sparsity Based Image Registration and Fusion Technique for Satellite Images


**Anil Naragonahalli ShambuGowda and Chandrappa Dasanapura Nanjundaiah**

**Abstract** Image registration (IR) is the basic preprocessing step that is widely used for performing different operations such as digital surface modelling, satellite image fusion, image mosaicking and so on, where pixel information from multiple images is merged into a single image for better representation. Existing image registration and fusion techniques achieve good spectral quality, but the spatial quality of the fused image is poor. This paper presents a Dynamic Gradient Sparsity based Image Registration and Fusion (DGS-IRF) technique which can retain good spectral and spatial features. The DGS-IRF technique performs a registration and fusion operation at the same time, which includes an inherent correlation of different pixels to improve registration accuracies under extreme intensity variations/distortions. Different qualitative measurement metrics such as correlation coefficient, standard deviation and root mean square error are used for validation of the DGS-IRF technique.

**Keywords** Correlation coefficient · Image fusion · Image mosaicking · Image registration · Intensity variations

## 1 Introduction

Image registration refers to aligning two or more image point sets that share the same scene in a common coordination system. It is widely used in remote sensing applications, including multi-spectral classification, environment surveillance, super-resolution reconstruction, change detection and outer space exploration [1, 2, 16]. With the advancement of high resolution (HR) hyperspectral data, registration of high resolution hyperspectral data collected from satellites possesses specific difficulties.

A. Naragonahalli ShambuGowda (✉) · C. Dasanapura Nanjundaiah
Department of ECE, East West College of Engineering Yelahanka, Bangalore 560064, India
e-mail: anilns@ewce.edu.in

C. Dasanapura Nanjundaiah
e-mail: dnchandrappa@sjbit.edu.in

C. Dasanapura Nanjundaiah
Department of ECE, SJB Institute of Technology, Bangalore 560060, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_38

The spatial data from HR hyperspectral data collected from satellites is progressively abundant in nature. Since the similarities between feature sets closeness among close to land object/item sets become increasingly obvious, feature matching is simply effected by the feature sets likeness. As a result, it will bring down the registration accuracy and affect the quality of the satellite image fusion [3]. Furthermore, HR hyperspectral data from satellites has a wide width and a high spatial resolution. As a result, it is difficult to process and store these images. Satellite image registration (SIR) is formed by considering a set of images where one image is selected as the reference image and the rest of the images are used as sensed/input images. Every sensed image is aligned with respect to the reference image in IR to establish the relationship between different pixels between image pairs and compute spatial feature variations between the input satellite images and the reference satellite image. However, in this work, the satellite image registration operation is done using one set of input and reference images. Existing SIR methodologies are categorized into Fourier-Melin transformation, dense-based and sparsity-based methods. The Fourier-Melin transformation-based methodology [4] is generally utilized for SIR with similarity-based transformation methods.

The paper is arranged as follows. The Sect. 2, discusses about various existing content image registration and fusion technique. In Sect. 3, dynamic gradient sparsity based image registration and fusion technique is presented. In Sect. 4, DGS-IRF performance is validated with state-of-art models. Finally, the research is concluded with future research direction.

## 2 Related Works

The sparsity-based methods extract features and matches identical (i.e., distinctive) features obtained from input and reference image; after that computes spatial variation among input and reference images in accordance to matched feature sets. The distinctive features such as point feature sets (i.e., gravities of regions, line intersections, and corners) and line features (i.e., edges) can be utilized for performing registration. Corner feature set are widely used feature for performing image registration, which can be extracted automatically or manual using KAZE, Oriented FAST and Rotated BRIEF (ORB) [5], Speeded-Up Robust Features (SURF), Scale-Invariant Feature Transform (SIFT), Features from Accelerated Segment Test (FAST), Harris [6] Local Global-SIFT (LG-SIFT) descriptor [7] etc. LG-SIFT based SIR methodologies that aim to minimize texture resemblance. Many existing SIR methods [7, 8], and highly rely on SIFT, or improved version of SIFT for achieving improved registration accuracies, because it is very efficient with respect to rotation and scaling [9]. In CNN-based [10] SIR method is presented, which works well for multi-temporal data and achieves robust registration performance by generating robust descriptor by progressively cumulating the inliers. However, doesn't work when image exhibit high transformation and achieves poor results considering complicated distortions.

In contrast to sparsity-based methodology, the dense-based methodology instead of detecting feature from image pair they directly establish feature by searching through spatial transformation of entire pixels of image pairs. Similarity metrics are modelled for quantifying the dependencies among image pairs such as MI (Mutual Information), Normalized Cross-correlation Coefficient (NCC), Spearman's Rho, Peak Signal to Noise Ratio (PSNR) and Root-Mean-Squared Error (RMSE). However, dense-based methodology is not efficient under illumination variation. Thus, achieves very poor registration and fusion outcomes [11]. For improving image fusion performance in presented optimal filter methodology; the model can improve spectral and spatial quality by minimizing tradeoffs function. In comparison with compressive sensing methodology the multi-resolution analysis (MRA) methodology can preserve better spectral information, however spatial information obtained not guarantees desired quality [12]. Recently, Deep learning has been applied for fusion operation, employed generative adversarial networks (GANs). Deep learning-based approach provide good fusion outcomes; however, requires huge amount of training samples, thus are not flexible in nature. Model-based methodology has achieved very good result in preserving spatial and spectral quality. In variational methodology namely P + XS image fusion, employed threshold function considering certain theory in order to obtain desired fusion quality [11]. The fusion operation is modelled as convex optimization problem using dynamic gradient sparsity and least square fit-ting, the model achieved good performance of preserving both spatial and spectral quality [12]. Although existing fusion model achieves good performance, still improving spatial quality and preserving spectral information [13, 14] is a challenging task in performing fusion of satellite images.

This work develops a dynamic gradient sparsity based image registration and fusion model. First we consider the fused image after performing down sampling to be closer to low resolution satellite image in order to keep spectral information. Then dynamic gradient sparsity feature is identified to improve spatial quality. Thus, achieve accurate registration even under severe intensity distortions. Further, the novelty of pro-posed work is, it considers correlation among different bands (no prior work has considered this). Further, optimize the cost using threshold function to solve computational complexity problems [14] for obtaining scalable performance for large dataset.

*The significance of using DGS-IRF Technique is described below*

The DGS-IRF will extract useful feature both spatial and as well as spectrally even with presence of severe distortion. Experiment conducted using standard dataset to improve overall image quality and reduce spectral distortion.

The DGS-IRF technique attain better RMSE (Root mean square error), correlation coefficient, and standard deviation outcome than existing image registration and fusion model; thus, proves the DGS-IRF can preserve sharp object boundaries on standard real-time datasets with pre-registration errors.

## 3 Methods: Dynamic Gradient Sparsity Based Image Registration and Fusion Technique

This section presents dynamic gradient sparsity (DGS) based image registration and fusion (IRF) technique for Satellite Images. The architecture of DGS-IRF technique is shown in Fig. 1. The model is designed to obtain good correlation and storage efficiency performance. This work consider that the down sampled fused image will be as similar to the original source satellite image, which is expressed as a least-square fitting for retaining its spectral information. Using geographical correlation among input reference and fused image, dynamic gradient sparsity features are identified, established, and used for enhancing spatial qualities. Using DGS aid in enhancing registration accuracies under extreme intensity variations/distortions. Further, the proposed satellite registration and fusion technique includes inherent correlation of different pixels which no prior work has used in literature.

### *3.1 Dynamic Gradient Sparsity Features*

For optimizing the fusion operation this paper presents an improved method using proximity gradient method. Particularly, this work optimizes the sub-problem in an efficient manner using gradient descent algorithm with backtracking and shrinkage thresholding methodologies [9] in an iterative manner. The proposed registration and fusion technique is scalable with respect to satellite benchmark as the model in

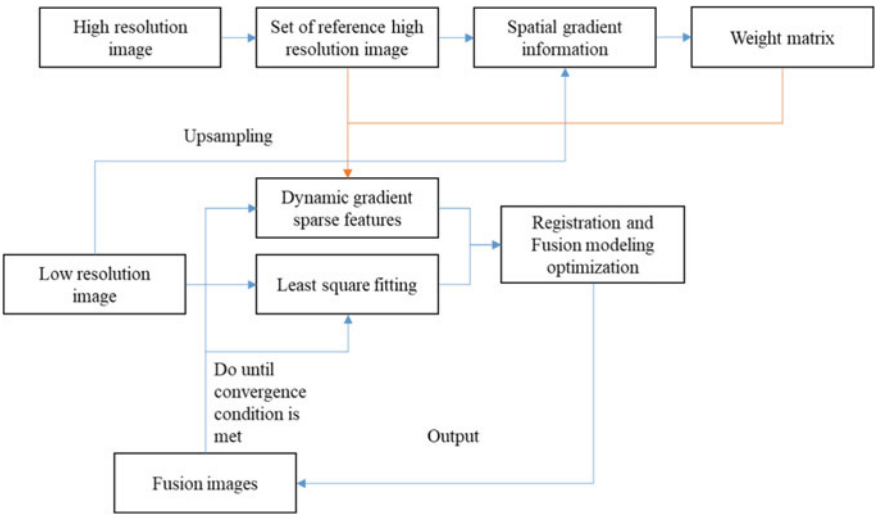



**Fig. 1** The architecture of dynamic gradient sparsity based image registration and fusion technique for satellite images

different iteration remains within linear computational complexities. Our improved satellite image registration and fusion technique achieves very good histogram performance through experimental study which shows its efficiency in improving image quality and reducing spectral distortions considering different metrics. Thus, improved satellite image registration and fusion technique can retain sharp object boundaries even with presence of error during pre-registration.

## *3.2 Least Square Fitting*

Let's assume a satellite image with low resolution of $N \in S^{\frac{n}{d}*\frac{o}{c}*t}$ and high resolution satellite image of $Q \in S^{n*o}$, where $o$ depicts number of column and $n$ depicts number of rows, $d$ represent a constant, and $t$ depicts band size. Majority of existing image registration and fusion technique up-sample the low resolution satellite image; then from the up-sampled image the model extract the spectral information. Nonetheless, the up-sampled low resolution is blurry and with less accuracy. So, this method considers the fused image only when down-sample image is closer to sensed low resolution image. For establishing correlation among the images least square fitting is used as shown in Eq. (1).

$$F_1 = \frac{1}{2}\|\beta Y - N\|_G^2, \tag{1}$$

where $\beta$ represent down-sampling function which is used to reduce spectral distortion outcomes. Local spectral information is retained for each low resolution satellite image pixel. Minimizing $F_1$ is very difficult because of low under-sampling rate and estimating $Y$ accurately is practically difficult, with no knowledge of prior information. However, high resolution satellite image provide prior knowledge such as clear edge features of terrestrial objects as it gives good geographical correlation among fused image $X$.

Existing methodologies aimed at mathematically modelling to capture such correlation. Very few methodologies have attained good performance in capturing such correlation features. The remote sensing image are generally composed of non-zeros with respect to edges and gradients incline to be sparse in nature as they often piecewise smooth. Along with, when they are well aligned, the edges position must be identical with respect to high resolution satellite image. Thus, it can be seen the sparse nature is dynamic in nature with respect to reference satellite image. This nature in the work is represented as dynamic gradient sparse (DGS) features.

### 3.3 Registration and Fusion Modelling Optimization

An improved fusion function for modelling DGS and group sparse property in simultaneous manner as shown in Eq. (2)

$$F_2 = \frac{\left\|\delta Y - \delta\overline{Q}\right\|_{2,1},}{\sum_j \sum_k \left(\delta_r Y_{j,k,e} - \delta_r Q_{j,k}\right)^2}, \tag{2}$$

where $r = 1, 2$ and $\overline{Q}$ and $Q$ depicts replicating $Q$ to $t$ bands. An important thing to be noticed here when $Q = 0$ (i.e., when reference satellite image is not present), the outcome is indistinguishable with respect to vector total variation (VTV). By merging Eq. (1) and Eq. (2) the satellite image fusion problem is described in Eq. (3).

$$\min_y \left\{ F(Y) = F_1 + \mu F_2 = \frac{1}{2}\|\beta Y - N\|_G^2 + \mu\left\|\delta Y - \delta\overline{Q}\right\|_{2,1} \right\}, \tag{3}$$

where $\mu$ depicts positive parameter value. The satellite image fusion methodology using Eq. (3) prerequisite high satellite image registration accuracies among high resolution satellite image and the low resolution satellite image, and removing false alignment is practically not good in preprocessing phase. For addressing such issues, this work registers the satellite images concurrently during the fusion process. First, for improving accuracies of satellite image registration the low resolution satellite image is sharpened to higher resolution. Second, for eliminating false alignment during fusion process the work run the two process in iterative manner till it converges. The proposed methodology can keep spatial information using DGS and false alignment will lead increase in sparse features nature of gradient. Thus, DGS is utilized for measuring similarities and the fusion function is refined for concurrent registration and fusion using Eq. (4).

$$F(Y, U) = \frac{1}{2}\|\beta Y - N\|_G^2 + \mu\left\|\delta Y - \delta\overline{Q}(U)\right\|_{2,1}, \tag{4}$$

where $U$ depicts the transformation that must be computed. The Eq. (4) describes cumulative cost of all pixels, a small value will give transformation outcomes without any overlapping of two satellite image. For avoiding small value this work rather uses gradient descent method for decreasing the cost value. This aid in avoiding wrong transformation of satellite image in iterative manner. Further, the model achieves good correlation coefficient performance with minimal storage space which is experimentally shown in next section.

## 4 Results

This section conducted experiment analysis for evaluating performance of proposed IR and fusion method over existing IR method [14]. For evaluating storage performance of image registration and fusion method correlation coefficient of histogram information is considered. Experiment is conducted on standard dataset obtained from [15]. The images are composed of both low resolution satellite image and high resolution satellite images of different bands. The input satellite images and corresponding histogram levels are shown in Fig. 3a, 4a, and 5a, the reference satellite images and its corresponding histogram level are shown in Fig. 3b, 4b, and 5b. The registered image outcomes are shown in Fig. 2. The fused satellite images and its corresponding histogram levels are shown in Fig. 3c, 4c, and 5c. From Fig. 3c, 4c, and 5c it is observed that the proposed DGF-IRF method achieves higher registration and fusion accuracies by retaining and combining large number of information from source and reference image into fused image.

In Table 1, comparative analysis of image registration outcomes achieved using proposed DGS-IRF and existing image registration model such as PR-GLS [16],

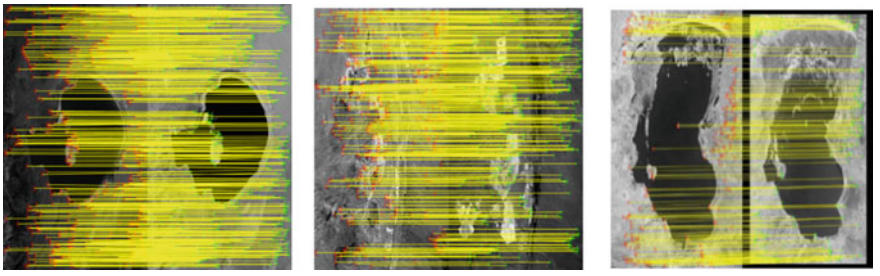


**Fig. 2** Standard satellite images used for carrying image registration

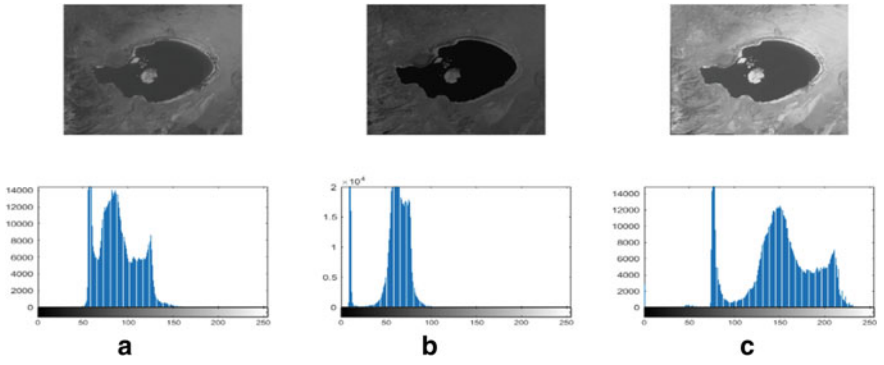



**Fig. 3** **a** Input image and its corresponding histogram **b** Reference image and its corresponding histogram **c** Fused image and its corresponding histogram

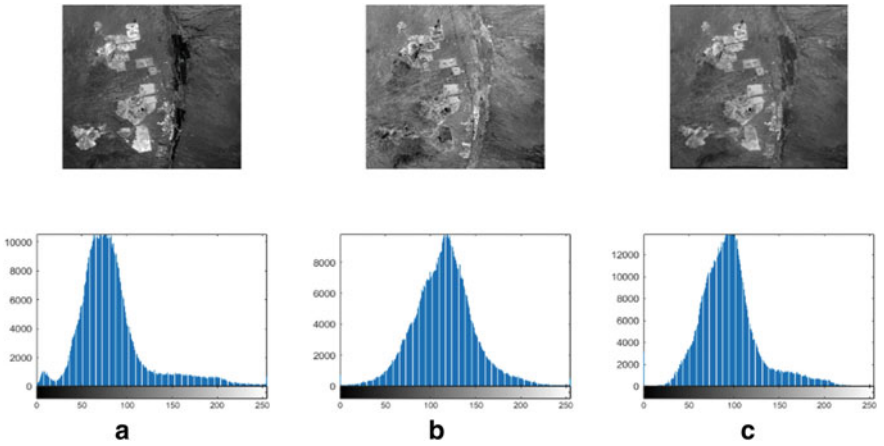



**Fig. 4** **a** Input image and its corresponding histogram **b** Reference image and its corresponding histogram **c** Fused image and its corresponding histogram

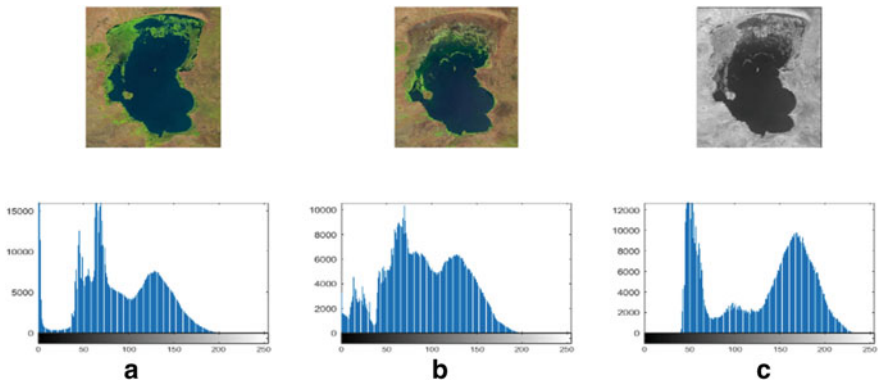



**Fig. 5** **a** Input image and its corresponding histogram **b** Reference image and its corresponding histogram **c** Fused image and its corresponding histogram

**Table 1** Comparative analysis of image registration outcomes using DGS-IRF and existing IRF methodologies

| Models | RMSE | Standard deviation |
|---|---|---|
| GL-CATE | 11.82 | 4.92 |
| PR-GLS | 9.36 | 7.17 |
| IR-CNN | 9.88 | 5.23 |
| DGS-IRF | 7.74 | 3.84 |

IR-CNN [10], and GL-CATE [17]. RMSE and standard deviation are metrics used for analyzing image registration outcomes. From Table 1, we can see DGS-IRF achieves much better RMSE and standard deviation outcomes in comparison with PR-GLS [16], IR-CNN [10], and GL-CATE [17]. Thus, will aid in improving quality of reconstructed image post fusion operation which is shown in Figs. 3, 4 and 5.

Further, the fusion outcome achieved by DGS-IRF model is compared with satellite image fusion based on modified central force optimization (MCFO) with Intensity-hue-saturation (IHS), high-pass filtering (HPF), and discrete wavelet transform (DWT) to qualitatively and quantitatively assess performance. Different performance metrics such as correlation coefficients (CC), root mean square error (RMSE) and standard deviation are used for performance evaluation. The Correlation Coefficient is used to quantify the similarities among the input and reference satellite images. The correlation coefficient value closer to "1" indicates very good performance. The correlation coefficient [14] performance is computed using Eq. (5).

$$CC(x, y) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(x_{i,j} - \overline{x}\right)\left(y_{i,j} - \overline{y}\right)}{\sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(x_{i,j} - \overline{x}\right)^2 \left(y_{i,j} - \overline{y}\right)^2}} \tag{5}$$

where $\overline{x}$, $\overline{y}$ indicates mean values of $x$, $y$, respectively and $x$, $y$ are the pixel values of input and the fused images, respectively.

The Root Mean Square Error (RMSE) among each resampled input image and the fused image measures the radiance changes of the pixel values. The RMSE is a good indicator of the spectral quality of fusion assessment. The RMSE value closer to "0" indicates very good performance. The RMSE metric [14] is computed using Eq. (6).

$$RMSE = \sqrt{E\left(D_{input} - D_{fused}\right)} \tag{6}$$

where $D$ indicates the pixel values, $D_{input}$ indicates input satellite image, and $D_{fused}$ indicates respective fused images outcomes.

The standard deviation is used for determining how much variation of the data is from the mean value. Higher SD value indicates better image registration and fusion performance. The SD [14] is computed using Eq. (7).

$$SD = \sqrt{\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} |f(i, j) - \mu|^2}{MN}} \tag{7}$$

where $M$ and $N$ indicates image dimension $f(i, j)$, and $\mu$ represents the mean value of pixel intensity.

**Table 2** Comparative analysis of DGS-IRF over existing IRF methodologies

| Models | RMSE | Correlation coefficient | Standard deviation |
|---|---|---|---|
| MCFO-IHS, 2020 | 0.0194 | 0.9634 | 32.3903 |
| MCFO-HPF, 2020 | 0.0150 | 0.9999 | 31.4327 |
| MCFO-DWT, 2020 | 0.0415 | 0.9942 | 35.8937 |
| DEDN, 2021 | 0.0881 | 0.9470 | – |
| DGS-IRF [Proposed] | 0.0031 | 0.9992 | 39.9456 |

In Table 2, comparative analysis of proposed DGS-IRF over existing IRF methodologies [14, 18] such as MCFO-HIS [14], MCFO-HPF [14], MCFO-DWT [14] and Deep Encoder-Decoder Network (DEDN) [18] is presented. From Table 2, it is observed that DGS-IRF achieves much better correlation coefficient performance than MCFO-HIS, MCFO-DWT and little lesser than MCFO-HPF; However, in terms of RMSE and standard deviation the DGS-IRF achieves much superior performance than that of MCFO-IHS, MCFO-HPF, MCFO-DWT and DEDN [18]. From the results obtained, it can be stated that the DGS-IRF preserves spectral information with high spatial quality when compared with existing IRF methodologies.

## 5 Discussion and Conclusion

The proposed methodology improves spatial quality of fused image by simultaneously registering and fusion operation by incorporating gradient prior information of high-resolution reference image and the low resolution input satellite image. The proposed registration and fusion methodologies achieves significant performance when compared with state-of-art registration and fusion methodologies. Initially, using DGS obtains sharp edges directly from the high resolution satellite images, which is proven to be good. Followed by, incorporating the local neighborhood correlation among several reference images or bands by jointly considering sharpening the satellite images, where standard methodologies carry out fusion operation band-by-band. Lastly, registering high resolution satellite image and low resolution satellite image is a challenging because of its dynamic spatial resolution, the proposed DGS-IRF methodologies register these images in simultaneous manner during fusion process. Experiment outcome shows the proposed DGS-IRF methodology can effectively minimize or remove misalignment in process of preregistration, and also offer superior registration and fused image quality than the existing methodologies. Thus, the proposed DGS-IRF methodologies can be adopted in various remote sensing applications such as change detection, area mapping, classification, etc. Future work would evaluate the proposed fusion model considering more diverse satellite images and performance metric and also compare with other fusion methodologies to enhance the fusion model.

## References


1. Chen S, Li X, Zhao L et al (2018) Medium-low resolution multisource remote sensing image registration based on SIFT and robust regional mutual information. Int J Remote Sens 39(10):3215–3242
2. You Y, Cao J, Zhou W (2020) A survey of change detection methods based on remote sensing images for multi-source and multi-objective scenarios. Remote Sens 12:2460. https://doi.org/10.3390/rs12152460
3. Meher B, Agrawal S, Panda R, Abraham A (2019) A survey on region based image fusion methods. Inf. Fusion 48:119–132
4. Rabatel G, Labbe S (2016) Registration of visible and near infrared unmanned aerial vehicle images based on Fourier-Mellin transform. Int J Adv Precis Agric 17:564–587
5. Liu H, Xiao G-F (2020) Remote sensing image registration based on improved KAZE and BRIEF descriptor. Int J Autom Comput 17(4):588–598
6. Paul S, Pati UC (2021) A comprehensive review on remote sensing image registration. Int J Remote Sens 42(14):5396–5432
7. Liu Z, Wang L, Wang X, Shen X, Li L (2019) Secure remote sensing image registration based on compressed sensing in cloud setting. IEEE Access 7:36516–36526
8. Yang K, Pan A, Yang Y, Zhang S, Ong SH, Tang H (2017) Remote sensing image registration using multiple image features. Remote Sens 9:581
9. Ma J, Junjun J, Zhou H, Zhao J, Guo X (2018) Guided locality preserving feature matching for remote sensing image registration. IEEE Trans Geosci Remote Sens 56(8):4435–4447
10. Yang Z, Dan T, Yang Y (2020) Multi-temporal remote sensing image registration using deep convolutional features. IEEE Access 6:38544–38555
11. Ma J, Ma Y, Li C (2019) Infrared and visible image fusion methods and applications: a survey. Inf Fusion 45:153–178
12. Chen C, Li Y, Wei L, Huang J (2014) Image fusion with local spectral consistency and dynamic gradient sparsity. IEEE Comput Vis Pattern Recogn. https://doi.org/10.1109/CVPR.2014.347
13. Yuan Y, Huang W, Wang X et al (2020) Automated accurate registration method between UAV image and Google satellite map. Multimed Tools Appl 79:16573–16591
14. Talal TM, Attiya G, Metwalli MR et al (2020) Satellite image fusion based on modified central force optimization. Multimed Tools Appl 79:21129–21154
15. Wang L, Feng Y, Gao Y, Wang Z, He M (2018) Compressed sensing reconstruction of hyperspectral images based on spectral unmixing. IEEE J Sel Top Appl Earth Observations Remote Sens 11(4):1266–1284
16. Zhang S, Yang Y, Yang K, Luo Y, Ong SH (2017) Point set registration with global-local correspondence and transformation estimation. In: IEEE international conference on computer vision, pp 2688–2696
17. Ma J, Zhao J, Yuille AL (2016) Non-rigid point set registration by preserving global and local structures. IEEE Trans Image Process 25(1):53–64
18. Hong D, Gao L, Hang R, Zhang B, Chanussot J (2021) Deep encoder-decoder networks for classification of hyperspectral and LiDAR data. IEEE Geosci Remote Sens Lett 19:1–5. https://doi.org/10.1109/LGRS.2020.3017414

# Effectiveness Analysis of Distance Measures for Graph Coloring Based View-Construction Approach In Multiview Ensemble Learning


**Sapna Kumari, Vipin Kumar, and Aditya Kumar**

**Abstract** The classification performance of multi-view learning is better than the traditional machine learning algorithms as stated in state-of-art. In this research, using the graph coloring technique, clustering of dissimilar features has been done for view construction in multiview ensemble learning. Ten similarity measures (correlation, cosine, Euclidean, Minkowski, Manhattan, Seuclidean, Sqeuclidean, Chebyshev, Canberra, and Bray Curtis) have been applied for the graph coloring-based view-construction approach and their impact is analyzed using the k-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Naive Bayesian (NB) classifier over the eight different datasets. Based on overall performance, the research recommends the correlation distance for the view-construction method. The proposed method Graph Coloring based View-Construction methods (GC-VC) has been compared with state-of-art methods. The analysis of results and their statistical analysis show that the GC-VC method has performed significantly better than other methods.

**Keywords** Classification · Distance measure · Graph coloring · Multi-view Ensemble Learning (MEL) · View construction method

## 1 Introduction

Multi-view learning is one of the emerging areas of machine learning. It learns by partitioning the features of a dataset, where each subset of the feature set is known as view [1]. The main objective of multi-view learning is utilizing information available from multiple feature subsets (views) collectively to enhance performance. Multi-view learning has been applied successfully in many areas like dimensionality reduction [2, 3], image recognition [2], multi-view clustering [4], human gait recognition [5], etc. The quantity and quality of the views affect the performance of multi-view learning for the given task. Ensuring the quality and quantity of the views for the given data is a complex task in multi-view learning. The number of views may be

S. Kumari · V. Kumar (✉) · A. Kumar
Department of Computer Science and Information Technology, Mahatma Gandhi Central University, Motihari, Bihar 845401, India
e-mail: rt.vipink@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_39

easily identified based on their natural source of the dataset usually. But it is a difficult task to identify the suitable number of views for single-sourced data. Because there are several ways to partitioning features set to get the views. Therefore, an effective multi-view learning approach is highly required that ensure the quantity along with the quality of the views. The graph coloring method colors the vertices of a graph in such a way that no two adjacent vertices have the same color. This method can be applied directly to get the views by imposing the edge constrained on the similarity-based weighted edge in the complete graph of the features set. Several distance methods may have a different group of vertices (same color) after edge removal through suitable constrained on edge. The grouping of the features can be considered as views of the dataset for the distance method. Therefore, it is an urgent requirement to identify the appropriate distance method(s) for the graph coloring method for view construction in multi-view ensemble learning (MEL).Generally, MEL has three phases like 1) view construction, 2) view evaluation, and 3) view ensemble. In this research, the different similarity measures have been applied for graph coloring in the view construction phase of MEL. The similarity measure in graph coloring methods affects the quantity and quality of views. Therefore, the objective of this research is to identify the best similarity measure for the graph coloring-based view construction method in MEL.

In the experiments, ten different standard distances are utilized like Correlation distance, Cosine distance, Euclidean distance, Minkowski, Manhattan, Seuclidean, Sqeuclidean, Chebyshev, Canberra, Bray Curtis distance [6]. The k-Nearest Neighbors (KNN), Naive Bayesian (NB), and Support vector machine (SVM) are deployed to learn from each view of the dataset. And eight standard datasets have been used for each distance measure. The quantitative and statistical analysis show that the correlation-based graph coloring approach for view construction in multi-view ensemble learning has performed the best among other distance measures. The previously proposed GC-VC method performance has also been compared with another state-of-art method. The results analysis shows an effective performance than the other state-of-art methods.

The organization of the paper is as follows: Sect. 1 and Sect. 2 are the introductions and related works, respectively. Section 3 deals with the Taxonomy of the Research. Experiments and Schematic Review have been shown in Sect. 4 and Sect. 5 respectively. Finally, Sect. 6 presents the conclusion of the research work.

## 2 Related Works

Multi-view learning obtains the generalized performance by prediction of the same sample through multiple views of the data. The suitable view can be formed by partitioning features (vertical partition) or samples (horizontal partition). In this work, the vertical partition is done and throughout the paper, partition based on attributes is considered as a view. Multiview learning is used for the different fields in machine learning like clustering [7], supervised learning [8], semi-supervised learning [9],

dimensionality reduction [10], active learning, ensemble learning [11]. The attention task of MEL is view construction. There are various view construction methods in state-of-art. In the feature set partitioning method, a disjoint partition of feature sets is being done and each subset will represent a view [12]. A random split of features is done when there is no proper principle of view construction. In bagging, the samples of the dataset are divided into k-parts with replacements [13]. Bagging is a horizontal partition method in multi-view learning. In the attribute bagging method, first, the best partition size is determined then the feature set partition is done of that size [14]. Prediction from multiple views can be ensembled by various ensemble methods such as Stacking, Arbiter tree, combiner tree, majority voting, performance weighting, Bayesian combination, vlogging, etc. [11].

## 3 Taxonomy of the Research

Graph Coloring Based Views Construction (GC-VC) is a potential method for view constructionin multi-view ensemble learning (MEL) [3]. A successful experiment of classifier comparison has been performed and the neural network has performed well. A major limitation of GC-VC was that only a correlation similarity measure wastaken. Therefore it is required to check the performance of various other similarity measures over the GC-VC method [3]. A comparative analysis has been done to find a better similarity measure for the GC-VC method in MEL. Graph constructed using the feature as a vertex, their coloring is suitable for multi-view learning because each color class may represent the views as a subset of features. It is a new approach in the field of multi-view ensemble learning [3].

Let, the dataset is $D_{m\times n}$ where the feature set is denoted as $X = \{X_i(v)\}_{i=1}^{n}$, $v \in \text{dom}(X_i)$ and labels are $y_i \in \text{dom}(y)$. The $i^{th}$ view of the dataset can be denoted $V_i = \{X_i\}_{i=1}^{q}$; $i, q \in n$. Suppose $V_i$ and $V_j$ are the views, then $\{X_i\}_{i=1}^{p} \cap \{X_i\}_{i=1}^{q} = \varnothing$ where $i \neq j$ and $p \neq q$. If £ is the classifier then their training accuracy $i^{th}$ views $A_i = £(V_i)$.

### *3.1 Graph Coloring Based Views Construction*

This approach has the following steps to construct the views of the dataset which are as follows:

- *STEP-1 (Similarity Matrix $M_s$ of the feature of the dataset):* first of all, similarity matrix $M_s$ of the features set is obtained using the distance method.however, different types of similarity coefficients may be utilized. The similarity between the two features $X_i \in X$ and $X_j \in X$ can be understood as a distance (Eq. 1):

$$d_{i,j} = \text{distances}(X_i, X_j) \tag{1}$$

where $d_{i,j} \geq 0$ and $d_{i,j} = 0$, if $X_i = X_j$; $j \in \{1, 2, 3, \ldots n\}$. It means, distance $d_{i,j}$ can be positive or zero only. The symmetric similarity matrix of all features have the following condition as $d_{i,j} = d_{j,i}$ and $d_{i,a} + d_{j,a} \geq d_{i,j}$ show that the sum of the distance from $d_{i,a}$ and $d_{j,a}$, can not be smaller than the distance $d_{i,j.}$ and then, L2 normalization techniques are utilized in the normal similarity matrix $M_s$ in the range [0, 1]. the framework of the method is shown in (Fig. 1).

- *STEP-2 (Drawing Graph):* A complete graph G is drawn using a normalized similarity matrix, where vertices are represented by features and the weight of each edge is normalized similarity measures values. the value within two nodes represents the similarity of features correspondingly.
- *STEP-3 (Obtaining the partial graph $G_c$ from complete graph G):* The correlated features in the same views may perform like redundant features. Therefore, the information of correlated features may not be utilized at their full capacity in the same view. It is required to have uncorrelated features in the same views for generalized performance. Therefore, the edge removal is done based on threshold t, defined within the range of (0,1). The edges having weight ($d_{i,j}$) below the threshold value will be removed. This process makes the partial graph.
- *STEP-4 (Deploying graph coloring method):* The objective of this process is to assign a color to each vertex of the partial graph $G_c$ obtained from the previous step-3. Coloring of the vertex means assigning the label to it. Here, the chromatic

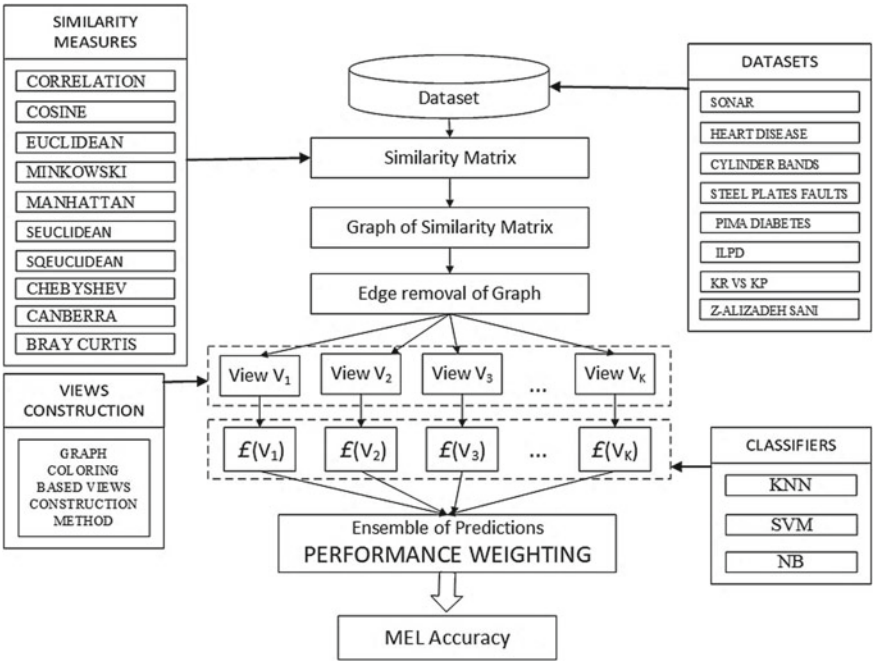



**Fig. 1** Taxonomy Diagram of view construction using graph coloring for multi-view ensemble learning

number $\chi(G_c)$ is the least number of distinct colors required to color all vertices. The vertex (feature) having the same color (label) are assigned to the same group, where the grouped vertices (a subset of feature) are considered as a view of the dataset. It means the number of views is equal to chromatic number $\chi(G_c)$, which is also ensuring that highly correlated features are not belonging to the same view, which minimizes the feature redundancy in views.

### *3.2 Evaluation of Obtained View Using the Classifier(s)*

Once we obtain the views of the dataset using the graph coloring method. Then, a classification algorithm learns from the view $V_i$. Each view $V_i$ is evaluated based on classification performance $A_i$ (accuracy) which is shown in (Eq. 2):

$$A_i = £(V_i), i \in \{1, 2, 3, \ldots \chi(G_c)\} \tag{2}$$

Then, the set of learners can be represented as $\{£(V_i)\}_{i=1}^{\chi(G_c)}$ and their respective performances are $\{A_i\}_{i=1}^{\chi(G_c)}$.

### *3.3 Ensemble of Prediction Obtained from View Learners*

The views are required to an ensemble of learner $£(V_i)$ for the collective performance which is shown in (Eq. 3):

$$C = \sum_{i=1}^{\chi(G_c)} \{w_i \times £(V_i)\} \tag{3}$$

where $\{w_i = A_i\}_{i=1}{}^{\chi}(G_c)$. The choice of threshold ($t_0$) governs the structure of the partial graph $G_c$ after edge removal from complete graph G. therefore, the chromatic number $\chi(G_c)$ is also varies according to the $G_c$. It means, obtaining the suitable threshold (t) is required to ensuring the better performance of MEL. The proposed framework repeatedly obtains the performance C for random values of threshold (t) and find the suitable threshold ($t_0$) with the effective performance of MEL which is shown in (Eq. 4):

$$t_0 = \left\{\text{argmax}_{t\in[0,1]}\{C_i, t\}_{i=1}^{k}\right\} \tag{4}$$

where k is the number of iterations to obtain the suitable threshold $t_0$ successfully. Then the learners $\{£(V_i)\}_{i=1}^{\chi(G_c)}$ obtained using threshold $t_0$ are used for predicting test datasets.

## 4 Experiments

In this experiment, eight real-world binary labeled datasets are utilized, where name of the datasets [samples × features]are as followed: Sonar [60 × 208], Heart Disease [13 × 270], Pima Diabetes [8 × 768], Cylinder Bands [39 × 512], Steel Plates faults [27 × 1941], Indian Liver Patient Dataset (ILPD) [10 × 583], King-Rook Vs king-Pawn (Kr Vs Kp) [36 × 3196] and Z-Alizadeh Sani [56 × 303] [15].

### *4.1 Experimental Setup*

All experiments have been performed on Colab (Google Collaboratory) using Python programming language. The eight real-world datasets are utilized to learn three different classifiers as KNN, NB, and SVM. Ten similarity measures (correlation, cosine, Euclidean, Minkowski, Manhattan, Seuclidean, Sqeuclidean, Chebyshev, Canberra, and Bray Curtis) are used to create the similarity matrices for each dataset. And views are constructed using a suitable threshold in the GC-VC method for each dataset. Each view of the dataset has been utilized to build the model individually with the same classifier (like KNN, NB, and SVM). The predicted class of each view is ensembled using the performance weighting ensemble method. The tenfold cross-validation has been used to get the overall prediction of each view. The threshold value is chosen out of 150 randomly generated threshold values on which MEL accuracy is maximum. Optimized parameter $k = 5$ has been used for the KNN classifier and the SVM classifier has an RBF kernel while learning the model. To obtain the ranking of different distance measures Friedman ranking test is performed. The non-parametric statistical analysis has been done with the help of an open-source tool i.e., the KEEL software tool [16].

### *4.2 Results*

The single view and MEL accuracy obtained by different distance measures for graph coloring-based view construction method using KNN, SVM, and NB are shown in Table 1. The row of the table represents the datasets corresponding to the three classifiers and the column represents different distance measures.

**Table 1** Classification performance of MEL using different distance measures for graph coloring-based view construction method

| Classifiers | Datasets | Single-View | Correlation | Cosine | Euclidean | Minkowski | Seuclidean | Sqeuclidean | Chebyshev | Canberra | Manhattan | Braycurtis |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SVM | **D1** | **0.6395** | **0.6875** | **0.6971** | **0.6634** | **0.7163** | **0.6826** | **0.7307** | **0.6875** | **0.6634** | **0.7355** | **0.6826** |
| | **D2** | **0.6604** | **0.7953** | **0.8217** | **0.8316** | **0.8184** | **0.8118** | **0.8316** | **0.8316** | **0.7887** | **0.8316** | **0.7557** |
| | **D3** | **0.7578** | **0.7578** | **0.7018** | **0.7617** | **0.7617** | **0.7617** | **0.6783** | **0.7604** | **0.7591** | **0.7617** | **0.7591** |
| | **D4** | **0.6495** | **0.6586** | **0.679** | **0.679** | **0.653** | **0.679** | **0.6512** | **0.679** | **0.679** | **0.679** | **0.679** |
| | **D5** | **0.6496** | **0.6558** | **0.6532** | **0.6532** | **0.6532** | **0.6532** | **0.6532** | **0.6532** | **0.6532** | **0.6532** | **0.6548** |
| | **D6** | **0.7135** | **0.7135** | **0.7135** | **0.7135** | **0.7135** | **0.7135** | **0.7135** | **0.7135** | **0.7135** | **0.7135** | **0.7135** |
| | **D7** | **0.9364** | **0.9455** | **0.9302** | **0.9367** | **0.9367** | **0.9367** | **0.9367** | **0.5222** | **0.9367** | **0.9367** | **0.9402** |
| | **D8** | **0.7129** | **0.7129** | **0.7128** | **0.7128** | **0.7128** | **0.7128** | **0.7128** | **0.7128** | **0.7128** | **0.7128** | **0.7128** |
| KNN | **D1** | **0.5952** | **0.673** | **0.6634** | **0.6586** | **0.6778** | **0.6634** | **0.6971** | **0.6923** | **0.6346** | **0.6875** | **0.6538** |
| | **D2** | **0.6535** | **0.759** | **0.7722** | **0.7722** | **0.7722** | **0.7755** | **0.7788** | **0.7524** | **0.7425** | **0.7722** | **0.7722** |
| | **D3** | **0.7213** | **0.7278** | **0.7122** | **0.7083** | **0.72** | **0.72** | **0.72** | **0.72** | **0.7356** | **0.72** | **0.7356** |
| | **D4** | **0.5811** | **0.6382** | **0.653** | **0.6307** | **0.6178** | **0.6307** | **0.6456** | **0.6085** | **0.6159** | **0.6363** | **0.6474** |
| | **D5** | **0.5739** | **0.6409** | **0.6537** | **0.6532** | **0.6517** | **0.6543** | **0.6558** | **0.6532** | **0.6496** | **0.6527** | **0.6563** |
| | **D6** | **0.6515** | **0.7118** | **0.7032** | **0.7118** | **0.7118** | **0.7221** | **0.7169** | **0.7049** | **0.7118** | **0.7118** | **0.7221** |
| | **D7** | **0.8022** | **0.8125** | **0.8025** | **0.8019** | **0.8019** | **0.8019** | **0.8019** | **0.5222** | **0.8019** | **0.8019** | **0.8025** |
| | **D8** | **0.6739** | **0.7359** | **0.7128** | **0.7128** | **0.7128** | **0.7128** | **0.7128** | **0.71287** | **0.7128** | **0.7128** | **0.7425** |

(continued)

**Table 1** (continued)

| Classifiers | Datasets | Single-View | Correlation | Cosine | Euclidean | Minkowski | Seuclidean | Sqeuclidean | Chebyshev | Canberra | Manhattan | Braycurtis |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NB | **D1** | **0.6071** | **0.6538** | **0.6634** | **0.6538** | **0.6442** | **0.6586** | **0.6634** | **0.6682** | **0.6346** | **0.6538** | **0.6634** |
| | **D2** | **0.805** | **0.825** | **0.8151** | **0.8184** | **0.8085** | **0.8184** | **0.8184** | **0.8085** | **0.8085** | **0.8085** | **0.8085** |
| | **D3** | **0.7564** | **0.7708** | **0.7408** | **0.7552** | **0.7552** | **0.7552** | **0.7421** | **0.7421** | **0.7643** | **0.7552** | **0.763** |
| | **D4** | **0.6518** | **0.679** | **0.6717** | **0.6567** | **0.6808** | **0.6549** | **0.6753** | **0.653** | **0.679** | **0.6808** | **0.6771** |
| | **D5** | **0.4972** | **0.6388** | **0.6532** | **0.6548** | **0.6785** | **0.544** | **0.6584** | **0.6532** | **0.6733** | **0.6537** | **0.6532** |
| | **D6** | **0.5573** | **0.7135** | **0.7135** | **0.7118** | **0.7118** | **0.7118** | **0.7118** | **0.7118** | **0.7204** | **0.7135** | **0.7118** |
| | **D7** | **0.6127** | **0.6304** | **0.6145** | **0.6279** | **0.6279** | **0.6279** | **0.6279** | **0.534** | **0.6279** | **0.6279** | **0.6223** |
| | **D8** | **0.7323** | **0.8151** | **0.7392** | **0.7425** | **0.7425** | **0.7524** | **0.7524** | **0.7194** | **0.7425** | **0.7425** | **0.7623** |

# 5 Schematic Review

## *5.1 Quantitative*

From Table 1, the classifier wise observations corresponding to distance measures are as followed:

**Performances of MEL Using SVM Classifier.** The performance of MEL using SVM classifier have better performance than single-view learning for all distance measures over six datasets, where performance on ILPD and Z-Alizadeh Sani datasets are equivalent to single-view learning. The Manhattan and Correlation distance measures are effective for MEL using SVM classifiers over 4 and 4 datasets, respectively. Therefore, both distance measures may be utilized for MEL for effective performance over single view learning.

**Performances of MEL Using KNN Classifier.** Six out of eight datasets, the KNN classifier has performed better than single-view learning overall distance measures. It has been also observed that Bray Curtis, Sqeuclidean, Cosine, and Correlation distance measures have been performed on 3, 1,1, and 2 datasets, respectively. Therefore, Bray Curtis distance measures may be recommended for the graph coloring-based view construction method for MEL.

**Performances of MEL Using NB Classifier.** MEL has performed better than single-view learning using NB classifier for all datasets, except the Pima Diabetes dataset which also shows the equivalence performance. So, the MEL approach using the GC-VC method is effective than single-view learning for SVM, KNN, and NB classifiers. The comparative analysis of distance measures for SVM, KNN, and NB classifiers shows that the Correlation distance measure has performed the best among the other nine measures. The effectiveness of distance measures for the GC-VC method has also been analyzed through nonparametric statistical analysis using the Friedman test in the next section.

## *5.2 Nonparametric Statistical Analysis Using Friedman Test*

For statistical analysis, the Friedman ranking test has been performed then post-hoc analysis got done. The ordering of distance measures has been done based on the result obtained by the Friedman ranking. The best distance, (i.e., having minimum Friedman rank value) is assigned as rank-1, second-best is assigned a rank-2. And worst distance is assigned a rank-10 similarly. In this way, ranking is done from 1 to 10 for all the distance measures. The ranking of distances is shown in (Fig. 2). using a bar graph. Friedman's raking test of n × n has been done using Keel software *[*https://sci2s.ugr.es/keel/description.php*.]* Friedman ranking is done for the data shown in Table 1. For SVM, the corresponding data from Table 1. is taken for Friedman

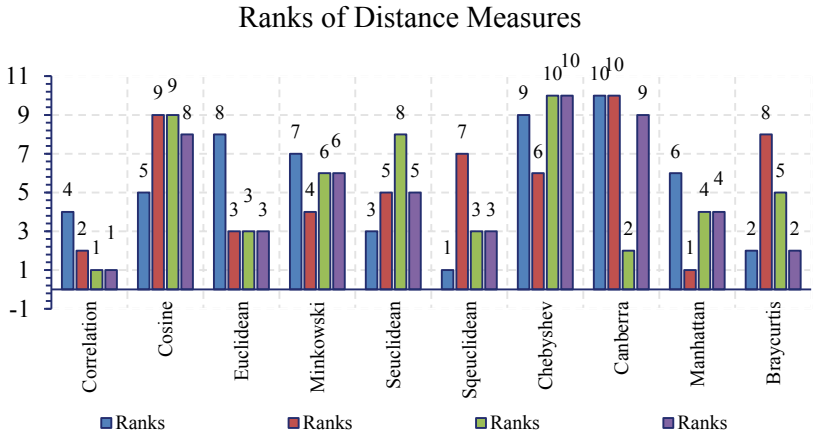



**Fig. 2** The overall ranking of ten distance measures corresponding to the classifiers used in MEL

ranking. Similarly, for NB and KNN, separately Friedman's ranking has been done. At last, the whole Table 1. has been taken for the overall ranking of all the distance measures. (Fig. 2). showsthe ranking of KNN, SVM, NB, and overall, from the left bar to right bar, respectively. After Friedman's ranking, we have done a comparison among distance measures by performing post-hoc analysis ata significance level of 0.10. From analysis, we can say that Bray Curtis distance is very dissimilar to Chebyshev distance i.e., their p-value is very lessthan 0.015. Correlation and Bray Curtis distances are highly similar (very high p-value 0.738597) based on post-hoc analysis. Therefore, it is recommended to use Correlation, then after Bray Curtis for this MELframework.

## *5.3 Comparison with Other Views Construction Methods*

The GC-VC method [3] has analyzed the effectiveness of the proposed method for single view learning. The comparison has been done on ten datasets where KNN, NB, SVM, Decision Tree (DT), and Neural Network (NN) classifiers are utilized for learning the models. The study shows that the GC-VC method has performed well on all ten-dataset using all five classifiers. The Friedman ranking of the proposed method and post-hoc analysis is also performed to provide the statistical proof for better performance of the proposed method. While the experiments, it has been observed that the distance measure is affecting the performance of the proposed method. Therefore, it is highly important to find out the best suitable distance measure for the proposed method. The extensive analysis for finding the suitable distance measure(s) was out of scope in the paper [3]. Therefore, distance measures are primarily focused on the methods GC-VC to get the suitable measures. The effectiveness of the proposed method has been analyzed using three classifiers (KNN, SVM,

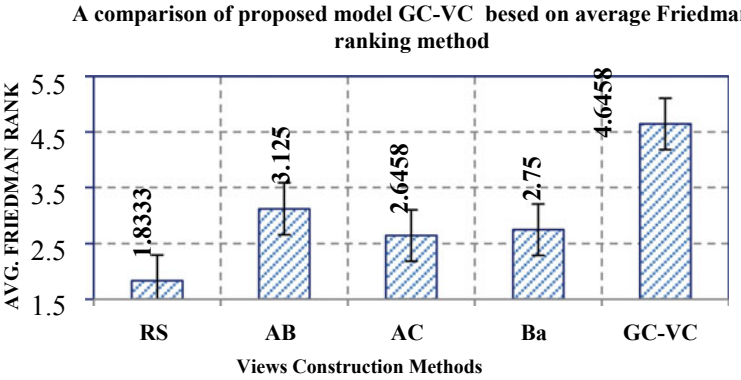



**Fig. 3** Comparison of the overall performance of GC-VC with state-of-art methods

and NB).The proposed view construction method GC-VC have been compared with a single view (Sv) as well as four states of the art view construction methods namely random split (RS) [12], Attribute Bagging (AB) [14], Attribute Clustering (AC) [16] and Bagging (Ba) [13] as per Fig. 3. For the Comparision of the potential of the GC-VC to other views construction methods, Friedman ranking has been used as shown in Fig. 4. The purposed method GC-VC using the correlation distance has significantly better performance than other views construction methods. Random split has lesser performance among all. Attribute Bagging is the second most performing model. From Fig. 3, it is evident that the GC-VC method has a better performance overall dataset for each classifier KNN, SVM, and NB.

## 6 Discussion

After doing a rigorous quantitative and statistical review of the results, the following conclusions have been made based on classifier wise distance measure recommendation for the graph coloring-based view-construction method of MEL:

- For KNN, it has been found that Sqeuclidean distance is performing best after that Bray Curtis is the second best. For NB, Correlation is ranked first then Canberra is ranked second. In the case of SVM, Manhattan has performed best than after Correlation measures. For three classifiers, Correlation is ranked first then after Bray Curtis is ranked second. Our suggestion is to use either Correlation or Bray Curtis for this MEL framework. Finally, we can conclude that correlation distance is the best performing distance for all 3 classifiers.
- The graphical representation of a comparison of the proposed method GC-VC (Fig. 3) shows that the method performances are effective overall dataset corresponding to the KNN, SVM, and NB classifiers. Further, the performances of

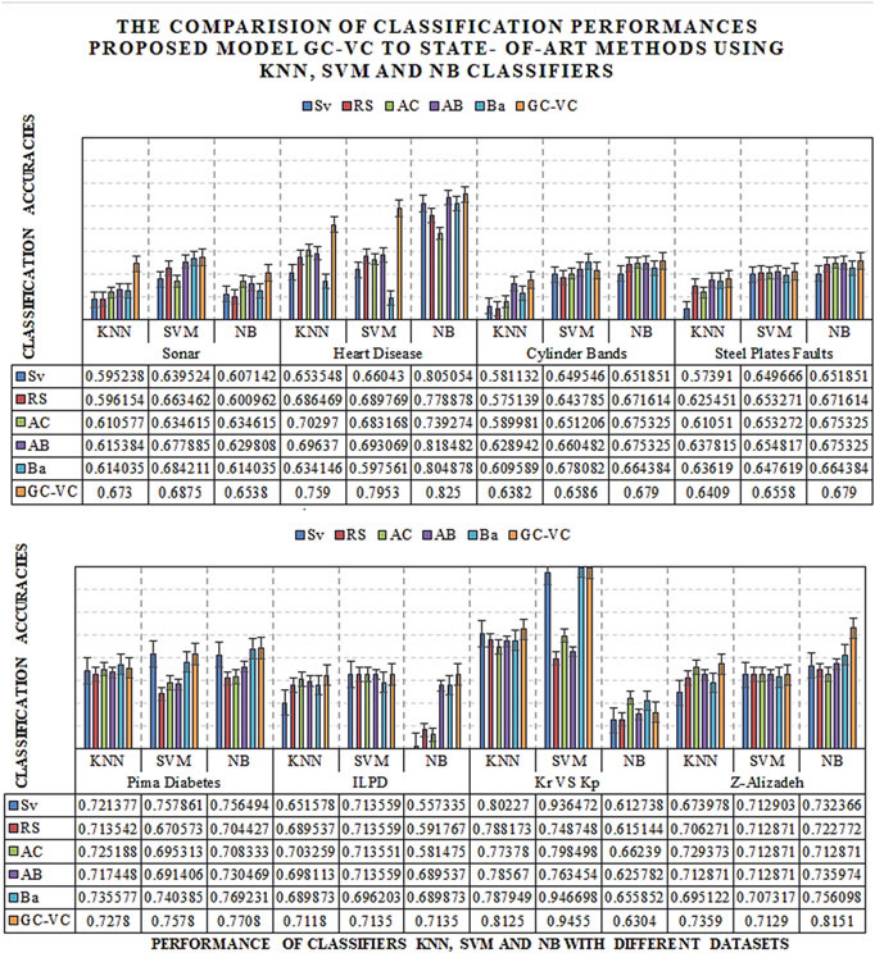



**Fig. 4** Comparison of MELclassification accuracy of GC-VC with state-of-art methods

the GC-VC method have been tested using the Friedman ranking test (shown in Fig. 4), where the method has the highest rank than other state-of-art methods.

- The above discussion concludes that the research work has successfully identified the best distance method i.e., Correlation or Bray Curtis for the GC-VC method and it is effective over other state-of-art methods.

## 7 Conclusion

The impact of different distance measures for graph coloring-based view construction in MEL is analyzed. Selecting suitable distance measures will significantly improve the performance of MEL. For each dataset and classifier, the classification performance of MEL is found to be better than a single view. To obtain the ranking of different distances, rigorous statistical analysis has been done. It has been concluded that the overall correlation distance measure has performed the best on the graph coloring-based view-construction approach for MEL. This paper can help in selecting the best distance measure for multi-view ensemble learning. It has also identified that the GC-VC method has better performance than the other state-of-art methods using correlation methods for KNN, SVM, and NB classifiers where the Friedman ranking of the methods for the comparison shows that the GC-VC method is better than other comparative methods. The better classification performance may be achieved by optimizing the threshold which can be analyzed in future research. Only ten similarity measures have been used, further, the work can be extended to more similarity measures. A further experiment can be performed on some high-dimensional data using more classifiers.

## References

1. Kumar V, Minz S (2015) Multi-view ensemble learning: an optimal feature set partitioning for high-dimensional data classification. Knowl Inf Syst 49(1):1–59
2. Hou C, Zhang C, Wu Y, Nie F (2010) Multiple view semi-supervised dimensionality reduction. Pattern Recogn 43(3):720–730
3. Kumar A, Kumar V, Kumari S (2021) The graph coloring based view construction method for multiview ensemble learning. In: 2nd International Conference on Secure Cyber Computing and Communication (ICSCCC-2021). IEEE Xplore
4. Tzortzis G, Likas A (2012) Kernel-based weighted multi-view clustering. In: IEEE 12th International Conference on Data Mining
5. Deng M, Wang C, Chen Q (2016) Human gait recognition based on deterministic learning through multiple views fusion. Pattern Recognit Lett 78(C):56–63
6. Deza E, Deza MM (2009) Encyclopedia of Distances, p 94. Springer, Heidelberg. https://doi.org/10.1007/978-3-642-30958-8
7. Mandru DB, Krishna YS (2018) Multi view cluster approach to explore multi objective attributes based on similarity measure for high dimensional data. Int J Appl Eng Res 13
8. Sun S (2013) A survey of multi-view machine learning. Neural Comput Appl (7–8):2031–2038
9. Sun S, Jin F, Tu W (2011) View construction for multi-view semi-supervised learning. In: Liu D, Zhang H, Polycarpou M, Alippi C, He H (eds.) Advances in Neural Networks–ISNN 2011. ISNN 2011. LNCS, vol 6675, pp 1–7. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-21105-8_69
10. Gönen M, Gönen GB, Gürgen F (2014) Bayesian multiview dimensionality reduction for learning predictive subspaces. In: ECAI 2014–21st European Conference on Artificial Intelligence
11. Rokach L (2010) Pattern classification using ensemble methods. World scientific
12. Ho TK (1998) The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 20(8):832–844

13. Breiman L (1996) Bagging predictors. Springer 24:123–140
14. Bryll R, Gutierrez-Osuna R, Quek F (2003) Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets. Elsevier 36(6):1291–1302
15. UCI machine learning repository. http://WWW.ics.uci.edu/~mlearn/databases. Accessed 18 Jan 2021
16. Di W, Crawford M (2012) View generation for multi-view maximum disagreement based active learning for hyperspectral image classification. IEEE Trans Geosci Remote Sens 50(5)

# Efficient Square Root Computation–An Analysis


**A. Sai Prasanna, J. Tejeswini, P. Keerthana, P. Yamini Raghavi, and J. P. Anita**

**Abstract** In this paper an analysis on the advantages and disadvantages of the common algorithms adopted for designing square root calculators on FPGA chips is performed. A comparison of methodology for square root calculation using the CORDIC with Non-Restoring Square root algorithm is presented. This comparison is used to make the computation of complex square roots more efficient as square root calculation is an integral part of the same. The architecture is designed using Verilog HDL language and implemented using Modelsim for simulation and the synthesis is realized using Zed board on Xilinx Vivado. Experimental results demonstrate better performance in terms of power for the case of hyperbolic CORDIC when compared with the non - restoring method.

**Keywords** Complex square root · CORDIC · Non-restoring · Square root · Verilog

## 1 Introduction

The need for square root calculation is seen in algorithms used for Interactive media applications like DSP and image processing, scientific computation and graphics. The computation effort required for square root calculation and the computational time is high. As it is an integral part of most algorithms, there is a need to reduce its computational complexity. Various algorithms have been put forth, namely Newton Raphson method as well as sequential algorithm (digit-by-digit method).

In the works of Li et al. [1], the authors examine the non-restoring square root algorithm being implemented on the FPGA. Although, the non - restoring algorithm architecture increases the area cost as it uses Carry select adder and add/sub units that cause unwanted increase in area for input data with large bit-width.

A. S. Prasanna · J. Tejeswini · P. Keerthana · P. Y. Raghavi · J. P. Anita (✉)
Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, India
e-mail: jp_anita@cb.amrita.edu

Amrita Vishwa Vidyapeetham, Coimbatore, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_40

The next aspect of the work is complex numbers. These numbers have been used for applications such as real-time system modelling and data representation, that includes concepts such as communication systems, electronic circuits and electromagnetism that form a basis to electronics and communication systems [2–4]. The real valued square-root computation mentioned earlier would not be suitable for compute complex square roots as it would require additional hardware. To overcome this coordinate rotation digital computer (CORDIC) [5] which has lower design complexity and higher performance compared with conventional architectures are presented [6].

In this literature, a comparison of methods involving CORDIC to compute complex square roots is studied and analysed. Section 2 describes the literature research carried out to understand various methods used for square root computation. Section 3 discusses the proposed work of computing complex square roots and the involved concepts. In Sect. 4 the results are provided. Section 5 provides the discussion of the results. Finally, the conclusion of the work is described in Sect. 6.

## 2 Related Works

There are different types of methods used in the implementation of square root function. Among those some of them are discussed in detail below.

### *2.1 Digit Recurrence Method*

This method is mainly used for division computation which is used in multimedia and graphics applications. Here, the only thing that was done was a sequential implementation of the combined division, square root algorithm. It is suitable for any base, and the way it proceeds depends on the base selected.

This algorithm uses table-lookups, limited-precision multipliers, adders. The design generated using this algorithm, is expected to have a smaller area, longer cycle time and also a desirable power dissipation which can be used for low-power applications. It is appropriate for implementation with a higher radix. It gives us the execution time, estimates of delays and cost when hardware implementation is done.

It cannot guarantee a fixed VLSI implementation. It is simple but practically tough. It has increased Hardware complexity and Latency when compared to other methods [7].

## 2.2 Newton Raphson Method

It's a rapid method to perform square root operations. This algorithm requires multiplication and either addition or subtraction during each iteration often implemented using Wallace tree adder. This step is succeeded by carry propagate adder. This procedure is not considered feasible to be implemented on FPGA as it requires a high gate count. The algorithm is meant in such a simple way that it doubles the quantity of accurate digits in each iteration and requires additional operations to see the root. It also requires rounding after each operation [8].

SRT redundant and non-redundant algorithms lead to wrong values at the last digit. The number of iterations is driven by the accuracy of initial estimation.

## 2.3 Non-restoring Algorithm

The non-restoring algorithm is an efficient way for calculating the root of a real-valued number. This doesn't require more arithmetic operations. Addition or subtraction supported the sign of the results of the previous iteration. The partial remainder and also an accurate resulting value is generated in each iteration. So less computational and execution time is required to work out the result. It also provides less complexity furthermore as exact results and doesn't require additional circuit for adjusting the result bit. No multiplier is required for the non-restoring algorithm to calculate the root. This algorithm is appropriate for the FPGA execution of the square root circuit. The proposed algorithm skips the restoring and extra addition steps of the restoring algorithm. The non-restoring algorithm is uncomplicated and can also be used for various designs of divider framework for effectual VLSI Signal processing applications which consume a low design area with lower computation error [9].

## 2.4 Cordic

CORDIC or Coordinate Rotational Digital Computer is a simple 2D geometry algorithm implementing shift and add operations. It is hardware efficient as many calculations can be done using these simple operations. The main disadvantage of this algorithm is the latency and throughput. Many new algorithms have been designed to handle these drawbacks [10]. The algorithm is mainly defined in two modes - The vectoring mode and the rotational mode.

The generalized CORDIC algorithm is given by (1), (2) and (3). The three main CORDIC algorithms are obtained by substituting the proper 'w' values like circular (w = 1), hyperbolic (w = –1) and linear (w = 0).

$$m_{i+1} = m_i - w.\sigma_i.2^{-i}.n_i \tag{1}$$

$$n_{i+1} = n_i + \sigma_i.2^{-i}.m_i \tag{2}$$

$$z_{i+1} = z_i - \sigma_i.\alpha_i \tag{3}$$

The scaling factor K is a constant value depending on the kind of algorithm it is being used like circular, hyperbolic, linear etc. CORDIC algorithm finds its applications in various functions and systems like trigonometric, vector rotations, hyperbolic, logarithmic, coprocessor units, clock recovery circuits, machine learning, image processing, robotics, multipliers, comparators etc. [11–15].

## 3 Methods

From the survey done, it was noted that CORDIC was most preferred due to its efficient means of hardware implementation and low power consumption. Therefore, this algorithm has been chosen to calculate the square root of a complex number. For our study, 2 cases are considered as shown in Fig. 1 and the algorithm is as in Fig. 2.

The novelty of the proposed work is a power based comparison of CORDIC algorithm and a combination of CORDIC with Non-restoring algorithm for the computation of complex square root. For computing complex square root, CORDIC has been

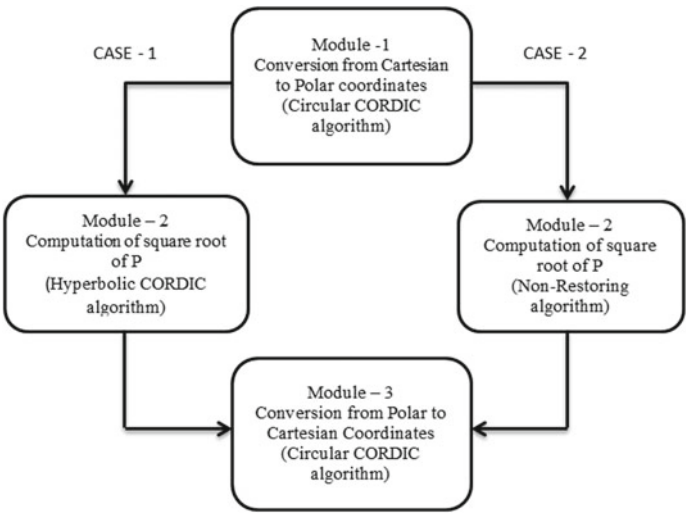



**Fig. 1** Flow diagram for the proposed work

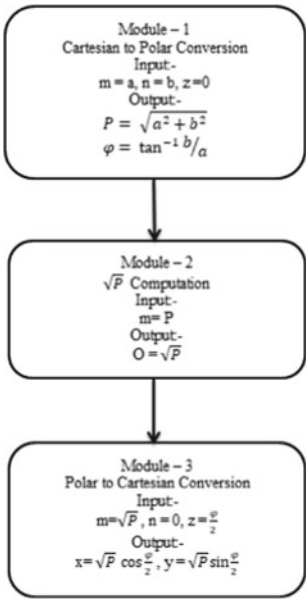



**Fig. 2** Algorithm for the proposed work

proved to be a better algorithm in terms of power consumption and in order to prove this, this study provides a comparison with Non-Restoring Square Root algorithm.

In Module 1, the input for which complex square root is to be computed is given in Cartesian form and this is converted to polar form, the arctan computation is done by CORDIC in vectoring mode. In Module 3, the data is converted from polar form back to Cartesian form, the sine and cosine values are computed by using CORDIC in rotational mode. These steps are based on a circular CORDIC algorithm and they are common for the analyzed cases. Inputs provided to the CORDIC algorithm in Rotational mode and Vectoring mode respectively, provides the results in the form of sine, cosine values and arctan values, due to the rotation of these inputs by predetermined microrotation angles until the output theta = 0 for rotational mode and y = 0 for vectoring mode. Circular CORDIC algorithm is derived from the generalized CORDIC algorithm by substituting the value w = 1. Therefore, the equations for circular CORDIC and its vectoring and rotational mode operations is as shown in Table 1.

In this paper, both the cases have the same module 1 involving circular CORDIC algorithm in vectoring mode in order to convert the input (a, b) from a + ib to $P = \sqrt{a^2 + b^2}$ and $\varphi = \tan^{-1}\left(b/a\right)$. Similarly, module 3 is the same for both the cases with the circular CORDIC algorithm implemented in rotational mode to convert the inputs $\sqrt{P}$ and $\varphi/2$ into $\sqrt{P}\cos\frac{\varphi}{2}$ and $\sqrt{P}\sin\frac{\varphi}{2}$. The module 2 i.e. the intermediate step involving computation of the square root of P, is implemented in two ways here as in case - 1 using Hyperbolic CORDIC algorithm and case-2 using Non -restoring algorithm.

**Table 1** Circular CORDIC – vectoring and rotational modes

| Equations | Mode | Vectoring mode | Rotational mode |
|---|---|---|---|
| $m_{i+1}$<br>$= m_i - \sigma_i.2^{-i}.n_i n_{i+1}$<br>$= n_i + \sigma_I.2^{-i}.m_I z_{i+1}$<br>$= z_i - \sigma_I.\alpha_i$ | $w = 1$ | $m_f = K\sqrt{m^2 + n^2}$<br>$n_f = 0$<br>$z_f = \tan^{-1} n/m$ | $m_f = K(m\cos z - n\sin z)$<br>$n_f = K(n\cos z + m\sin z)$<br>$z_f = 0$ |

**Table 2** Hyperbolic CORDIC – vector and rotational modes

| Equations | Mode | Vectoring mode | Rotational mode |
|---|---|---|---|
| $m_{i+1}$<br>$= m_i - \sigma_i.2^{-i}.n_i n_{i+1}$<br>$= n_i + \sigma_I.2^{-i}.m_I z_{i+1}$<br>$= z_i - \sigma_I.\alpha_i$ | $w = -1$ | $m_f = K\sqrt{m^2 - n^2}$<br>$n_f = 0$<br>$z_f = \tanh^{-1} n/m$ | $m_f = K(m\cosh z - n\sinh z)$<br>$n_f = K(n\cosh z + m\sinh z)$<br>$z_f = 0$ |

**Case - 1:** Hyperbolic CORDIC algorithm is derived from the generalized CORDIC algorithm by substituting the value w = –1. Therefore, the equations for Hyperbolic CORDIC and its vectoring and rotational mode operations are as shown in Table 2. To calculate the square root of P, the hyperbolic CORDIC is implemented in the vectoring mode. Before P is utilized in hyperbolic CORDIC, its value should be less than 1. For P > 1, P is right shifted by l bits as from (base paper). Now, the inputs P + 0.25 and P – 0.25 are given to hyperbolic CORDIC to obtain a scaled down version of the square root of P. This answer is left shifted by l/2 bits to obtain the actual square root of P.

**Case - 2:** Non-restoring algorithm is employed to compute the square root of P. The remainder after every iteration or cycle is the main focus of the algorithm. After every iteration, a result bit is produced. Simple addition or subtraction operations are done based on the sign of the preceding operation result. Each iteration employs the outcome of the preceding cycle. In the final iteration, if the outcome is positive then it's the final answer else the resulting outcome is processed further by addition to obtain the solution.

## 4 Results

Both the cases were implemented in Verilog and the codes were synthesized in Xilinx Vivado using ZedBoard which is a low-cost development board for the XilinxZynq-7000. The elementary board of Z-7007S of speed grade –1 was utilized. The focus

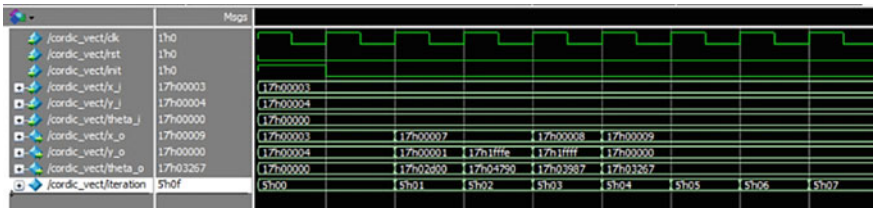



**Fig. 3** Vectoring mode – inputs [x = 3, y = 4, theta = 0]


**Fig. 4** Rotational mode – inputs [x = 3, y = 0, theta = 03267]

of this paper is an algorithm consuming less power for square root computation with other factors like delay, frequency etc. considered secondary. Hence, these secondary values were not noted.

Figure 3 and Fig. 4 provides the simulation for Vectoring mode and Rotational mode operation in CORDIC. In Vectoring mode, the given x and y input values are rotated by small pre-determined angles until the value of y co-ordinate becomes zero. Therefore, the Cartesian input values are converted to polar form. In Rotational Mode, the polar form inputs are rotated by pre-determined micro rotations until the angle value becomes zero. Therefore, the polar form inputs are converted to Cartesian form.

The Utilization reports and the primitive tables for both the cases are as in Table 3 and Table4.The power reports are as shown in Table 5.

## 5 Discussion

16-bit hexadecimal inputs were utilized for the work. Hence the CORDIC code iterates 16 times and once the right output is achieved at any iteration then the rest of the iterations will carry the same output until the last iteration. In Module 1, calculations are done using vectoring mode as in Fig. 3. and inputs given are in the Cartesian form i.e., $x = 3$; $y = 4$; $z = 0$.The output obtained is in the polar form i.e., $x_o = 9$; $y_o = 0$; $z_o = 03267$.The square root has to performed on the output obtained in polar form and therefore output of 9 is converted to 3 and sent as input to the next module (i.e. module 3). The calculated square root of $x_o$ and the angle will be the inputs for the rotational mode as in Fig. 4. i.e., $x = 3$; $y = 0$; $z = 03267$.

**Table 3** Synthesis report for case – 1 and case – 2

| On-Chip | Power (W) | | Functional units used | | Available | | Utilization (%) | |
|---|---|---|---|---|---|---|---|---|
| | Case 1 | Case 2 | Case 1 | Case 2 | Case 1 | Case 2 | Case 1 | Case 2 |
| Slice logic | 6.742 | 20.211 | 471 | 4340 | – | – | – | – |
| LUTas logic | 6.675 | 15.138 | 308 | 1774 | 63,400 | 63,400 | 0.49 | 3.33 |
| CARRY4 | – | 3.264 | – | 518 | – | 13,300 | – | 3.89 |
| Register | 0.071 | 1.803 | 75 | 1953 | 126,800 | 126,800 | 0.06 | 1.84 |
| BUFG | 0.006 | 0.006 | 1 | 1 | 32 | 32 | 3.13 | 3.13 |
| Others | 0.000 | 0.000 | 4 | 5 | – | – | – | – |
| Signals | 9.704 | 19.052 | 422 | 2509 | – | – | – | – |
| I/O | 53.358 | 58.805 | 144 | 110 | 210 | 210 | 54.29 | 55 |
| Static power | 0.841 | 1.039 | – | – | – | – | – | – |
| Total | 74.635 | 99.107 | – | – | – | – | – | – |

**Table 4** Primitive table for case – 1 and case – 2

| Ref. Name | Functional units Used | | Functional category | |
|---|---|---|---|---|
| | Case 1 | Case 2 | Case 1 | Case 2 |
| LUT6 | 191 | 32 | LUT | LUT |
| OBUF | 96 | 65 | IO | IO |
| LUT5 | 75 | – | LUT | LUT |
| FDRE | 75 | 1953 | FLOP & LATCH | FLOP & LATCH |
| LUT4 | 57 | 37 | LUT | LUT |
| LUT3 | 42 | 984 | LUT | LUT |
| LUT2 | 22 | 665 | LUT | LUT |
| IBUF | 18 | 45 | IO | IO |
| LUT1 | 5 | 177 | LUT | LUT |
| BUFG | 1 | 1 | CLOCK | CLOCK |
| CARRY4 | – | 518 | | CARRY LOGIC |

**Table 5** Comparison of power reports

| | Case – 1 | Case - 2 |
|---|---|---|
| Total on chip Power (W) | 74.635 | 99.107 |
| Dynamic Power (W) | 73.794 | 98.068 |
| Device Static (W) | 0.841 | 1.039 |
| Effective TJA ( C/W) | 4.6 | 11.5 |
| Junction Temperature (C) | 125 | 125 |

**Table 6** Comparison with existing literature

| | [4] | [5] | [6] | [10] | [11] | Proposed |
|---|---|---|---|---|---|---|
| Total on chip power | 3.95 mW | 0.7453 mW | 1.203 mW | 0.8908 mW | 81.65 μW | 74.635 W |
| Dynamic power | – | 744.635 μW | – | 889.876 μW | – | 73.794 W |

The output obtained is the initial Cartesian form input i.e., $x_o = 3$; $y_o = 4$; $z_o = 0$. The synthesis results for both cases were analyzed. Using CORDIC alone the on-chip power obtained was 74.635 W, compared to the combination of CORDIC and Non-Restoring algorithm's on-chip power of 99.107 W. Table 6 gives the comparison of the proposed work with existing literature. This difference in the power values is due to unavailability of sophisticated software for implementation and had been computed with the open source tools available.

## 6 Conclusion

In this paper, different methods to compute square root were analyzed and complex square root computation was implemented utilizing CORDIC and Non-Restoring algorithms. From the above discussion it is understood that CORDIC comparatively is a more efficient algorithm for the square root and complex square root computation. Further work can be done to reduce the power and increase the throughput utilizing pipelining and other efficient methods.

## References

1. Li Y, Chu W (1996) A new non-restoring square root algorithm and its VLSI implementations. In: Proceedings of international conference on computer designs: VLSI in computers and processors, Austin, pp 538–544, October 1996
2. Ko K-I, Yu F (2007) On the complexity of computing the logarithm and square root functions on a complex domain. J Complex 23(1):2–24
3. Wang D, Ercegovac MD (2009) A design of complex square root for FPGA implementation. Proc SPIE 7444:74440L
4. Chen H, Wu R, Lu Z, Fu Y, Li L, Yu Z (2021) A general methodology and architecture for arbitrary complex number Nth root computation. In: 2021 IEEE international symposium on circuits and systems (ISCAS), pp 1–5. https://doi.org/10.1109/ISCAS51556.2021.9401720
5. Mopuri S, Acharyya A (2017) Low-complexity methodology for complex square-root computation. IEEE Trans Very Large Scale Integr (VLSI) Syst 25(11):3255–3259
6. Chen H, Yu Z, Zhang Y, Lu Z, Fu Y, Li L (2021) Low-complexity high-precision method and architecture for computing the logarithm of complex numbers. IEEE Trans Circuits Syst I Regular Pap 68(8):3293–3304. https://doi.org/10.1109/TCSI.2021.3081517

7. Ercegovac MD, Muller J (2003) Digit-recurrence algorithms for division and square root with limited precision primitives. In: The thirty-seventh asilomar conference on signals, systems and computers, vol 2, pp 1440–1444
8. Kachhwal P Rout BC (2014) Novel square root algorithm and its FPGA implementation. In: International conference on signal propagation and computer technology, pp 158–162
9. Panda S (2016) FPGA-VHDL implementation of pipelined square root circuit for VLSI signal processing applications. Int J Comput Appl 142(5):20–24
10. Meher PK, Valls J, Juang T, Sridharan K, Maharatna K (2009) 50 years of CORDIC: algorithms, architectures, and applications. IEEE Trans Circuits Syst 56(9):1893–1907
11. Krishna A, Anusree Raj LS, Priyadarsini G, Raghul S, Ramesh SR (2019) A low power binary square rooter using reversible logic. In: International Conference on Advanced Computing & Communication Systems (ICACCS), pp 619–623. https://doi.org/10.1109/ICACCS.2019.8728490
12. Somayajulu PK, Ramesh SR (2020) Area and power efficient 64-BitBooth multiplier. In: International conference on advanced computing and communication systems, pp 721–724
13. Lachireddy D, Ramesh SR (2020) Power and delay efficient ALU using vedic multiplier. In: Sengodan T, Murugappan M, Misra S (eds) Advances in Electrical and Computer Technologies. LNEE, vol 672, pp 703–713. Springer, Singapore. https://doi.org/10.1007/978-981-15-5558-9_61
14. Aakash S, Anisha A, Das GJ, Abhiram T, Anita JP (2017) Design of a low power, high speed double tail comparator. In: Proceedings international conference on circuit, power and computing technologies
15. Abhiram T, Ashwin T, Sivaprasad B, Aakash S, Anita JP (2017) Modified carry select adder for power and area reduction. In: Proceedings international conference on circuit, power and computing technologies
16. Raveendran A, Jean S, Mervin J, Vivian D, Selvakumar D (2020) A novel parametrized fused division and square-root POSIT arithmetic architecture. In: 33rd international conference on VLSI design and 19th international conference on embedded systems (VLSID), pp 207–212. https://doi.org/10.1109/VLSID49098.2020.00053

# Email Spam Detection Using Machine Learning and Feature Optimization Method


**Naseeb Grewal, Rahul Nijhawan, and Ankush Mittal**

**Abstract** Email communication has been stated to be the most affordable, economical, and quickest communication way nowadays. Although there are significant perks of emails, sadly, its practice has been baffled by the vast amount of unsolicited and often deceitful emails. And these extreme amounts of spam are decreasing the essence of information present on the Internet and causing concern among users. Many spam detection models have been suggested and experimented with within the literature; however, the listed accuracy showed that more could be done in this area to improve accuracy. This work describes how to detect spam emails using machine learning based on words, numbers, and characters in the emails' content. We have used some prevalent machine learning models (Naive Bayes', Neural Network, K-NN, Tree, Logistic Regression) and compared them to classify emails. This paper's primary focus is obtaining optimal accuracy with the help of limited features out of 57.

**Keywords** Machine learning in email · Spam emails · Spam emails detection

## 1 Introduction

Spam email and its harms on people and businesses are reported worldwide. Spammers are trying to further their cause and steal the identities of people by using these spam emails. Recent studies have shown that people are losing their confidence in communicating through email due to spam. Statistics show that globally 14.5 billion spam emails are conveyed every day or in other words, 45% of all emails sent in a day are spam. A thorough work by the Radicati Research Group Inc., a research organization from Palo Alto, California, shows that the damage caused by spam to businesses is almost $20.5 billion every year in terms of productivity, and predictions

N. Grewal (✉) · A. Mittal
Indian Institute of Technology Roorkee, Roorkee, India
e-mail: naseebgrewal44@gmail.com

R. Nijhawan
University of Petroleum and Engineering, Dehradun, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_41

for future costs does not seem hopeful as well. Speculations show that in merely four years in future approximately 58 billion junk emails will be sent every day, and it will harm $198 billion annually [1]. Therefore, many people who were aware of this problem felt this need to develop spam detection methods.

Thus, to deal with the spam problems, one can take either of two ways, either using knowledge engineering or using machine learning. But machine learning shows great potential in dealing with email filtering because, unlike the knowledge engineering approach, it does not need consistent modification; instead, it takes a sample of pre-classified emails and training on that data to learn using a particular algorithm. Then the trained model is used to predict the unclassified emails. Machine learning has been studied extensively and can use several email filtering algorithms like the Naive Bayes method, K-nearest neighbour, Tree, Neural Networks, Logistics Regression.

Awad et al. [2] show that Naive Bayes and Artificial Neural Network works pretty solidly to classify spam emails with 99% accuracy using Naive Bayes' and concluded that in the future hybrid systems will be the most efficient way to generate anti-spam filters. Still, there were 21,700 attributes in their data set, and it can be very challenging to work with these many attributes.

Our work found out that our data has 57 features to see all models' performance. Initially, we used all the features to train the models. But it was not an efficient way to deal with so many features; then, to improve the task's productivity and efficiency, we only used the top 10 features, hence reducing the work by a significant factor.

The description of our paper is as follows: Sect. 1 is Introduction of the paper. Section 2 summarizes the efforts made in the relevant field. Section 3 gives description of the data set that is used in this particular paper. Section 4 provides the brief overview of our model. Section 5 offers a short description of all the machine learning classification algorithms that we used in this paper. Section 6 mentions the proposed approach. Results and discussion are introduced in Sect. 7. In the end, the work is concluded with Sect. 8.

## 2 Related Work

Machine learning knowledge has been applied in past years in email classification, and there are some research papers also regarding the use of machine learning in email classification. Hamid et al. [3] proposed the use of the Whale optimization algorithm (WOA) for feature selection and then applying rotation forest afterward to minimize false positives. Their work was auspicious in delivering performance (99.9% accuracy with shallow False positive rate). Hamid et al. [4, 5], reviewed ten machine learning classification methods and their application in the field of email classification. Their work involved a comprehensive understanding of email filtration and the basic structure of email spam filters and machine learning algorithms to approach the problem statement of email classification. Jawale et al. [6], show that their work on spam email classification is unique. It is based on the hybrid system which uses the accuracy of the SVM (Support Vector Machine) model and

classification speed of Naive Bayes' model and their hybrid system generated results (97.57% accuracy on the testing dataset) that were more improved than produced by either of the models alone. Manaa et al. [7] show that the accuracy of their model was not affected by the percentage of data included. Their method shows comparatively better results in accuracy matrix than k-means model. Douzi et al. [8] show that their proposed method's performance is very compelling and relatively better when compared with Bag-of-Words (BOW) and Paragraph Vector-Distributed Memory (PV-DM), two well-known deep learning methods using three different classifiers. Sahni et al. [9] show that their model produced approximately 98% accuracy using Naive Bayes Model and their insight about how this model perform much better with emails from single source account than from several source accounts. Hari K.C. et al. [10] show that their work compared output of three different classifiers and effectively choose the approach that showed maximum f1 score value for classification of emails. Rheem et al. [11] show that their hybrid ensemble approach provided the best results (94%) out of all other classifiers which they used in their work including Naive Bayes, Decision Tree and Ensemble learning model.

## 3 Dataset Description

### *3.1 Data Extraction*

We extracted the data from the URL [12] (UCI Machine Learning Repository).

### *3.2 Dataset Description*

Data set has 4601 instances (1813 spam = 39.4%) and 58 attributes (57 continuous and 1 class label). Name and description of features is given in the file named spambase.names.

Information about source, past usage, attribute information and attribute statistics is provided in spambase.documentation file and value of all 57 features is given in a separate file called spambase.data. All these files spambase.names, spam-base.documentation and spambase.data are found in the URL mentioned in Sect. 3.1.

## 4 Model Description

Figure 1 given below shows the overview of our model. First, gathering of the files that we found in repository, and then manipulating their data to create a CSV file

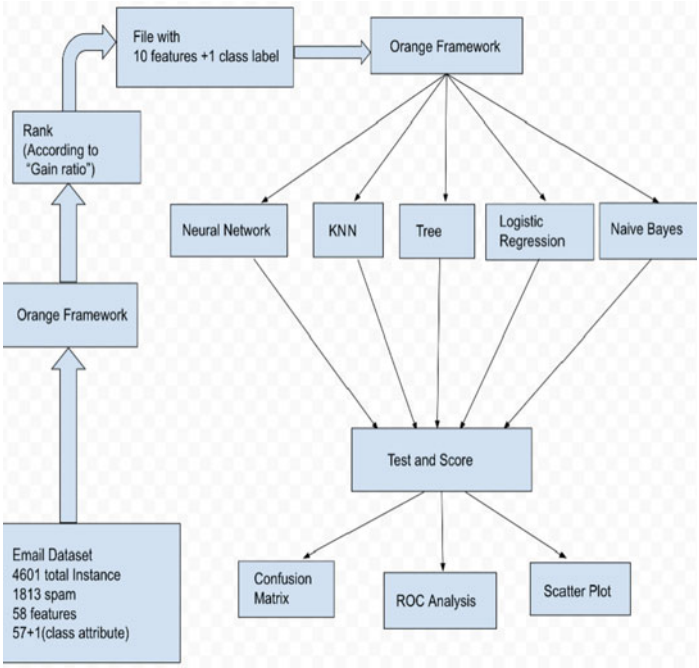



**Fig. 1** Figure shows the model of the proposed approach

which will be used to generate the model. Then, process this CSV file with all 57 features through Orange Framework for model generation. Using the "Rank" icon on all features, we arrange them in decreasing order of "Information gain ratio". Once the list of all features in descending order is obtained, then next step is to select the top 10 features, and afterwards create another CSV file of these 10 features. Then, after passing this file through the Orange Framework again, we train the different machine learning classification methods like Neural Networks, Decision Tree, Logistic Regression, K Nearest Neighbor and Naive Bayes'. After training these different models, their accuracy, recall, and precision were tested using "Test and Score" widget. And finally, visualization of the results was the last step achieved using the tools like confusion matrix, ROC analysis, and Scatter Plot.

## 5 Methods

### *5.1 Neural Networks*

Neurons are the basis of this algorithm. Layers of neuron units are present, and each unit takes an input vector and molds it into an output vector. Every unit actively

participates and applies a (mostly nonlinear) function to the input and passes the following layer's resultant output. The network is generally fed forward, but the previous layer doesn't get any feedback from the next layer. Each signal is assigned some weighting during the time it moves between two units, and by manipulating the weight values in the training phase, a neural network adapts to the given task [13–15].

Here is a short description of how a neural network algorithm works [16]:

Input $x$ goes to a neuron, then it does some mathematical manipulations on the input $x$, and gives an output. One might say three things are happening. First, multiplication of every input with some weight given in Eqs. (1), (2):

$$x_1 = x_1 \times w_1 \tag{1}$$

$$x_2 = x_2 \times w_2 \tag{2}$$

Then, the addition of weighted inputs with $a$ bias $b$:

$$(x_1 \times w_1) + (x_2 \times w_2) + b$$

And at last, the sum is processed through an activation function given in Eq. (3)

$$y = f(x_1 \times w_1 + x_2 \times w_2 + b) \tag{3}$$

## *5.2 Naive Bayes'*

This classifier assumes no relation between the class's two features. Even if there was some dependence on each other, still probability is a mutually exclusive set of all those features, and these features independently contribute to probability. Extensive data sets are preferred for the Naive Bayes' algorithm, and this model is rather easier to build. This may seem to be a simple model, but still, it is admitted to surpass highly recognized classification methods.

Here is the description of Naive Bayes' classifier [17]:

In machine learning, one need to elect the most suitable hypothesis ($h$) for given data ($d$). Considering a classification predicament, our hypothesis ($h$) might be the class label for a new data instance ($d$). One of the simplest means of picking the best presumable hypothesis based on the data we have; we can apply that data as our preceding awareness about the issue. Bayes' theorem calculates the possibility of an event transpiring based on the likelihood of another event that already befell. Below equation is mathematical representation of Bayes' theorem: [18–20] and given in Eq. (4).

$$P(h|d) = P(d|h) \times P(h)/P(d) \tag{4}$$

$\boldsymbol{P(h|d)}$ is the probability of hypothesis $h$ based on data $d$. It is known as the posterior probability.

$\boldsymbol{P(d|h)}$ is probability of the data $d$ given that hypothesis $h$ was true.

$\boldsymbol{P(h)}$ is the odds of hypothesis $h$ being true (no matter what the data). This is called the prior probability of $h$.

$\boldsymbol{P(d)}$ is the data's probability (no matter what the hypothesis).

As it is quite clear that our interest is in anticipating the posterior probability of $P(h|d)$ from the prior probability $P(h)$ with $P(d)$ and $P(d|h)$.

For many different hypotheses, estimate the posterior probability, and afterwards, select the hypothesis with maximum probability. It is known as the maximum aposteriori hypothesis (MAP). This will look like Eqs. (5), (6) and (7):

$$MAP(h) = Max(P(h|d)) \tag{5}$$

$$MAP(h) = Max((P(dh) \times P(h))/P(d)) \tag{6}$$

$$MAP(h) = Max(P(d|h) * P(h)) \tag{7}$$

For probability estimation, use the normalizing term $P(d)$.

If in each class, number of features are even, then probability of every class will be equal (e.g., value of $P(h)$). Also, this term will be a constant in final equation, hence, the resultant equation will look like this Eq. (8):

$$MAP(h) = Max(P(d|h)) \tag{8}$$

### *5.3 K Nearest Neighbours*

This is a supervised classification. Drawing a new point as it glances at the identified points most proximal to that new locality (those are its nearest neighbours), and it has those bystanders' vote, so label of majority of neighbours becomes the identification of new point ($k$ represents the number of neighbours it checks) [13].

Here is the description of the KNN classification method [21–23]:

Algorithm: Assume there are m training data samples, and $p$ unknown points.

1. Make an array $arr[]$ of data point to store the training samples.
2. For $i = 0$ to $m$:
3. Find out the Euclidean distance $d(arr[i], p)$.
4. Create a set $S$ of the $K$ smallest distances obtained. Every single one of these distances points to an already classified data point.

5. Return the majority label among $S$.

### 5.4 Decision Tree

It's an acyclic graph which can be practiced to execute decisions. For every branch point of the graph, a distinct feature j of the feature vector is researched. If the value of feature vector comes out less than a specific threshold, then follow the left branch; otherwise, go after the right branch. As one reaches the leaf node, the conclusion regarding the example is formed to get the answer to question "To which class does the example belong?" [24, 25].

### 5.5 Logistic Regression

Basic introduction to algorithms is described ahead [26]. In this algorithm, one need to plot $y_i$ as a linear variation of $x_i$; still, with a binary $y_i$, it is not elementary. The linear aggregation of traits such as wxi + b is a function that stretches from negative infinity to positive infinity, while $y_i$ got only two possible values.

$Y_i$ corresponds to the output classification of feature vector $x_i$ given by the model and $w$ is the assigned weight to the input feature vector, and $b$ is the bias value.

If we term 0 as negative label and 1 as positive label, then only thing required is to get an uncomplicated continuous function whose co-domain lies between (0, 1). For this type of case, if value given by model regarding the feature vector $x$ is very near to 0, then a negative label is given to $x$; otherwise, the label of the example is positive. A function that possesses the similar characteristic is the standard logistic function (also known as the Sigmoid function) given in Eq. (9):

$$f(x) = 1/\left(1 + e^{-x}\right) \tag{9}$$

If optimization of $x$ and $b$ is done accordingly, we will be able to deduce the output of $f(x)$ as the odds of $y_i$ being positive. For instance, if it's value is greater than or equal to the threshold 0.5, it can be said that $x$'s class is positive; else, it's negative. In practice, the threshold's value could be different based on the obstacle. So, equation of the model comes out to be something like this Eq. (10):

$$f_{w,b}(x) = 1/\left(1 + e^{-(wx+b)}\right) \tag{10}$$

Description of the above equation is given here [27].

# 6 Proposed Approach

## *6.1 Implementation*

**Step 1: Data Organizing:** Given dataset has data and attributes in different files. So, files were thoroughly examined in case data was missing (none found) and then successfully merged to get the desired CSV file that contains 58 attributes and 4601 instances.

**Step 2: Dividing Dataset:** Using the orange software widget "Data Sampler", data was divided into a ratio of 7:3 for training data and test data purposes.

Total emails = 4601 (spam = 1813 or 39.4%), where Training set emails = 3221, Testing set emails = 1380.

**Step 3: Feature Selection:** In this step, we try to shortlist the features to achieve optimum accuracy. Using the Rank widget of orange, selection of the top 10 features based on their "Information gain ratio" (Descending order) was done. Table 1 given below provides the name of those 10 features.

**Step 4: Training the Models:** First, divide the dataset into training data and test data. Then, with the help of practice data, train the all five methods.

**Step 5: Performance Evaluation:** After training the models, test data was used to see the performance, such as accuracy, precision, recall, etc., and results were obtained using widgets "Test and Score" and "Predictions" represented in Eqs. (11), (12) and (13).

$$Precision(P) = TP/(TP + FP) \tag{11}$$

$$Recall(R) = TP/(TP + FN) \tag{12}$$

$$Accuracy(A) = (TP + TN)/(TP + TN + FP + FN) \tag{13}$$

**Table 1** Table shows list of the top 10 features

| | |
|---|---|
| 1 | **word_freq_remove** |
| 2 | **word_freq_money** |
| 3 | **char_freq_$** |
| 4 | **word_freq_000** |
| 5 | **char_freq_!** |
| 6 | **word_freq_free** |
| 7 | **word_freq_hp** |
| 8 | **word_freq_hpl** |
| 9 | **word_freq_george** |
| 10 | **word_freq_credit** |

where,

$TP$ = email classified as spam and was initially spam.
$TN$ = email classified as non-spam was originally a non-spam.
$FP$ = email classified as spam, but actually, email was non-spam.
$FN$ = email classified as non-spam but was spam.

### *6.2 Visualization*

Finally, results were visualized using widgets "ROC Analysis" and "Confusion Matrix".

**Confusion Matrix.** The confusion matrix is a table that summarizes how successful the classification model is at predicting examples belonging to various classes. In confusion matrix, there are two axis: one is predicted label and the other is actual label. The label estimated by the model is predicted label.

**The Area Under the ROC Curve (AUC).** The ROC curve (ROC stands for "receiver operating characteristic," the term comes from radar engineering) is a method used to appraise the performance of classification models. ROC curves use an aggregation of the true positive rate (the proportion of positive examples predicted correctly, defined precisely as recall) and false positive rate (the ratio of negative examples predicted incorrectly) to construct a brief picture of the classification performance.

The true positive rate (TPR) and the false positive rate (FPR) are respectively defined as in Eqs. (14) and (15):

$$TPR = TP \times (TP + FN) \tag{14}$$

$$FPR = TP/(FP + TN) \tag{15}$$

## 7 Results

In this section, we present the comparison of the performance of all five methods in terms of accuracy, precision, recall. Tables 2 and 3 summarizes the performance of all five models by selecting the top 10 features.

In terms of accuracy, Decision Tree shows the maximum percentage in training data but falls to the second last position in testing data as Neural Network came on top while being in the middle in the case of the training dataset. Naive Bayes and Logistic Regression showed approximately the same accuracy in both testing and training. But in the case of KNN, there was something abnormal. It showed a very appealing accuracy in training but worked poorly in test data.

**Table 2** Table shows performance on the training dataset of the top 10 features

| Algorithm | Accuracy (%) | Precision (%) | Recall (%) |
|---|---|---|---|
| Tree | 97.9 | 94.8 | 94.8 |
| KNN | 97.6 | 92.8 | 92.8 |
| Neural Network | 97.4 | 92.2 | 92.1 |
| Naive Bayes' | 96.1 | 89.5 | 89.4 |
| Logistic Regression | 95.7 | 89.6 | 89.4 |

**Table 3** Table shows performance on the testing dataset of the top 10 features

| Algorithm | Accuracy (%) | Precision (%) | Recall (%) |
|---|---|---|---|
| Neural Network | 96.5 | 91.5 | 91.4 |
| Naive Bayes' | 96.4 | 91.4 | 91.3 |
| Logistic Regression | 95.8 | 91.2 | 91.0 |
| Tree | 94.4 | 90.5 | 90.6 |
| KNN | 90.8 | 90.3 | 90.4 |

In the case of precision, the Tree model has the maximum percentage in training but Neural Network stole the1st place in test data. KNN showed a decline in precision from training to testing, while Naive Bayes' precision value actually increased in the test dataset. Recall values of all models came out to be very close to their precision values.

## *7.1 Comparison with Other Works*

Rheem et al. [11] have done the similar work in email classification area, and they also used the same dataset that we have used in our work. Their hybrid ensemble gave the best accuracy (94%) out of all other methods which they used in their work using the mixture of feature selection and ensemble model while in our work Neural Networks gave the top accuracy of (96.5%) where we only included the top 10 feature to analyze the performance of our model. Precision and recall value of their approach is same as their accuracy but in our method, it shows a little decrease in the value of these quantities.

Girase et al. [28] show their work on two different datasets, ling corpus dataset and anron dataset. Their first and also the bigger dataset (ling corpus) contain 962 emails where each email has 3000 feature vectors. In their work, they show that for ling corpus dataset Logistic regression provided the maximum accuracy of 98.07%

**Table 4** Table shows comparison of our model with other two other paper

| Girase et al. [28] | | | Rheem et al. [11] | | Our paper | |
|---|---|---|---|---|---|---|
| Model | Ling corpus dataset accuracy (%) | Enron dataset accuracy (%) | Model | Accuracy (%) | Model | Accuracy (%) |
| SVM | 96.15 | 97.4 | Naive Bayes' | 79.28 | Neural Network | 96.5 |
| Naive Bayes' | 96.15 | 98 | Decision Tree | 92.97 | Naive Bayes' | 96.4 |
| Logistic Regression | 98.07 | 97.6 | Ensemble | 90.06 | Logistic Regression | 95.8 |
| Random Forest | 95 | 93.6 | Hybrid Ensemble | 94.41 | Tree | 94.4 |
| | | | | | KNN | 90.8 |

and in case of Enron dataset the accuracy came out to be 98% but it was showed by Naive Bayes' method. While in our paper we used over 4601 emails and merely 10 feature and still we were able to obtain 96.5% accuracy. Table 4 given below shows the accuracy comparison of our model with two recent year papers mentioned above.

In conclusion, the merit of our model is not only in terms of better accuracy but also in terms of execution time. While the other recent works use full data with a lot of features, we utilize minimum number of features. Our work is also very useful as we require less data to learn as the number of features is less (François et al. [29]) in comparison to other recent works.

## 8 Conclusion

Some very famous machine learning classification methods are studied and reviewed in this paper and their application to the spam email detection problem. Brief descriptions of algorithms are presented and their performance comparison is also presented. Including all features, Neural Network showed the highest and excellent accuracy of 100% on the training dataset and 98.3% accuracy on the testing dataset. The recall seems to show the least value for all models except Logistic Regression among all three factors recall, accuracy, and precision, respectively. But when only 10 features were taken out of 57, Neural Network still showed an amazing accuracy of 96.5% on testing data but on training data Tree model showed the best accuracy of 97.9%. In the future to improve the performance of Neural Networks and logistic regression perhaps more research needs to be done either on the number of neuron layers that can produce the optimum result in this method or combining the two methods to get better results.

## References


1. Last visited 9th May 2021. https://www.spamlaws.com/spam-stats.html
2. Awad WA, Elseuofi SM et al (2011) Machine learning methods for spam E-mail classification
3. Shafi MA, Hamid M, Shuaib JK, Alhassan I, Idris O, Adebayo (2019) Whale optimization algorithm-based email spam feature selection method using Rotation Forest Algorithm for Classification
4. Shafi MA, Hamid H, Chiroma EG, Dada JS, Bassi et al (2019) Machine learning for email spam filtering: review, approaches and open research problems
5. Nijhawan R (2017) A deep learning hybrid CNN framework approach for vegetation cover mapping using deep features. In: 13th international conference on signal-image technology & internet-based systems (SITIS)
6. Diksha S, Jawale AG, Mahajan RK, Shinkar VV et al (2018) Hybrid spam detection using Machine Learning
7. Manaa ME, Obaid AJ, Dosh MH et al (2021) Unsupervised approach for email spam filtering using data mining
8. Douzi S, AlShahwan F, Lemoudden M, Ouahidi B et al (2020) Hybrid email spam detection model using artificial intelligence. Int J Mach Learn Comput 316–322
9. Sahni R et al (2021)Analysis of naive bayes algorithm for email spam filtering. Int J Modern Trends Sci Technol 05–09
10. Hari KC et al (2021) Comparative analysis and prediction of spam emails classification using supervised machine learning techniques. Int Res J Modern Eng Technol Sci
11. Ablel Rheem DM, Ibrahim AO, Kasim S, Almazroi AA, Ismail MA et al (2020) Hybrid feature selection and ensemble learning method for spam email classification. Int J Adv Trends Comput Sci Eng
12. Last visited 9th May 2021. https://archiveicsuciedu/ml/datasets/spambase
13. Last visited 9th May 2021. https://medium.com/@Mandysidana/machine-learning-types-of-classification-9497bd4f2e14
14. Varshni D (2019) Pneumonia detection using CNN based feature extraction. In: 2019 IEEE international conference on electrical, computer and communication technologies (ICECCT)
15. Nijhawan R (2017) An integrated deep learning framework approach for nail disease identification. In: 13th international conference on signal-image technology & internet-based systems (SITIS)
16. Maind SB, Wankar P (2014) Research paper on basic of artificial neural network. Int J Recent Innov Trends Comput Commun 2(1):96–100
17. Rish I (2001) An empirical study of the Naive Bayes classifier. In: IJCAI 2001 workshop on empirical methods in artificial intelligence, vol 3, no 22
18. Burkov A et al (2019) The hundred-page machine learning book
19. Nijhawan R et al (2019) A futuristic deep learning framework approach for land useland cover classification using remote sensing imagery. In: Advanced computing and communication technologies. Springer, Singapore
20. Nijhawan R, Srivastava I, Shukla P et al (2017) Land cover classification using supervised and unsupervised learning techniques. In: 2017 international conference on computational intelligence in data science (ICCIDS)
21. Cheng D et al (2014) KNN algorithm with data-driven k value. In: International conference on advanced data mining and applications. Springer
22. Nijhawan R et al (2019) A novel deep learning framework approach for natural calamities detection. Springer, Singapore
23. Gupta S et al (2019) Classification of lesions in retinal fundus images for diabetic retinopathy using transfer learning. In: 2019 international conference on information technology (ICIT)
24. Brijain M et al (2021) A survey on decision tree algorithm for classification
25. Rawat S, Singh A, Bisht R, Nijhawan (2019) A classifier approach using deep learning for human activity recognition. In: Fifth international conference on image information processing (ICIIP)

26. Bhning D (1992) Multinomial logistic regression algorithm. Ann Inst Stat Math 44(1):197–200. https://doi.org/10.1007/bf00048682. https://dx.doi.org/10.1007/bf00048682
27. Last visited 2nd June 2021. https://www.stat.cmu.edu/~cshalizi/uADA/12/lectures/ch12.pdf
28. Girase PSTR, Patidar MK, Kushwaha MR, Yadav MM (2020) Spam detection in email through comparison of different classifiers
29. Verleysen M, François D et al (2005) The curse of dimensionality in data mining and time series prediction. In: International work-conference on artificial neural networks. Springer, Heidelberg

# Enhancement of Data Between Devices in Wi-Fi Networks Using Security Key


**C. Amarsingh Feroz, K. Lakshmi Narayanan, Aiswarya Kannan, R. Santhana Krishnan, Y. Harold Robinson, and K. Precila**

**Abstract** Due to the rapid growth of internet of things and mobile internet applications, the traditional centralized cloud computing environment has many issues such as low spectral efficiency and higher latency. In order to address these concerns, numerous edge computing technologies are created from dissimilar backgrounds for improving spectral efficiency, decreasing latency, and for supporting machine type communication. Edge computing, a kind of cloud computing, depends on the cooperation of tools instead of sending the details to remotely located servers. In the computational aspect, a Neural network analyses the gap of the user interests, which abuses the transfer based collaboration game to divide users into multiple fog socios. Edge server deployment is described as an issue of optimization of multi-objective constraint. This work proposed to reduce access delay between mobile consumer and edge server. Workload balance of edge servers is ensured. Experimental result showed the efficacy of the proposed techniques based on the base station dataset of Shanghai Telecom.

**Keywords** Cellular network · Computation capacity · Device-to-device communication · Dynamic source routing protocol · Internet of things · Mobile edge computing · Resource allocation

C. Amarsingh Feroz (✉) · K. Lakshmi Narayanan · K. Precila
Department of ECE, Francis Xavier Engineering College, Tirunelveli, India
e-mail: amarferoz@gmail.com

K. Lakshmi Narayanan
e-mail: drkln86@gmail.com

A. Kannan
SRM TRP Engineering College, Trichy, Tamilnadu, India
e-mail: Aiswarya.k@trp.srmtrichy.edu.iny

R. Santhana Krishnan
Department of ECE, SCAD College of Engineering and Technology, Tirunelveli, India
e-mail: santhanakrishnan86@gmail.com

Y. Harold Robinson
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India
e-mail: yhrobinphd@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_42

## 1 Introduction

To meet requirements of wireless data services, edge computing is the ideal choice for the Internet of Things (IoT) data services and its corresponding applications. In this project, device-to-device (D2D) communications is used as basic building block for communication and its corresponding computation. The interest of users is different and analyzed by building a neural network in this project [1]. It helps to divide users into multiple fog communities to achieve effectiveness of transfer Device to Device (D2D) typically helps to communicate directly with adjacent mobile devices. D2D interaction extends and improves system performance, enriches category functionality and variety of applications. Scarceness of frequency spectrum is due to huge number of new wireless systems resulting in greater demand of bandwidth [2–4]. Surprisingly, most measurement programs have shown that most of the time, almost all of the spectrum allocated is used inefficiently. Cognitive radio networks (CRNs) was a reliable system for bandwidth usage. In edge computing, data processing is done by the by the device and also a personal computer or workstation are used to avoid being transmitted to a data center. CRN users are divided into primary users (PUs) and secondary users (SUs). A PU in other words a licensed user, has the advantage of being given the highest priority for spectrum access and to abstain from harmful interference. An SU accesses the licensed spectrum with interweave spectrum access in the absence of PU. The effect is e-sensing techniques are used by SU's to pick the unused spectrum for communication on its own [4–8].

The SU reach the licensed frequency bands concurrently in the overlay model, in order to manipulate knowledge of PU's in codebooks and messages. The SU carefully divides the power transmitted into two sections based upon this information; one section is used to help the PU's communication while the other is used for its own communication [9]. Power allocation policies has to be optimum and cognitive cooperative radio networks (CCRNs) were arranged to increase network bandwidth, use of frequency, and coverage of SU. Based on this a performance analysis was done for ergodic capacity, outage probability, throughput, symbol error probability, stable transmission condition, as well as delay of packet transmission is also taken into account. Otherwise, this secondary user simply stays silent. Thus in the proposed protocol, the advantages of a two-path relay channel can be well exploited to serve the primary users as well as facilitate the bi-directional transmissions between the two secondary users. Based on the classification with regard to space, the SU can transmit in the gray and white spaces, and it is not allowed to operate in the section of black space when PU is active. Depending on the principles opening of spectrum, nearby definition of cognitive radio (CR) has an ability to understand [10, 11]. It helps to know from the environment and to statistical variations based on internal states.

## 2 Related Works

Previously, edge computing is an extension of cloud computing, has proposed trying to offload the Internet of Things (IoT) data services and applications to meet the rapidly rising wireless core network requirements. We establish the notion of using communication device to interface (D2D). Cloud computing is a system in which the data is transferred directly from the cloud server.D2D enlarges and increases system capacity as well as it enriches service category. For the purpose of computation, we use a neural network which would transfer data by dividing users into multiple fog nodes. Pustokhina et al. [12] implemented a new effective model; deep neural networks in edge computing enabled IoMT systems. Deng et al. [13] has discussed the research road map and the core concepts that deliver essential background for future research in IoT using edge computing. In addition to this, Alfakih et al. [14] developed a new reinforcement learning based state action reward state action approach for resolving the resource management issue [15] in edge computing and makes the off-loading decision effective by reducing computing time delay, system cost and energy consumption.

## 3 Proposed Methodology

### *3.1 Vehicular Communications*

**Spectra.** Vehicular to vehicular communication has competitors from cable television and other competing technologies which vie for sharing of the frequency spectrum for high-speed internet access. Several sections of the radio spectrum are regulated for the productive use of the finite radio spectrum by regulatory bodies. The growing use of engine-specific wireless communications systems should require V2V communication system spectrum availability. As a consequence, FCC has allotted 75 MHz of spectrum for V2V and V2I radio spectrum at 5.9 GHz in the US.

**VANET Challenges.** Challenges faced while implementing security features in VANET are huge. Issues are channel itself is dynamic with often disconnections, vehicular arrival and departure is instantaneous, for emergency and safety messages wireless channels are used, making it vulnerable to various threats and attacks. Here we characterize VANET's unique characteristics and recognize some serious problems.

**Security, Privacy, and Safety.** Privacy and security concerns in VANETs are of fundamental importance. For Vehicle safety protocols, trust, efficiency and high mobility.

**Cognition Cycle.** The two major features of CR were first identified here: cognitive ability and re-configurability. First we explain briefly the interpretation of the CR

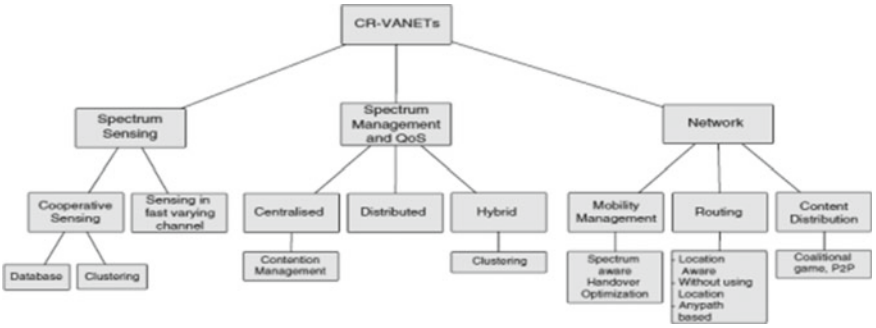



**Fig. 1** Cognition cycle process in DSRC Channel

perception process as well as some CR-VANET particulars. The function of CR enabled device is to adapt to its environment. CR components are predominantly radio, processor, database of intelligence, engine and learning, tools and optimization, and engine to reasoning. CR has strengths both for the mental and configurability. Cognitive awareness allows CR to feel and gather information from its background, and secondary users, for starters, may figure out the best spectrum available. The Cognition cycle process in DSRC Channel is mentioned in Fig. 1.

Spectrum hole can be defined as portion of spectrum, a primary/licensed user does not use at the same place and time instant. Spectrum sensing and signal detection are the techniques to identify a spectrum hole.

## *3.2 Network Layer: Mobility Management*

Based on the descending order of received signal strength, The MS tries to find an apt cell by scanning through the list of channels. It selects the BCCH channel that satisfies the requirements.

**Criteria for Cell Selection.** Requirements a cell should satisfy are:

- Choose a cell of the selected PLMN. It has to be checked by the mobile station.
- The selected cell should not be blocked. The PLMN operator can bar mobile stations to access certain cells, E.g. cells used for traffic handover. BCCH broadcasts this information.
- The radio path loss set by the PLMN operator should be above a threshold.
- MS enters a "limited service" and emergency calls are made if no appropriate cell is found.

### *3.3 Network Configuration*

The different areas of network are PLMN area, MSC/VLR Service area and Location area;

**PLMN Area.** PLMN provider offers services like land mobile communication to the public. Inside a PLMN area, the function of mobile user is to set up calls to another user of the same/another network.

**Location Area.** If PLMN knows the likely position of MSs which are active in their own coverage area, unnecessary network wide paging broadcasts can be avoided. In order to find the likely positions of MS divide the network into Location Areas (LAs).

- Control of BTS in a particular area can be accomplished by one or more BS
- Corresponding BSCs are controlled by corresponding MSC in a particular area.

### *3.4 Security*

Issues are many. Some of them are bandwidth, small device size, dynamic topology and limited battery life. Attacks occur in two ways. Passive attack: changes in transmitted data do not happen. Nevertheless, allows unauthorized user to know the message. Active attack: Attacker either prevents or modifies the message. Identification of malicious node: pattern of dropped packets, more energy consumption, bandwidth usage more, unreliable packets, more delay, often frequent connection break and routing loops.

**Power Control.** Factors associated with power control are node capacity limitation, dynamic infrastructure, energy constraint and utilization of channel Power maintenance is due to:

- Low capacity condition: nodes are active during transmission of packet and inactive when they wait for packet reception.
- Transmission power control: Factors associated with it are error rate, transmission range and interference. More transmission power means transmission range is increased and hop count is reduced.
- Power aware routing: Purpose of routing protocols are to allow more network life time. This is done by reducing battery life time of the nodes.

**Numerical Results.** A random network with single source S, one destination D1, NP = 16 primary nodes, both primary and secondary users have equal transmitting power. i.e. EP = ES, which yields P = S t = − 5 db. source S and destination D1 are positioned in the center of two opposite sides Placement of random nodes are uniform in a square area. The value of policies to be optimized are set at = 0.99, which is the norm for static networks. For primary transmissions power allocated

is computed by obtaining the largest that satisfy Pout, SS (dS) = S for S = 0.15 and a distance d = 0.2. Plotting of performance of routing protocols are primary end-to-end throughput (5) vs energy consumption of the primary usually expressed in dB, 10 log10 E (k, RP, Q). The value of RP = 3 bits/s/Hz is set, RS = 1 bits/s/Hz and NS = 16. The resulting output is [0, 1] for (Optimal) and K in {0,..., NS} The energy spent by the two schemes are equal. But for K _ 1, K-OSLA has a bit higher energy consumption compared to K-Closer. The reason for this primary nodes are given preference in the metric taken. For Primary packets, more transmission power is used resulting in more related measures.

## 4 Results and Discussions

Rectangular node was generated using 50 sensor nodes and one connector node in this project, as stated in Fig. 2. By using ANN algorithm, shortest path between target and destination node is determined. Square shaped form of Communication is exhibited to avoid Traffic.

The Fig. 3 shows the forming of anchor modules between the node with origin and the node on anchor. The overlaid wireless networks for both structures are with PPP modeling. The base station (BS) is affiliated with each mobile user (MU), thereby representing the Voronoi cell in the cellular network The circularly defined area around each BS, with radius c0, represents the nano particle-inner area. The outside of the circular area has been the cell-edge area in each cell. Potential secondary users in each cell in compensation for a fraction of the disjoint spectrum band can actively


**Fig. 2** Communications between various nodes

help downlink communications in the cell-edge. Growing SU has departed d meters away from either a fixed sender, and the ellipse pairs both together. The performance analysis of the proposed and the existing model [14] is analyzed in terms of packet delivery ratio, throughput and packet drop, which is graphically stated in Fig. 4, 5 and 6, and the tabulation is given in Table 1 in terms of throughput.

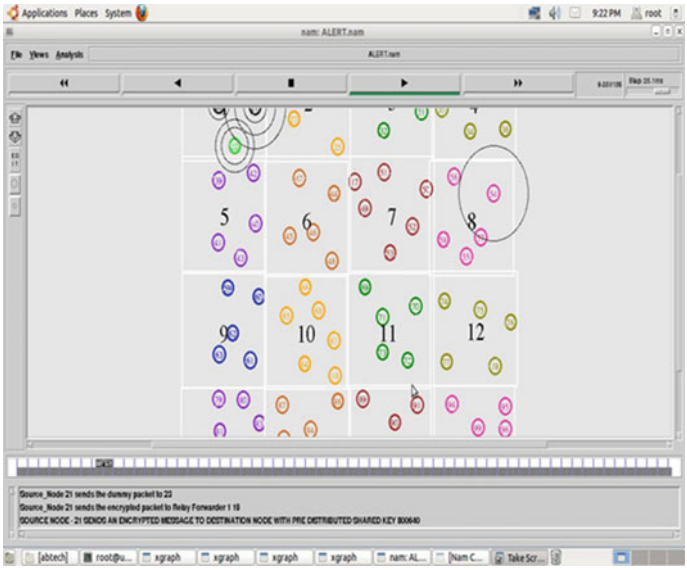



**Fig. 3** Anchor node formations

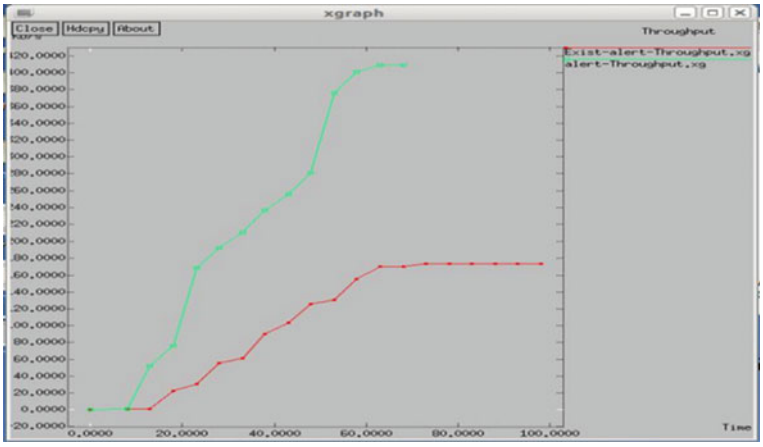



**Fig. 4** Throughput comparison

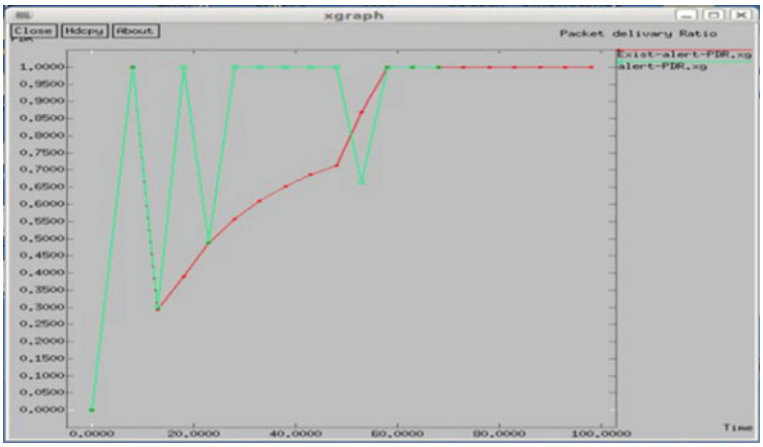



**Fig. 5** Packet delivery ratio comparison

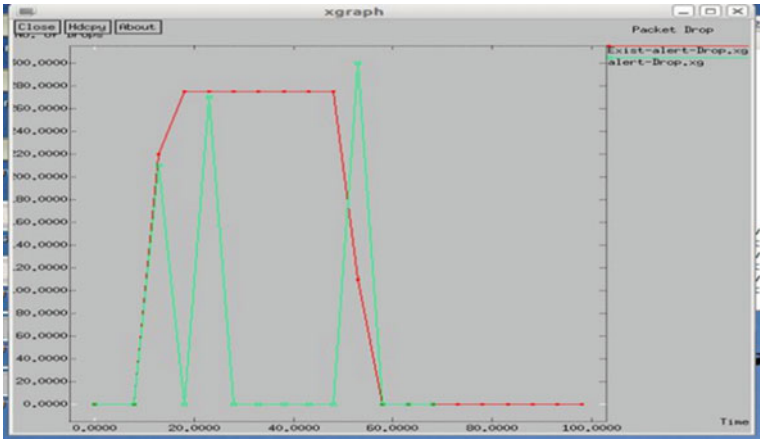



**Fig. 6** Packet drop comparison

**Table 1** Comparative analysis in terms of throughput

| Throughput | | |
|---|---|---|
| Number of nodes | Existing model [13] | Proposed |
| 0 | 0 | **0** |
| 20 | 80 | **30** |
| 40 | 240 | **80** |
| 60 | 380 | **140** |
| 80 | 400 | **140** |
| 100 | 400 | **140** |

## 5 Conclusion

To increase the downlink rate of SN, we formulated the optimal control optimization concern after this. The SINR restriction of Pus is put forth, so the conventional interference power restriction to shield PUs from harmful interference is done. In order to solve the problem, an iterative approach based on convex approximation was done. If there is more number of antennas the implications are analyzed for the success of PN and SN. Results inferred show that the SN can tolerate higher sum amounts with SINR limitations more than power generation constraints without intrusion. Also in this work, we focus only on the issue of the SU's power allocation.

## References

1. You KH, Chae H (2016) Energy-efficient mobile multi- threading powered by microwave transmitters transfer. IEEE J Sel Areas Commun 34(5):1757–1771 (2016). IEEE Trans Commun 64(10):4268–4282
2. Mao Y, Zhang J, Song SH, Letaief KB (2016) Power-delay tradeoff in multi-user mobile-edge computer networks. In: Proceedings the IEEE world society conference, Washington, DC, USA
3. Liu J, Mao Y, Zhang J, Letaief KB (2016) Delay-optimal scheduling of graphics rendering for mobile computing systems. In: Proceedings of the International IEEE 2016. Proc. IEEE Int.
4. Chen X, Pu L, Gao L, Wu W, Wu D (2017) Exploiting massive D2D collaboration for energy-efficient mobile edge computing. IEEE Wirel Commun 24(4):64–71
5. Hao S, Yang O, Muta H, Gacanin H, Furukawa H (2016) Sensing-based spectrum sharing of oil and natural gas for collective cognitive radio networks. Ins Electron Inf Commun 99(8):1763–1771
6. Wang AQ, Mancuso V (2014) Cellular network device to device survey. IEEE Commun. Tuts. Surveys 16(4):1801–1819
7. You Y, Huang K, Chae H (2017) Efficient mobile edge offloading resource allocation. IEEE Trans Wirel Commun, 1397–1411 (2017)
8. Dai L, Gao X, Su X, Han S, Lin CI, Wang Z (2015) Low-complexity soft-output signal detection based on Gauss-Seidel method for uplink multi-user large-scale MIMO systems. IEEE Trans Veh Technol 64(10):4839–4845
9. Feroz CA, Narayanan KL, Jayaraman G, Devaraj AS, Precila K (2020) Transmission of data between devices via mobile edge processing in wi-fi networks. Int J Sci Technol Res 9(3)
10. Narayanan KL, Krishnan RS, Julie EG, et al (2021) Machine learning based detection and a novel EC-BRTT algorithm based prevention of DoS attacks in wireless sensor networks. Wirel Pers Commun (2021). https://doi.org/10.1007/s11277-021-08277-7
11. Robinson YH, Vimal S, Julie EG, et al (2021) 3-Dimensional manifold and machine learning based localization algorithm for wireless sensor networks. Wirel Pers Commun. https://doi.org/10.1007/s11277-021-08291-9
12. Pustokhina IV, Pustokhin DA, Gupta D, Khanna A, Shankar K, Nguyen GN (2020) An effective training scheme for deep neural network in edge computing enabled Internet of medical things (IoMT) systems. IEEE Access 8:107112–107123
13. Deng S, Zhao H, Fang W, Yin J, Dustdar S, Zomaya AY (2020) Edge intelligence: The confluence of edge computing and artificial intelligence. IEEE Internet Things J 7(8):7457–7469

14. Alfakih T, Hassan MM, Gumaei A, Savaglio C, Fortino G (2020) Task offloading and resource allocation for mobile edge computing by deep reinforcement learning based on SARSA. IEEE Access 8:54074–54084
15. Ning H, Li Y, Shi F, Yang LT (2020) Heterogeneous edge computing open platforms and tools for internet of things. Futur Gener Comput Syst 106:67–76

# Extraction of Dataset for Indian Sign Language Recognition from News Video


**Pooja Goswami and S. Padmavathi**

**Abstract** Sign language recognition can provide excellent platform for the hearing-impaired people in society. In day-to-day life physically challenged people face problem to communicate with normal people. There are many research works done for sign language recognition, but there is, lack of availability of dataset for the research work which leads to less development of tools for training and testing for the recognition of signs. This paper proposes an algorithm that takes offline NEWS video stream to create dataset of sign language video with the corresponding text. The signs are recognized from the gesture of the person in the news-video with the help of lip movement, facial expression. Therefore, those signs will be mapped to the audio text in the video up to sentence level. The proposed algorithm takes input of offline video and apply various techniques such as extraction, segmentation, template matching, speech conversion to generate dataset which consist of sentence level segmented videos which will be pushed to pre-trained model to get the output for the sign language which will be compared with the audio text and comparison score will be calculated. Therefore, this technique will help to enhance the dataset for the future research for Sign Language recognition.

**Keywords** Dataset generation · Face detection · Facial expression · Lip-Movement · Lipnet · Segmentation · Sign Language Recognition

## 1 Introduction

Sign language is a part of hearing-impaired people's daily life. Their communication with people is restricted as there exist very minimum percent of people who can

P. Goswami · S. Padmavathi (✉)
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: s_padhmavathi@cb.amrita.edu

P. Goswami
e-mail: cb.en.p2cse19016@cb.students.amrita.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_43

understand and use it. According to survey, five percent of people out of world population are suffering from the disability of hearing. They require translation to communicate with other people who don't know sign language. This necessitates automatic sign language recognition tool.

The major challenge for sign language recognition is the dataset involved for the work. Two common methods related to data set are 1) Create dataset for different signs manually 2) Use the existing limited dataset for the work. Sign language recognition algorithms usually addresses a limited set of signs. The development of tools for sign language recognition requires huge dataset which have the variations with respect to skin tone, gesture etc. that can be used for training different model. Automatic generation of dataset from the sign language videos could accelerate the research on sign language recognition. One can generate the dataset required for their work and need not invest huge amount of time for creating dataset. This paper aims in creating Indian Sign Language dataset from offline news video and automatically matching word corresponding to sign depicted in news video using Lipnet.

In Sign language hand gesture, lip movement, facial expression plays important role in conveying information. In many algorithms, Indian Sign language recognition is addressed as a gesture recognition task. The news video should be segmented based on the news, sentence and words of the sign language. The sign language gestures should be extracted and mapped to the corresponding words in the spoken. In this proposed method, the complete video called the parent news video is segmented into different news clip videos by extracting the frame containing the same News headline displayed at the bottom. In the next stage, the segregated news video is segmented into sentence level video by using shot boundary detection algorithm. From the complete frame the ISL action performer is extracted and stored as separate videos using Grabcut algorithm. The audio stream for each video segment is extracted and converted to text and stored in a text file. The significant keywords are filtered using NLP techniques like stop word removal and POS tags. Since pretrained models for ISL does not exist and the lip movement closely imitates the words spoken, Lipnet is chosen for ISL word recognition. Lipnet is a pretrained model available for word identification based on lip movement. The face region of the ISL performer is extracted using Haar Cascade classifier and this video is given as input to Lipnet. The words predicted from Lipnet is compared with extracted significant words. Though the score is less for the words recognized, it could be improved by combining with hand gestures.

## 2 Related Work

Indian sign language recognition comes under one of the most demanding research works for society, gathering dataset for the recognition of sign is most challenging part of the work as a result a big portion of time is allocated for gathering or creating dataset.

Video pre-processing that is conversion of video to images and cropping the region of interest and matching template is explained in [1]. For pulling out text from the images by using optical character recognition extraction, where Tejas et al. [2] have fetched character from number plate and pushed it to neural network, they achieved 81% of accuracy for the first trial. Shrenika et al. [3] explained the working of matching template in sign language recognition, by pushing input as image containing sign and by using OpenCV lib of matching-template the signs are matched and corresponding text related to the sign is given as result, the limitation is found in dataset which is on character level.

For abrupt change in video background the most commonly and widely used method in computer vision is shot detection algorithm. Mas et al. [4] has explained shot boundary detection algorithm based on histogram color difference which is able to get the abrupt shot boundaries and smooth shot boundaries for temporal color variation. Another technique for detection of shots in video is G-SURGE method which takes spatial distribution of colors and apply the algorithm to detect abrupt and gradual shots in video which is explained in [5]. The aim is to provide simple and fast algorithm that work on real time environment.

Initial step of dataset preparation involves detection of region of interest. Sharifara et al. [6] has explained the detection of face based on Haar Cascade classifier and segmentation of facial region. To detect the facial part from the face region Viola Jones Algorithm can be applied which is explained in [7] this algorithm is tested on different poses and achieved 92% accuracy score. Feature extraction on images is done by pulling the region of interest from the image, and this is done widely in image processing by segmentation, there are many methods used for segmentation. Dixit et al. [8] has explained global thresholding algorithm, another segmentation technique that is The Chan and Vese model is discussed in [9]. Park et al. [10] has explained the grabcut algorithm to pull human region from the image by giving reference of skeleton image to detect human body in image, where it has showed better performance than naïve grabcut algorithm.

Filtration of spoken text is done by Natural language processing [NLP] techniques which is discussed in [12]. Model used in neural network comes as pre-trained or one can design it from scratch. Model such as CNN, CNN + LSTM, Lipnet etc. are used for sign language recognition and some are pre-trained on database which can be used to pull out result with low cost of execution. For tracking lip movement in region of interest Lipnet model is used in [13]. This model is used to get output from variable length of sentences. This is trained on GRID Corpus. Lip reading based on different neural network model is described in [14]. Word recognition from CHMM model on Grid corpus is shown in [15] Also, for identifying the signs from the video of ISL performer manual dataset is being created in [16] which is pushed to CNN model pre-trained on ImageNet followed by LSTM network therefore, the demerit of this method is on lack of dataset. For Identifying sign language with the help of body gesture, hand is tracked using K-means clustering and can be used as additional feature for the dataset as explained in [17]. They achieved good accuracy for signs but the challenge was on dataset which had been created manually for training and

testing. For detecting gesture with the help of fingertip is explained in [18] which can be used in future to enhance the dataset with help of gesture.

In all the paper which results in recognition of sign, the main challenge is visible on the dataset for training and testing, hence there it requires an algorithm which can be used to create dataset by video which consumes less time for aggregating dataset.

## 3 Methodology

Dataset for ISL recognition is created from offline news video. Input video is subjected to various image pre-processing steps and the region of interest is extracted from the frames of video. The news video frames generally contain news reader or the new clips. The sign language performer is included as a part of every frame in a small window embedded in the main news video. The news text scrolls at the bottom with appropriate subheading. The locations of the sign language performer and text are usually fixed at certain position in the entire video. For extracting the sign language, the video has to be broken down to word level video. The single video stream has to be segmented at various level such as news segments, and news sentence before reaching the final form. Simple video processing algorithms may not give the accurate segmentation of news segments due to presence of embedded news clips. Hence the major steps involved can be subdivided into following tasks.

- Extracting different news segment from one news video (parent)
- Extracting sentence level video from child news video
- Audio Analysis

The steps involved in each task is briefly depicted in the architecture diagram shown in Fig. 1. First part deals with extracting news segment by pulling out the headline text from the video and mapping it with the video frames. Second part of the work explains the extraction of Indian Sign Language (ISL) performer region in news segment for the dataset which is pushed to pre-trained Lipnet model. Third part of the work involves the processing of audio to text and applying NLP techniques for extracting the ISL words.

### *3.1 Architecture*

#### 3.1.1 Extraction of News

Simple video processing algorithms may not give the accurate segmentation of news segments due to the presence of embedded news clips. This paper proposes shot boundary detection algorithm to identify the news clips. It also uses the headlines of the scrolling news text to identify the news segment. From the video frames the headline text is extracted based on the fixed location and stored as a video.

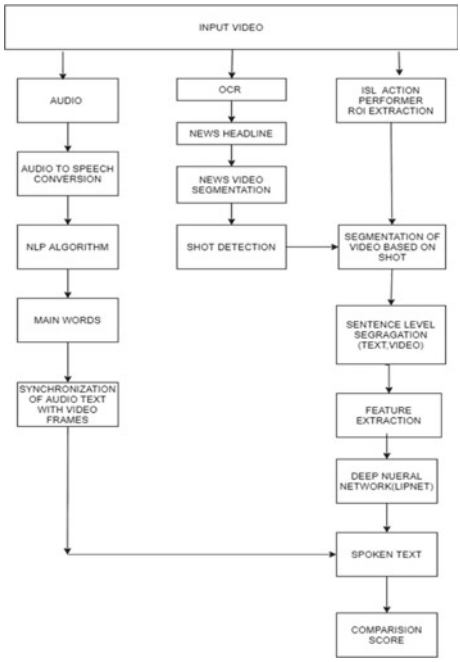



**Fig. 1** Complete flow chart for creation of dataset

OCR algorithms are applied on this video and the corresponding text are stored in a document with the line number corresponding to frame number. The redundant headlines are removed by string matching algorithms and the corresponding line numbers are used for extracting news segment. The unique headlines are stored in a document for further processing.

The result is enhanced by applying shot boundary detection algorithm on the news segment video. This will segment the news clip and also refine the extraction of news segment when the headlines are missing or changing. Shot detection algorithm is used to find shots within the news since the change of scene within the news results in change of sentence. Hence the detection of shots is taken as segmented video from news.

Shot boundary detection algorithm works on the base knowledge of discontinuity or variability of the intensity level of the frames. By taking frame n as a reference, the difference in intensity value is calculated with n + k frame where k > 1. By computing F (n, n + k) that is absolute difference of frames as shown in Eq. (1), which is then compared with the threshold value if the F (n, n + k) is above threshold it takes as shot.

$$\mathrm{F}\ (\mathrm{n},\ \mathrm{n}+\mathrm{k}) = \sum \left| I_{n(a,b)} - I_{n+k(a,b)} \right| \tag{1}$$

### 3.2 Extraction of ISL Performer NEWS

From the complete frame of news video, the ISL performer region is cropped based on the static coordinates and the video is stored separately. Based on the new segments extracted using shot boundary detection as discussed in previous section, the ISL performer video is also segmented accordingly. This implies that sentence level segregation of news video is applied for both complete frame video and ISL performer video. Then from the ISL performer video, the face of the ISL performer is detected by using Haar-cascade classifier for frontal face. Once the facial region of ISL performer is detected, grabcut algorithm is applied for background removal and fetching only the facial part in the video. The video with only the facial region is given as input to Lipnet model and store the output in csv file for later comparison.

There exists different model for lip reading but those are trained on phonemes, character but Lipnet is a pre-trained model on GRID Corpus which has various sentences from different speaker, once the input is given as video it gives sentence as output. Therefore, the Lipnet model is taken for experimenting the sample videos for which it gives desired sentence level prediction.

### 3.3 Pre-processing on Audio

The audio associated with the video is given to google speech translator. The speech converted text is stored in a text document. After getting the audio text of news video, NLP techniques is to filter the audio text, sign language involve no stop word for conveying message, Therefore stop word removal, POS tagger removes the stop words from the audio text and store the main words in text document.

### 3.4 Comparison of Spoken Text to Predicted Output of Lipnet

The spoken audio text is compared with the predicted text from Lipnet for depicting the words from segmented video which can be used for breaking sentence level video to word level in future. Comparison score can be calculated with respect to filtered audio text and predicted output text from Lipnet which is expressed in Eq. 2.

$$Comparision\ Score = \frac{Total\ Predicted\ Words}{Total\ Number\ of\ Words\ in\ Audio\ text} \tag{2}$$

# 4 Dataset

## *4.1 Dataset Creation*

The generated dataset has videos in it such as extracted news segment video, and from news segment video sentence level video is fetched by the use of shot boundary detection algorithm also the ISL performer sentence level video is fetched in same way which can be used to extract feature such as face by using segmentations technique, which is pushed to Lipnet and compared with the spoken text.

After applying the matching template algorithm all the frames with same headline is detected as shown in Fig. 2(a). All the frames collected from matching template algorithm are combined together to form a video as depicted in Fig. 2(b).

The news segment is pushed to shot boundary detection algorithm where the shot is detected, according to the shots taken the frame of ISL performer region as shown in Fig. 3(a) which is fetched out with the static coordinates is combined together as shown in Fig. 3(b).

Figure 4 depicts the result of face detection by using Haar cascade classifier and background subtraction using grabcut algorithm.

The audio from the corresponding video is converted into text as shown in Fig. 6(a). Only main words from audio text document is stored in text document which is shown


(a) (b)

**Fig. 2** **a** Frame of matched news. **b** Segmented news video


(a) (b)

**Fig. 3** **a** ISL performer frame. **b** Shot video of ISL action performer

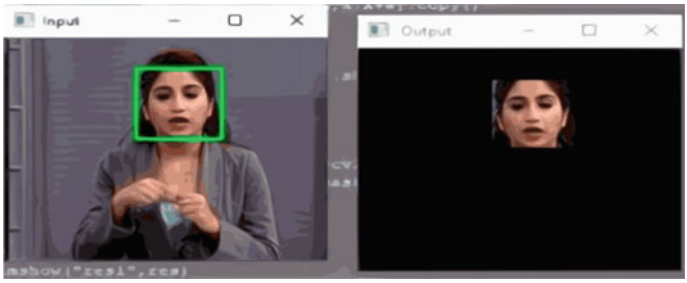



**Fig. 4** Face detection and background elimination

in Fig. 6(b). Scrolling text from the corresponding video is extracted and stored in document of one-to-one relation of line with frame of video and stored in text document to keep track of frame number. The result after removing the redundancy in text document by using string filtration method. Since the processed video, corresponding headline, spoken text are already prepared, the processed video is being pushed Lipnet to get predicted words as shown in Fig. 5. Which will be later used for segmenting the video into word level. With the sentence level segmented video 18.79% of words are matched with audio text.

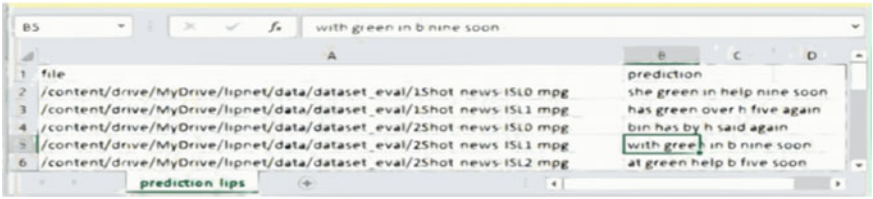



**Fig. 5** Word Prediction from Lipnet

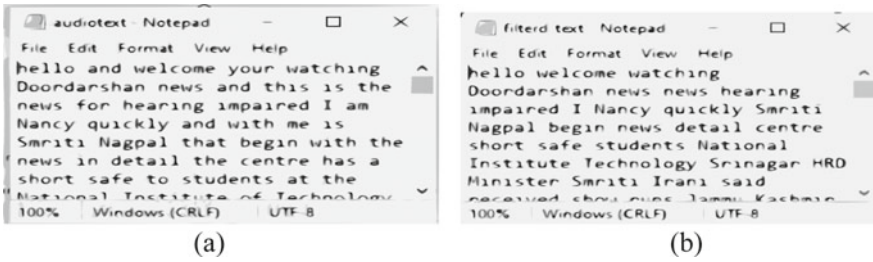



**Fig. 6** **a** Audio text of corresponding video. **b** Filtered audio text

## 5 Comparison and Discussion

Table 2 presents the state-of -the Art comparison, where the accuracy rate for different model is shown.

The complete work is done on segmenting the video to sentence level, which can be experimented with Lipnet model. There exist many pre-trained models but as per the requirement the need of the model in this project is to experiment the segmented video on sentence level by reading lip of the ISL performer. Therefore, the hand region is not segregated, only the facial region is segmented and saved for the sample generated videos. Since the work is based on mouth region, From all listed model in Table 2. Lipnet trained on Grid Corpus is the most suitable model for the experimentation with the video. Based on the outcome the videos it can be broken down in future to word level, so that the word level dataset can get trained with their respective labels Table 1. Shows the comparison of the spoken text from video to the predicted text from Lipnet, where in some segmented video words are getting matched and vice-versa. After experimenting the complete set of videos with Lipnet the total matched words are found to be 87 with the true audio text that has count of words 463.Therefore the comparison score for the words comes out to be 18.79%.

The comparison is not done with other models, since this work is related with generation of dataset from offline news video. Once the videos are broken down to word level the generated data can be trained and compared with different neural model for better accuracy.

**Table 1** Comparison of true text and Predicted text

| Video name | Spoken text (True value) | Predicted text |
|---|---|---|
| 1Shot-news-ISL0.mpg | Smriti Irani has said she has received the show runs from Jammu and Kashmir Chief Minister | She green in help nine soon |
| 1Shot-news-ISL1.mpg | Home minister Rajnath Singh have a discussion over the matter | Has green over h five again |
| 2Shot-news-ISL0.mpg | The matter with the J&K Chief Minister Mehbooba Mufti the chief minister | Bin has by h said again |

**Table 2** State-of-the-Art-Comparison

| Dataset | State-of-the-Art | Model | Rate [%]/Comparison score |
|---|---|---|---|
| GRID Corpus | Assael et al. [13]<br>Gergen et al. [15] | Lipnet | 95.2%<br>94.23% |
| Proposed Work | Assael et al. [13] | Lipnet | 18.79% (Comparison score) |

## 6 Conclusion and Future Work

With the advancement of the work in deep learning and computer vision our work gives the effectiveness of having automated system for generation of dataset for sign language recognition. As it makes easier for the researches to have their own dataset without creating it manually. The main objective of project is to create algorithm which segments the news video and generate dataset for sign language recognition so that one can have variation in their dataset and can have large dataset for their research work. This paper proposes algorithm for extracting sentence level video from complete news video. It also segments the sign language performer region from the video and maps the text converted from audio to the Sign language performer video. It involves technique such as extraction of text, segmentation, Longest string matching, Shot boundary detection, Haar cascade classifier, speech to text converter and NLP preprocessing on text. To get more understanding of sign language in future the processed video can be segmented for hand gesture. The future work involves exploration on segmenting the sentence level video into word level to get more specific dataset for the sign language recognition.

## References

1. Badhe PC, Kulkarni V (2015) Indian sign language translator using gesture recognition algorithm. In: 2015 IEEE international conference on computer graphics, vision and information security (CGVIS), Bhubaneshwar, Odisha, India, pp 195–200. https://doi.org/10.1109/CGVIS.2015.7449921
2. Tejas B, Omkar D, Rutuja D, Prajakta K, Bhakti P (2017) Number plate recognition and document verification using feature extraction OCR algorithm. In: 2017 international conference on intelligent computing and control systems (ICICCS), pp 1317–1320. https://doi.org/10.1109/ICCONS.2017.8250683
3. Shrenika S, Madhu Bala M (2020) Sign language recognition using template matching technique. In: 2020 international conference on computer science, engineering and applications (ICCSEA), pp 1–5. https://doi.org/10.1109/ICCSEA49143.2020.9132899
4. Mas J, Fernandez G (2013) Video shot boundary detection based on color histogram. In: IEEE second international conference on image information (ICIIP), pp 476–480
5. Deepak CR, Babu RU, Kumar KB Krishnan CBR Shot boundary detection using color correlogram and Gauge-SURF descriptors. In: 2013 fourth international conference on computing, communications and networking technologies (ICCCNT), pp 1–5. https://doi.org/10.1109/ICCCNT.2013.6726510
6. Sharifara A, Mohd Rahim MS, Anisi Y (2014) A general review of human face detection including a study of neural networks and Haar feature-based cascade classifier in face detection. In: 2014 international symposium on biometrics and security technologies (ISBAST), pp 73–78. https://doi.org/10.1109/ISBAST.2014.7013097
7. Vikram K, Padmavathi S (2017) Facial parts detection using Viola Jones algorithm. In: 2017 4th international conference on advanced computing and communication systems (ICACCS), pp 1–4. https://doi.org/10.1109/ICACCS.2017.8014636
8. Dixit K, Jalal AS (2013) Automatic Indian sign language recognition system. In: 2013 3rd IEEE international advance computing conference (IACC), pp 883–887. https://doi.org/10.1109/IAdCC.2013.6514343

9. Shangeetha RK, Karthik KS (2013) Indian sign language gesture segmentation using active contour. IJECCE 4(3):761–766
10. Park S, Yoo J (2014) Human segmentation based on GrabCut in real-time video sequences. In: 2014 IEEE international conference on consumer electronics (ICCE), pp 111–112. https://doi.org/10.1109/ICCE.2014.6775931
11. Nirkin Y, Masi I, Tran Tuan A, Hassner T, Medioni G (2018) On face segmentation, face swapping, and face perception. In: 2018 13th IEEE international conference on automatic face and gesture recognition (FG 2018), pp 98–105. https://doi.org/10.1109/FG.2018.00024
12. Hasan MR, Maliha M, Arifuzzaman M (2019) Sentiment analysis with NLP on Twitter data. In: 2019 international conference on computer, communication, chemical, materials and electronic engineering (IC4ME2), pp 1–4. https://doi.org/10.1109/IC4ME247184.2019.9036670
13. Assael YM et al (2016) Lipnet: End-to-end sentence-level lipreading. arXiv preprint arXiv:1611.01599 (2016)
14. Sindhura P, Preethi SJ, Niranjana KB (2018) Convolutional neural networks for predicting words: a lip-reading system. In: 2018 international conference on electrical, electronics, communication, computer, and optimization techniques (ICEECCOT), pp 929–933. https://doi.org/10.1109/ICEECCOT43722.2018.9001505.
15. Gergen S, Zeiler S, Abdelaziz AH, Nickel R, Kolossa D (2016) Dynamic stream weighting for turbo- decoding-based audiovisual ASR. In: Interspeech, pp 2135–2139
16. Jayadeep G, Vishnupriya NV, Venugopal V, Vishnu S, Geetha M (2020) Mudra: convolutional neural network based Indian sign language translator for banks. In: 2020 4th international conference on intelligent computing and control systems (ICICCS), Madurai, India, pp 1228–1232. https://doi.org/10.1109/ICICCS48265.2020.9121144
17. Keshari T, Palaniswamy S (2019) Emotion recognition using feature-level fusion of facial expressions and body gestures. In: 2019 international conference on communication and electronics systems (ICCES), pp 1184–1189. https://doi.org/10.1109/ICCES45898.2019.9002175
18. Shangeetha RK, Valliammai V, Padmavathi S (2012) Computer vision based approach for Indian Sign Language character recognition. In: 2012 International conference on machine vision and image processing (MVIP), pp 181–184. https://doi.org/10.1109/MVIP.2012.6428790

# Feature Fusion of LBP, HELBP & RD-LBP for Face Recognition


**Shekhar Karanwal and Manoj Diwakar**

**Abstract** As literature suggests that single descriptor fails to produce desired results therefore the proposed work launches novel Fused Local Descriptor (FLD), by merging 3 state of art local descriptors. The services of FLD was taken from LBP, HELBP & RD-LBP. As all these 3 descriptors adopts different concepts therefore their integration is done. In LBP, the gray value (neighbors of patch ($3 \times 3$ size)) are involved in comparison to center gray value. In HELBP, the gray value (horizontal neighbors of patch ($3 \times 5$ size)) are involved in comparison to center gray value and in RD-LBP, the different scale pixels are compared. PCA is bring next for reduction and classification is further done by SVMs. FLD attains staggering rates than either of LBP, HELBP & RD-LBP. FLD also proves its significance in front of several other literatures based ones. ORL & GT are used for evaluation in MATLAB R2018a environment.

**Keywords** Horizontal Elliptical LBP (HELBP) · Local Binary Pattern (LBP) · PCA · Radial Difference LBP (RD-LBP) · SVMs

## 1 Introduction

In literature plentiful of work has been done in Face Recognition (FR) & Facial Expression Recognition (FER), in tough conditions. These tough conditions are occurred due to intrapersonal changes. In these conditions, the performance accomplished by local descriptors are astonishing. Local descriptors work on distinct patches of the image for feature extraction. Some authors apply directly to image & some employ after pre-processing techniques. The aggregate size is built by joining features region wise (extracted from original input image or the transformed image evolved after deploying local descriptor concept). The literature work suggests clearly that fusion of local descriptors outperform the alone descriptor results & many others. The prime objective of all these developed descriptor is to achieve discriminativity in

S. Karanwal (✉) · M. Diwakar
CSE Department, Graphic Era University (Deemed), Dehradun, India
e-mail: shekhar.karanwal@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_44

uncontrolled conditions. By observing that fusion of 2 or more descriptors improve the accuracy therefore the invented work launch the novel descriptor by fusing the 3 state of art descriptors. The main objective of the fused descriptor is to produce the discriminativity in uncontrolled conditions. Results shows the effectiveness of the proposed fused descriptor. The related work and the description of the other utilized descriptors (for making the feature size of robust fused descriptor) are mentioned in the upcoming sections.

## 2 Related Works

Santosh et al. [1] develops the novel feature for FER by deploying the fusion of HOG& LBP. The HOG feature extraction is carried out from distinct landmarks of face such as mouth, eyes, nose & face (by usage of gradients & orientations of pixel value), which are concatenated for full feature size making. Similarly, extraction of LBP features is carried out. The fused method defeats alone result of HOG, LBP & others. Chengeta et al. [2] gives the FER survey by utilizing LBP & Local Directional Pattern (LDP). LBP is applied to attain local features & LDP is employed for directional features capturing. Both LBP & LDP feature extraction is done from small regions, which are merged into single feature representation. The merged technique provides healthier outcomes than many LBP variants. Sarangi et al. [3] presented his approach for Ear recognition by integrating features of Pyramid HOG (PHOG) & LDP. Spatial shape features are acquired by usage of HOG & local texture features are captured by utilizing LDP. PCA is adopted further for compact size. Dabagh et al. [4] merged features of Gabor & LBP to make discriminative face analysis. In Gabor, face shape encoding is done for broad range of scales and in LBP, local details are captured. Further CCA is used for selecting discriminative features. Annalakshmi et al. [5] invented hybrid technique for classification of gender, by usage of Spatially Enhanced LBP (SLBP) & HOG. Due to micro-patterns texture representation & local shape (by catching of gradient structures) this combination outclasses many descriptors. Reddy et al. [6] proposed novel FR by using LBP, LVP& Gabor LVP (GLVP). The features obtain from all are integrated to manufacture the discriminant size for matching. The merged one gives better consequences than alone ones. Hazgui et al. [7] discovered the new method for texture by merging LBP & HOG. First detection (of patch) is carried out to eliminate the undesired details. Then LBP & HOG extracted features are joined. LBP captures the local details and HOG captures the gradient details. The fusion scheme appears to be effective in distinct challenges. Yang et al. [8] presented the effective fusion scheme for FER by merging LBP, LDN & EOH features. The texture features are attained by LBP; the intensity details are given by LDN & Edge Orientation Histograms (EOH) are used as feature extractor. The rate gain by fusion scheme is higher than alone. Siddharth et al. [9] discovered heterogeneous FR by joining MB-LBP & MLBP. For catching of structural features local descriptors are consumed. The fused technique reaches stupendous results. Wang et al. [10] makes use of Local Difference Pattern (LDP) &

HOG in FR, by joining them. The local patterns are extracted by use of LDP & edge features are extracted by use of HOG. The rate is enhanced after feature fusion. Sharifnejad et al. [11] provide the fusion method for FER by merging ELBP & PHOG. ELBP size is formed from distinct face areas &PHOG size is formed from distinct cells (locally) of several pyramid levels. ELBP + PHOG attains the superb results. Zhang et al. [12] presented its FER based on GWT, LBP & LPQ. GWT is employed initially for gabor feature extraction and then LBP & LPQ are used for encoding gabor image. The fusion leads big size so 2 stage PCA-LDA is used for compaction. Many methods are outclassed by launched method.

This work launches novel descriptor so-called Fused Local Descriptor (FLD), by merging 3 state of art local descriptors. The services of FLD was taken from LBP [13], HELBP [14] & RD-LBP [15]. As all these 3 descriptors adopts different concepts therefore their integration is done. In LBP, the gray value (neighbors of patch ($3 \times 3$ size)) are involved in comparison to center gray value. In HELBP, the gray value (horizontal neighbors of patch ($3 \times 5$ size)) are involved in comparison to center gray value & in RD-LBP, different scale pixels are compared. PCA [16] is bring next for reduction with classification by SVMs [17]. FLD attains staggering rates than LBP, HELBP & RD-LBP. FLD also proves its significance in front of various literature based ones. ORL [18] & GT [19] are the 2 datasets utilized in MATLAB R2018a environment. The work remaining is as follows: Related works are discussed in Sect. 2, Methods are explored in Sect. 3, Experiments are conducted in Sect. 4 with Discussions in Sect. 5 & Conclusion in Sect. 6.

## 3 Methods

### *3.1 Local Binary Pattern (LBP)*

In LBP [13], gray value (neighbors of patch size ($3 \times 3$)) are involved in comparison to center gray value. The 1 value is set to those areas where gray neighbors have large or same gray value to gray value (of center), else 0 is issued. The length of pattern produced is 8 bits, after comparison. From pattern length, the LBP code is deduced by the weights grant. This procedure continues for all places, which yields the LBP image. The region wise ($3 \times 3$) extracting features from LBP image pulls of LBP size of 2304, as each regional size is 256. Equation 1 states LBP concept for single position. M, N, $V_{N,m}$ & $V_{2,2}$ are neighbors length, radius, distinct places of each & center place.

$$\mathrm{LBP}_{M,N}(x_c) = \sum_{m=0}^{M-1} k\left(V_{N,m} - V_{2,2}\right)2^m, k(y) = \begin{pmatrix} 1 & y \geq 0 \\ 0 & y < 0 \end{pmatrix} \tag{1}$$

### 3.2 Horizontal Elliptical LBP (HELBP)

In HELBP [14], gray value (of horizontal neighbors of patch size (3 × 5)) are involved in comparison to center gray value. The 1 is set to those areas where gray neighbors have large or same gray value to gray value (of center), else 0 is issued. The length of pattern produced is 8 bits, after comparison. From pattern length, HELBP code is deduced by weights grant. This process continuation in all places form HELBP image. The region wise (3 × 3) extracting features make HELBP size of 2304, as region size is 256. Equation 2 states HELBP concept for one place. M, $N_1$, $N_2$, $V_{N_1,N_2,m}$ & $V_c$ are neighbors length, radius ($N_1$), radius ($N_2$), distinct places of each & center place.

$$\mathrm{HELBP}_{M,N_1,N_2}(x_c) = \sum_{m=0}^{M-1} k\left(V_{N_1,N_2,m} - V_c\right)2^m, k(y) = \begin{pmatrix} 1 & y \geq 0 \\ 0 & y < 0 \end{pmatrix} \quad (2)$$

### 3.3 Radial Difference LBP (RD-LBP)

In RD-LBP [15], there is computation of radial differences among different scales of mxn image patch. In this work RD-LBP is examined by using 5 × 5 patch. Specifically, $N_1$ scale pixels are differentiated (subtracted) from $N_2$ scale pixels. The 1 value is set to those areas where gray differences have large or same gray value to 0, else 0 is issued. The length of pattern produced is 8 bits, after comparison. From pattern length, RD-LBP code is deduced by the weights grant. This procedure continues for all places, which yields RD-LBP image. The region wise (3 × 3) extracting features from RD-LBP image pulls of full RD-LBP size of 2304, as each regional size is 256. Equation 3 states RD-LBP concept for single position. M, $N_1$, $N_2$, $V_{N_1,m}$ & $V_{N_2,m}$ are neighbors length, radius ($N_1$), radius ($N_2$), pixel places at $N_1$& pixel places at $N_2$.

$$\mathrm{RD-LBP}_{M,N_2,N_1} = \sum_{m=0}^{M-1} k(V_{N_2,m} - V_{N_1,m})2^m, k(y) = \begin{pmatrix} 1 & y \geq 0 \\ 0 & y < 0 \end{pmatrix} \quad (3)$$

### 3.4 Fused Local Descriptor (FLD)

The LBP, HELBP & RD-LBP all builds the feature size of 2304 so FLD size is [2304 2304 2304] = **6912**. Figure 1 display the FLD illustration for one position. The global approach PCA is bring into use for size compression with SVMs for matching. Figure 2 gives the full architecture of the proposed method.

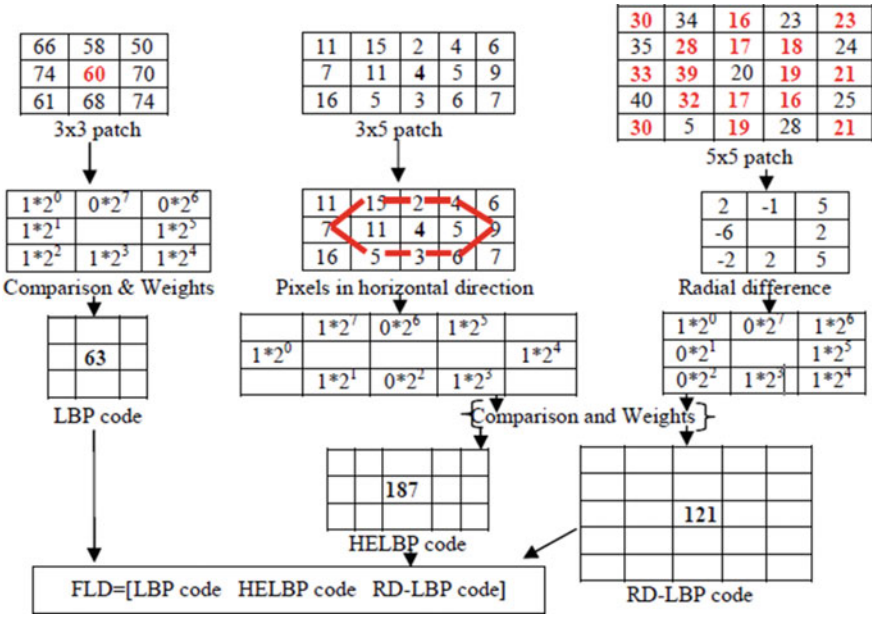



**Fig. 1** FLD demonstration for single position

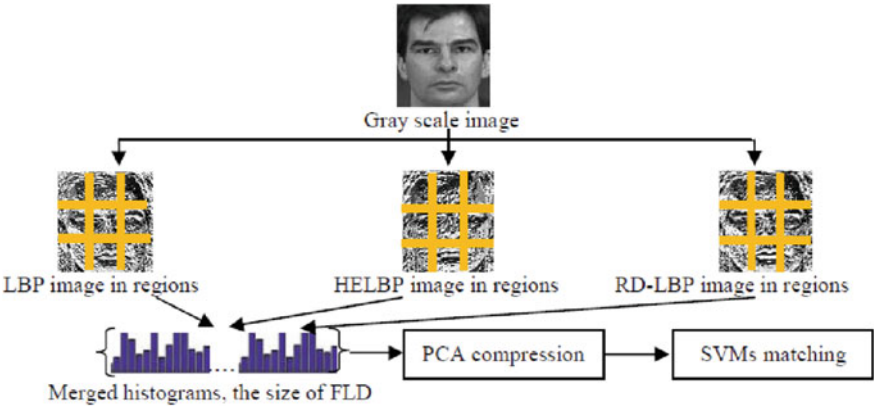



**Fig. 2** Proposed method full architecture

# 4 Experiments

## *4.1 Datasets Utilized*

The 2 used datasets are ORL & GT. Table 1 conveys all the details of the 2 datasets. Figure 3 displays some ORL & GT images.

**Table 1** ORL & GT datasets illustration

| Datasets | Subjects | Subject images | Size | Total | Conditions |
|---|---|---|---|---|---|
| ORL | 40 | 10 | 112 × 92 | 400 | Pose, illumination & emotion |
| GT | 50 | 15 | Varying | 750 | Illumination, emotion, pose & scale |


**Fig. 3** Some ORL & GT images

**Table 2** Particulars concerning the feature size

| Descriptors | 1 Regional size | Complete size | PCA reduction | |
|---|---|---|---|---|
| | | | **ORL** | **GT** |
| | | | PCA | PCA |
| LBP | 256 | 2304 | 33 | 41 |
| HELBP | 256 | 2304 | 33 | 41 |
| RD-LBP | 256 | 2304 | 33 | 41 |
| **FLD** | **256*3** | **6912** | 33 | 41 |

## *4.2 Descriptors Particulars Concerning the Feature Size*

Due to huge original image size it becomes obligatory to lower down the image size, for cost minimization. On ORL, the image size is set to 53 × 48 straight way. On GT, same is carried out after color conversion to gray. Rest details are found in Table 2.

### *4.3 Rate Assessment*

The recognition rate (in %) is assessed by the following formula utilization: $\left[\mathrm{RR} = \frac{T_{st} - I_{mf}}{T_{st}} * 100\right]$. $T_{rst}$ reveals training size, $T_{st}$ reveals test size, $I_{mf}$ are the incorrect matches found & RR reveals the recognition rate. On ORL, $T_{rst}$ consider values from 1:4 & $T_{st}$ have the remaining values which are 9:6. The upmost rate is attained after 20 executions. The FLD pulls off better rates than other ones. Table 3 provides all the results assessment. On GT, $T_{rst}$ consider values from 6:9 & $T_{st}$ have the remaining values which are 9:6. The upmost rate is attained after 25 executions. The FLD pulls off better rates than the other ones. Table 4 provides all the results assessment. The obtained rates clearly illustrate the FLD superiority. From the results gained it can be easily concluded that by joining the multiple features the rate increases.

**Table 3** Results assessment on ORL

| | RBF performance | | | | POLY performance | | | |
|---|---|---|---|---|---|---|---|---|
| | $T_{rst}$ essentials | | | | $T_{rst}$ essentials | | | |
| | $\mathbf{T_{rst} = 1}$ | $\mathbf{T_{rst} = 2}$ | $\mathbf{T_{rst} = 3}$ | $\mathbf{T_{rst} = 4}$ | $\mathbf{T_{rst} = 1}$ | $\mathbf{T_{rst} = 2}$ | $\mathbf{T_{rst} = 3}$ | $\mathbf{T_{rst} = 4}$ |
| All descriptors | RR | | | | RR | | | |
| LBP | 76.66 | 89.68 | 93.92 | 97.50 | 73.88 | 89.06 | 93.21 | 97.50 |
| HELBP | 78.05 | 90.31 | 94.28 | 97.91 | 74.16 | 89.06 | 93.92 | 97.08 |
| RD-LBP | 80.27 | 91.87 | 95.71 | 98.33 | 76.94 | 90.00 | 94.64 | 97.50 |
| **FLD** | **84.16** | **94.06** | **97.85** | **99.16** | **80.83** | **92.18** | **96.07** | **98.75** |

**Table 4** Results assessment on GT

| | RBF performance | | | | POLY performance | | | |
|---|---|---|---|---|---|---|---|---|
| | $T_{rst}$ essentials | | | | $T_{rst}$ essentials | | | |
| | $\mathbf{T_{rst} = 6}$ | $\mathbf{T_{rst} = 7}$ | $\mathbf{T_{rst} = 8}$ | $\mathbf{T_{rst} = 9}$ | $\mathbf{T_{rst} = 6}$ | $\mathbf{T_{rst} = 7}$ | $\mathbf{T_{rst} = 8}$ | $\mathbf{T_{rst} = 9}$ |
| All descriptors | RR | | | | RR | | | |
| LBP | 84.66 | 85.50 | 87.42 | 88.33 | 82.22 | 84.25 | 86.57 | 87.66 |
| HELBP | 84.22 | 86.50 | 87.42 | 88.66 | 81.55 | 84.75 | 86.28 | 87.66 |
| RD-LBP | 84.88 | 86.75 | 88.85 | 90.00 | 82.66 | 86.25 | 87.14 | 88.33 |
| **FLD** | **88.00** | **88.25** | **90.00** | **91.33** | **85.77** | **87.25** | **88.85** | **90.00** |

**Table 5** FLD comparison with other literature ones on ORL & GT datasets

| On ORL | $T_{rst}$ essentials | | | | On GT | $T_{rst}$ essentials | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | | **6** | **7** | **8** | **9** |
| All methods | RR in % | | | | All methods | RR in % | | | |
| NCDB-LBPac [20] | 80.2 | 93.7 | 97.1 | 99.1 | NCDB-LBPac [20] | 83.3 | 85.7 | 88.0 | 89.3 |
| NCDB-LBPc [20] | 81.1 | 92.5 | 96.7 | 98.7 | NCDB-LBPc [20] | 83.3 | 86.5 | 87.1 | 90.0 |
| OD-LBP [21] | 79.4 | 92.5 | 97.1 | N/A | OD-LBP [21] | 87.3 | N/A | N/A | N/A |
| LC-LBP [13] | 51.6 | 68.7 | 80.0 | 82.9 | LC-LBP [13] | N/A | 68.2 | 71.7 | 73.3 |
| VELBP [13] | 61.1 | 78.4 | 86.4 | 92.9 | VELBP [13] | N/A | 74.5 | 76.8 | 77.6 |
| RLRRP [22] | N/A | N/A | 92.5 | 96.0 | CZZBP [24] | 81.3 | 82.7 | 85.1 | 87.3 |
| MSRL [22] | N/A | N/A | 93.7 | 95.9 | CMBZZBP [24] | 84.4 | 87.7 | 88.5 | 89.3 |
| Extend Face_log [23] | N/A | N/A | 88.2 | N/A | EIT2KBSRM [25] | 71.1 | 72.5 | 77.4 | N/A |
| A3M_MSDs [23] | N/A | N/A | 86.5 | N/A | MWCSRFR [26] | 80.2 | 83.5 | N/A | N/A |
| **FLD** | **84.1** | **94.0** | **97.8** | **99.1** | **FLD** | **88.0** | **88.2** | **90.0** | **91.3** |

### *4.4 Comparing the FLD Results with Literature Ones*

The proposed FLD (obtained from RBF) is involved in comparing its results with numerous methods. FLD totally outstrip the all methods. Table 5 presents the whole comparison on ORL & GT.

## 5 Discussions

By merging 3 descriptors size, the discriminative descriptor i.e. FLD is gained. With solitary ones, the FLD beats the rate of 18 methods (from literature). Among 18, 9 belongs to ORL and 9 belongs to GT. FLD reach the maximum rate of 99.16% on ORL and 91.33% on GT, which is quite staggering. All simulation is conducted in the MATLAB R2018a environment. The impact of the proposed method can be deduced from the results achieved.

## 6 Conclusion

This work proves that by integrating the features of 3 local descriptors (i.e. LBP, HELBP & RD-LBP) there is immense increase in the recognition rate. The integrated descriptor is termed as FLD. As integration (region wise) yields the vast feature size so PCA service is taken for the reduction. Then SVMs is availed for classification. The proposed FLD defeats fully the results of individual & many others (from literature). By observing the results attained by FLD, the future work comes up with the novel

local descriptor in uncontrolled conditions. In addition, the FLD ability will also test on diverse applications.

## References


1. Santosh M, Sharma A (2021) Fusion of multi representation and multi descriptors for facial expression recognition. In: IOPCS: MSE, pp 1–12
2. Chengeta K, Viriri S (2018) Facial expression recognition: a survey on local binary and local directional patterns. In: Proceedings of ICCCI, pp 513–522
3. Sarangi PP, Mishra BSP, Dehuri S (2019) Fusion of PHOG and LDP local descriptors for kernel-based ear biometric recognition. Multimedia Tools Appl 78:9595–9623
4. Dabagh MZNA, Ahmad MI, Isa MNM, Anwar SA (2020) Face recognition system based on fusion features of local methods using CCA. In: Proccedings of IEEC
5. Annalakshmi M, Roomi SMM, Naveedh AS (2019) A hybrid technique for gender classification with SLBP & HOG features. Cluster Comput 22:11–20
6. Reddy N, Rao M, Satyanarayana C (2019) A novel face recognition system by the combination of multiple feature descriptors. Int Arab J Inf Tech 16(4):669–676
7. Hazgui M, Ghazouani H, Barhoumi W (2022) Genetic programming-based fusion of HOG and LBP features for fully automated texture classification. Visual Comput 38(2):457–476
8. Yang L, Ban X, Li Y, Yang G (2018) Multiple features fusion for facial expression recognition based on ELM. Int J Emb Syst 10(3):181–187
9. Siddharth KP, Kisku DR (2017) Heterogeneous face identification by fusion of local descriptors. In: 7th IACC, pp 886–894
10. Wang H, Zhang DS, Miao ZH (2018) Fusion of LDB and HOG for face recognition. In: 37th CCC
11. Sharifnejad M, Shahbahram A, Akoushideh A, Hassanpour RZ (2021) Facial expression recognition using a combination of enhanced local binary pattern and pyramid histogram of oriented gradients features extraction. IET Image Process 15(2):468–478
12. Zhang B, Liu G, Xie G (2016) Facial expression recognition using LBP and LPQ based on Gabor wavelet transform. In: Proceedings of ICCC, pp 365–369
13. Karanwal S (2021) A comparative study of 14 state of art descriptors for FR. Multimedia Tool Appl 80:12195–12234
14. Hatibaruah R, Nath VK, Saikia KJ, Hazarika D (2019) Elliptical local binary co-occurrence pattern for face image retrieval. J Stat Manag Syst 22(2):223–236
15. Wang M, Zhang Y (2020) Real time texture extraction based on the improved median robust extended LBP. In: ICCPR, pp 280–287
16. Jia Y, Liu H, Hou J, Kwong S, Zhang Q (2021) Semisupervised affinity matrix learning via dual-channel information recovery. IEEE Trans Cybner, 1–12
17. Ma Z, Li B (2020) A DDoS attack detection method based on SVM & K-NN in SDN environment. Int J Comput Sci Eng 23(3):224–234
18. Karanwal S (2021) Graph based structure binary pattern for face analysis. Optik 241:166965
19. http://www.anefian.com/research/face_reco.htm
20. Karanwal S, Diwakar M (2021) Neighborhood and center difference based LBP for face recognition. Pattern Anal Appl 24:741–761
21. Karanwal S, Diwakar M (2021) OD-LBP: orthogonal difference local binary pattern for face recognition. Dig Signal Process 110:102948
22. Lu J, Lin J, Lai Z, Wang H, Zhou J (2021) Target redirected regression with dynamic neighborhood structure. Inf Sci 544:564–584
23. Gao J, Li L, Guo B (2020) A new extend face representation method for face recognition. Neural Process Lett 51:473–486

24. Karanwal S, Diwakar M (2021) Two novel color local descriptors for face recognition. Optik 226:166007
25. Yadav S, Vishwakarma VP (2019) Extended interval type-II and kernel based sparse representation method for face recognition. Expert Syst Appl 116:265–274
26. Alobaidi T, Mikhael WB (2019) Facial recognition system employing transform implementations of SR method. In: Proceedings of IMSCC, pp 880–883

# Global Best Guided Binary Crow Search Algorithm for Feature Selection


**Unnati Agarwal and Tirath Prasad Sahu**

**Abstract** Feature selection is a universal combinatorial optimization problem that is used to enhance the characteristics of high-dimensional datasets by eliminating redundant data and selecting prominent features to generate acceptable classification performance. In optimization problems, the objective function can have several local optima, but the ultimate aim is to identify global optima or values close to global optimum. The Crow Search Algorithm (CSA) is a recently suggested metaheuristic algorithm, implemented to feature selection issues systematically. It is witnessed that the CSA solution search equation is suitable in exploration but poor in exploitation. In this paper, a global best-guided solution to CSA (G-CSA) is proposed and applied to pick the optimum feature subset in a wrapper mode to boost exploitation for classification purposes. The efficiency of the proposed methodology is examined on twelve standard UCI datasets. When comparing experimental results, it is clear that the proposed algorithm is superior to its challengers.

**Keywords** Classification · Crow Search Algorithm · Feature selection · Optimization

## 1 Introduction

Feature selection (FS) issues are explored by researchers within the areas of statistics, data mining, and pattern classification for several years. With no past statistics, it is hard to decide which feature is important and which is not. Thus, a variety of features such as important, obsolete or redundant features are stored in a data set. Not only are these obsolete or redundant features ineffective, but they can also degrade classification results. FS aims to maintain classification learning output while removing certain features [1].

U. Agarwal (✉) · T. P. Sahu
Department of Information Technology, National Institute of Technology Raipur, Raipur, India
e-mail: unnatiagarwal31194@gmail.com

T. P. Sahu
e-mail: tpsahu.it@nitrr.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_45

Many approaches for dealing with FS issues have been suggested so far. They can be divided into three categories: filters, wrappers, and embedded methods. Filters pick feature subsets that are independent of the predictor as part of the pre-processing level. The sequential forward selection method (SFS), the sequential backward selection method (SBS), and others are examples of filter methods [2]. Embedded methods pick features during learning and are normally customized to particular learning algorithms. Wrappers use learning algorithms as a black box for their statistical precise ranking of feature subset [3].

## 2 Related Work

Recently, metaheuristic technology for optimization in FS problems has been used to help wrappers locate optimal function subsets, so that they can obtain results for specific global search approaches. Some of these metaheuristic feature selection methods include Particle Swarm Optimization (PSO) [2], Whale Optimization Algorithm (WOA) [3] and Jaya Algorithm [4, 5].

A nature-inspired algorithm known as the Crow Search Algorithm (CSA) was recently proposed by Askarzadeh in 2016. Since CSA is a facile and effective algorithm; it is used in a variety of engineering fields. Many optimization problems have yielded promising results for CSA, including fractional order PID controller [6], energy management in smart grid [7], crop identification in agriculture [8], and many others.

Apart from the applications stated above, CSA has been efficaciously used in the arena of FS. The properties of Grey Wolf Optimization (GWO) and CSA were combined in [9] to solve the FS problem as well as additional unrestrained function optimization issues.

Exploration and exploitation are the two key phases of nature-inspired algorithms that seek to increase the algorithm's convergence speed and/or escape local optima while looking for a goal. During the exploitation process, search agents appear to move locally in the search room. During the exploration process, on the other hand, they are encouraged to travel in unpredictable ways. Since the awareness probability (AP) of the crow is the key parameter responsible for exploration and exploitation in CSA, it suffers from the issue of getting stuck in local minima. The primary objective is to counterbalance exploitation and exploration. This is accomplished by adding new methods for modifying the algorithm or combining two algorithms. In 2019, the authors hybridized the CSA with chaos theory to address the FS problem [10]. Similarly in 2020, the above-mentioned algorithm was further integrated with fuzzy c-means algorithm for FS problems of medical diagnosis [11]. In 2020, to update the position of crows, [12] proposes a definition of time-varying flight length. The locations of the crows were revised based on AP using a PSO-based group-oriented CSA [13]. To stabilize the trade-off between exploration and exploitation, a global best-led technique was introduced, which took into account the global best location of the crows [14].

However, the CSA's full potential is yet to be realized, and it can be used to build a modern and enhanced FS strategy. As a result, the main intention of this research is to develop a new CSA-based FS technique. To avoid being restrained in local minima, when updating the position of a specific crow, the global best location of the overall population (crows) is taken into consideration.

The rest of the paper is organized as follows: Sect. 3 describes the CSA. Section 4 discusses the proposed approach in depth. The results and conclusions are reported in Sects. 5 and 6, respectively.

## 3 Crow Search Algorithm

CSA is an evolutionary search algorithm inspired by nature, focused on the actions of a flock of crows. Crows are clever birds that exist in groups and search for food in their region, memorizing the best food source they find. Crows have a propensity to investigate new food sources by trailing another crow to learn where it eats and stealing from there. When a crow detects that another crow is pursuing it, it moves to an alternate location to deceive the chasing crow. The pseudocode of CSA is given in Algorithm 1.

Assume a d-dimensional world in which N crows (flock size) move through space in search of the best food source at any known time $t$ (generation). The location of crow $i$ in this generation $t$ is defined as given in Eq. (1):

$$X^{i,t} = \left[X_1^{i,t}, X_2^{i,t}, x_3^{i,t} \ldots\ldots., X_d^{i,t}\right] \tag{1}$$

where $i$ goes from 1 to $N$ and $t$ from 1 to $tmax$ (the maximum number of iterations). Since crows memorize the location of their food source so the best location of crow $i$ is defined as given in Eq. (2):

$$M^{i,t} = \left[M_1^{i,t}, M_2^{i,t}, M_3^{i,t} \ldots\ldots., M_d^{i,t}\right] \tag{2}$$

The algorithm will revise the location of crows by two methods, assuming the succeeding scenario: crow $j$ is hovering near it secrete place $M^{j,t}$ and crow $i$ choose to chase it to figure out its best food supply, which it will then steal. Two events could occur in this situation:

- Phase 1: Crow j is unaware that crow i is chasing it.

At the same time, crow $i$ approaches crow $j$, now crow $i$'s new location is changed by:

$$X^{i,t} + r_i * fl * \left(M^{j,t} - X^{i,t}\right) \tag{3}$$

where $r_i$ a random number with uniform distribution between [0,1] and flight length (*fl*) denotes the distance each crow covers in a solo flight.

- Phase 2: Crow j is aware that crow i is chasing it.

As an outcome, crow *j* tries to divert crow *i*'s attention away from its hiding place by moving elsewhere in the search room. The new location of crow *i* is determined at random. The following expression can also be used to express the two cases:

$$X^{i,t+1} = \begin{cases} X^{i,t} + r_i * fl * \left(M^{j,t} - X^{i,t}\right) & \text{if}_{\text{j}} \geq AP \\ a\,random\,positon & otherwise \end{cases} \tag{4}$$

where *AP* represents the awareness probability of crow *j* in generation *t*.

**Algorithm 1: Crow Search Algorithm**

```
Generate the population of N crows arbitrarily in a d
dimensional search space.
Calculate the position and memory of each crow.
Set awareness probability AP and flight length fl
while iteration < t_max
    For i = 1 to N
      Arbitrarily choose a crow j to trail crow i
      if  r_j ≥ AP
          X^{i,t+1} = X^{i,t} + r_i * fl * (M^{j,t} − X^{i,t})
      else
          X^{i,t+1} = any arbitrary position in the search
          space
      end if
    End for
Examine the viability of new position
Calculate the new location of the crow
Revise the memory of the crow
end while
```

## 4 Global Best Guided Binary CSA (G-BCSA)

In general, the ability of a population-based algorithm to maintain a balance between exploitation and exploration is used to evaluate its results. The locus of crow *i* is revised in Phase 1 by considering the memory of crow *j*. As a result, the algorithm's exploration capability is improved. However, it increases the likelihood of being stuck in a local optima stage. As a result, this algorithm is better at exploration than exploitation. The introduction of a global best driven methodology alleviates the

CSA's dilemma [14]. As defined in Eq. (5), the location of crow $i$ is modified by considering the optimal point of the corresponding iteration.

$$X^{i,t+1} = X^{i,t} + r_i * fl * \left(M^{j,t} - X^{i,t}\right) + r_i * \left(xbest - X^{i,t}\right) \tag{5}$$

where *xbest* denotes the global best value of the iteration and rand a random number between [0,1]. The pseudocode of G-BCSA is given in Algorithm 2.

A transfer function is an essential part of metaheuristic-based feature selection algorithms since it maps the continuous search space to the binary search space [3]. The transfer function determines the likelihood of a binary solution element being changed from 0 to 1 and vice versa. S-shaped and V-shaped transfer functions are the two most common types used in literature. The transfer function used in this study is defined in Eq. (6). The continuous to binary conversion is performed as per Eq. (7).

$$T\left(X^{i,t+1}\right) = \frac{1}{1 + e^{-X^{i,t}}} \tag{6}$$

$$X^{i,t+1} = \begin{cases} 1 \; if \; rand < T\left(X^{i,t+1}\right) \\ 0 \; if \; rand \geq T\left(X^{i,t+1}\right) \end{cases} \tag{7}$$

**Algorithm 2: Global Best Guided Binary Crow Search Algorithm**

```
Generate, the flock of N crows arbitrarily in binary
vector.
Calculate the position of each crow.
Set the memory of each crow.
Set awareness probability AP and flight length fl
while iteration < t_max
   for i = 1 to N
     Arbitrarily choose a crow j to trail crow i
     if  r_j ≥ AP
         X^{i,t+1} = X^{i,t} + r_i * fl * (M^{j,t} − X^{i,t}) + r_i * (xbest − X^{i,t})
         Convert X^{i,t+1} to binary according to eq (7)
     else
         X^{i,t+1} = any arbitrary position in the binary
         search space
     end if
   end for
Examine the viability of the new position
Calculate the new location of the crow
Revise the memory of the crow
end while
```

### 4.1 Fitness Function

FS issue has two goals. The first goal is to enhance the precision of classification by focusing on the key characteristics. The second goal is to pick the lowest possible number of features possible. The fitness function defined in Eq. (8):

$$\text{Fitness(X)} = \alpha\gamma_R(X) + \beta\frac{|X|}{|N|} \tag{8}$$

where $\gamma_R(X)$ is the error rate of classification, $|X|$ is the total amount of chosen features, $|N|$ is the cumulative number of features, $\alpha$ and $\beta$ are two constant parameters that determine the relative value of the two FS goals. Since the values of α and β are within the range [0,1], the correlation between α and β i.e., $\beta = (1 - \alpha)$ is supposed to be preserved. The values of α and β are adjusted as found in [3].

## 5 Results and Discussion

A set of algorithms, including the binary PSO, binary PSO with X-shaped transfer function (BPSO-X), binary WOA, binary CSA, binary CSA with time varying flight length (BCSA-TVFL) are contrasted to new proposed G-BCSA approach. The K-NN classifier ensures that the chosen features are the best suited. $K = 5$ is used in this study. The evaluation scheme for each dataset is a fivefold cross validation scheme. Python3 is used to execute this project and Matplotlib is used to plot the graphs.

### 5.1 Dataset Description

The proposed method is evaluated on 12 benchmark datasets shown in Table 1.

**Table 1** Datasets utilized in the experiments

| S.no | Name | Features | Samples |
|---|---|---|---|
| 1 | Amphibians | 23 | 189 |
| 2 | Breastcancer | 10 | 683 |
| 3 | Diabetes | 8 | 786 |
| 4 | Haberman | 3 | 306 |
| 5 | Immunotherapy | 7 | 90 |
| 6 | Ionosphere | 33 | 351 |
| 7 | Libras movement | 91 | 360 |
| 8 | Leukemia | 7129 | 72 |
| 9 | Mammographic masses | 6 | 961 |
| 10 | Sonar | 60 | 208 |
| 11 | Spambase | 58 | 4601 |
| 12 | Waveform | 40 | 5000 |

**Table 2** Parameter Values

| Algorithm | Parameters |
|---|---|
| BPSO, BPSO-X | Inertia w = 0.9, c1 = 0.5, c2 = 0.5 |
| BCSA | Flight Length = 2, Awareness Probability = 0.1 |
| BCSA-TVFL | Awareness Probability = 0.1 |

## 5.2 *Parameter Settings*

The parameters of optimizers are carefully chosen through a series of trial-and-error processes to determine the best possible settings. Table 2 lists the parameters that were used in all of the experiments.

## 5.3 *Experimental Results*

Table 3 presents the experimental findings of the average and standard deviation of classification accuracy for 30 independent runs. Table 4 compares the average fitness measure and standard deviation values. Table 5 shows the average feature selection ratio for all the approaches. Figure 1 shows the convergence curves of G-BCSA and other algorithms on the 12 datasets.

Table 3 shows that the proposed G-BCSA has achieved 75% times higher accuracy than its competitors among all the datasets. It attained low classification accuracy for only 3 datasets namely diabetes, ionosphere and waveform. Table 4 demonstrates that G-BCSA has attained the ideal fitness value among competitors 58.33% of the

**Table 3** Average and standard deviation to compare the classification accuracy of G-BCSA with competitors

| Dataset | | BPSO | BWOA | BPSO-X | BCSA | BCSA-TVFL | G-BCSA |
|---|---|---|---|---|---|---|---|
| Amphibians | Avg | 0.9707 | 0.9562 | 0.9968 | 0.9848 | 0.9973 | **0.9981** |
| | Std | 0.0085 | 0.0132 | 0.0131 | 0.0120 | 0.0107 | 0.0067 |
| Breastcancer | Avg | 0.9989 | 0.9984 | 0.9998 | 1.0 | 1.0 | **1.0** |
| | Std | 0.0065 | 0.0031 | 0.0014 | 0.0 | 0.0 | 0.0 |
| Diabetes | Avg | **0.8156** | 0.8186 | 0.8079 | 0.8140 | 0.8146 | 0.8116 |
| | Std | 0.0195 | 0.0148 | 0.0116 | 0.0140 | 0.0088 | 0.0123 |
| Haberman | Avg | 0.8620 | 0.8517 | 0.8891 | 0.8820 | 0.8835 | **0.8849** |
| | Std | 0.0300 | 0.0197 | 0.0319 | 0.0188 | 0.0255 | 0.0334 |
| Immunotherapy | Avg | 0.9802 | 0.9988 | 0.9977 | 0.9964 | 0.9988 | **0.9990** |
| | Std | 0.0262 | 0.0111 | 0.0175 | 0.0132 | 0.0078 | 0.0185 |
| Ionosphere | Avg | 0.9523 | 0.9691 | 0.9766 | 0.9690 | **0.9769** | 0.9676 |
| | Std | 0.0099 | 0.0099 | 0.0157 | 0.0072 | 0.0128 | 0.0094 |
| Libras movement | Avg | 0.8847 | 0.9054 | 0.8989 | 0.9146 | 0.8994 | **0.9148** |
| | Std | 0.0207 | 0.0182 | 0.0148 | 0.0123 | 0.0104 | 0.0187 |
| Leukemia | Avg | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | **1.0** |
| | Std | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Mammographic masses | Avg | 0.8981 | 0.8972 | 0.8787 | 0.8817 | 0.8871 | **0.8998** |
| | Std | 0.0106 | 0.0086 | 0.0081 | 0.0087 | 0.0090 | 0.0088 |
| Sonar | Avg | 0.9523 | 0.9617 | 0.9752 | 0.9683 | 0.9680 | **0.9759** |
| | Std | 0.0131 | 0.0148 | 0.0266 | 0.0129 | 0.0152 | 0.0114 |
| Spambase | Avg | 0.9200 | 0.9278 | 0.9044 | 0.9267 | 0.9323 | **0.9360** |
| | Std | 0.0085 | 0.0066 | 0.0054 | 0.0041 | 0.0026 | 0.0045 |
| Waveform | Avg | 0.8265 | 0.8341 | **0.8381** | 0.8269 | 0.8353 | 0.8212 |
| | Std | 0.0053 | 0.0054 | 0.0044 | 0.0079 | 0.0015 | 0.0064 |

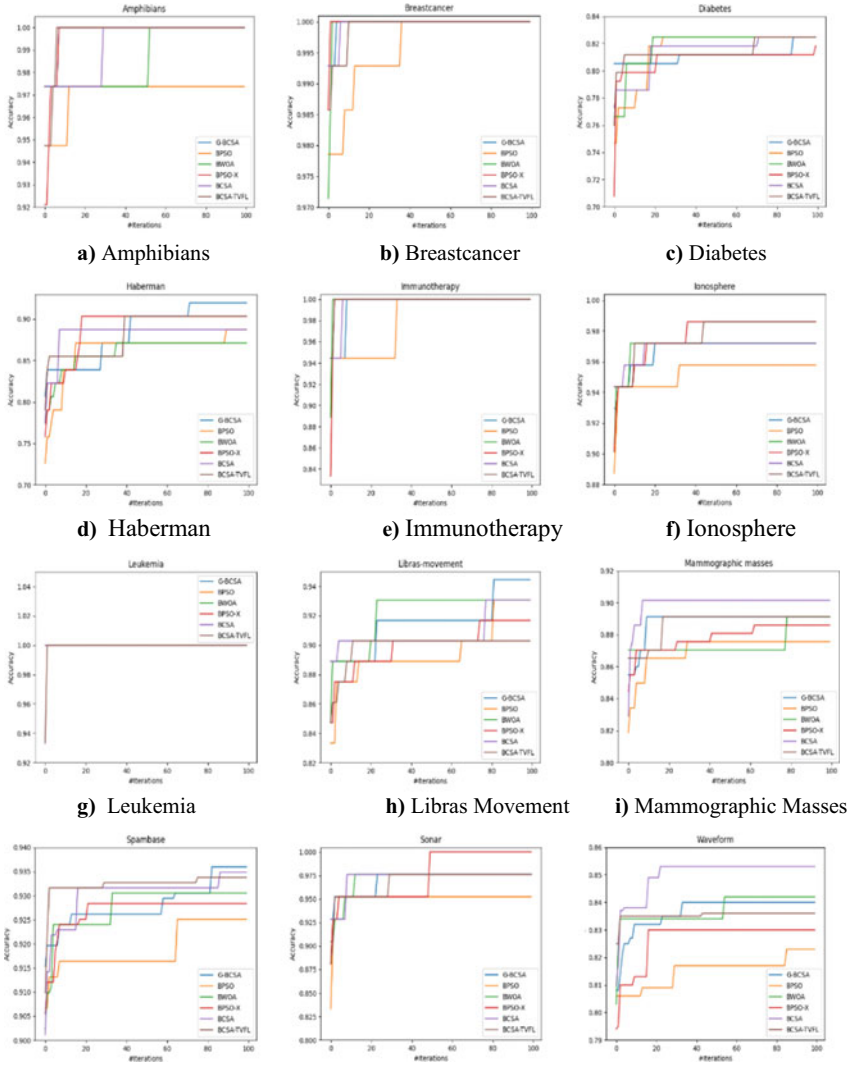
time. Following it BPSO-X has achieved best fitness value 33.33% among others. Table 5 reveals that G-BCSA outperforms other methods on majority of datasets in terms of feature selection ratio. G-BCSA obtained smallest feature selection ratio in seven out of twelve datasets. BPSO-X has achieved lowest feature selection ratio in 3 datasets, BPSO in one dataset and BCSA in one dataset but the classification accuracy is also low in these datasets.

**Table 4** **A**verage and standard deviation to compare the fitness of G-BCSA with competitors

| Dataset | | BPSO | BWOA | BPSO-X | BCSA | BCSA-TVFL | G-BCSA |
|---|---|---|---|---|---|---|---|
| Amphibians | Avg | 0.0042 | 0.0260 | **0.0033** | 0.0090 | 0.0066 | 0.0071 |
| | Std | 0.0166 | 0.0064 | 0.0131 | 0.0010 | 0.0104 | 0.0068 |
| Breastcancer | Avg | 0.0115 | 0.0126 | 0.0055 | 0.0055 | 0.0055 | **0.0019** |
| | Std | 0.0053 | 0.0019 | 0.0028 | 0.0014 | 0.0031 | 0.0066 |
| Diabetes | Avg | 0.1862 | 0.1721 | 0.1887 | 0.1850 | 0.1798 | **0.1715** |
| | Std | 0.0167 | 0.0179 | 0.0116 | 0.0136 | 0.0088 | 0.0059 |
| Haberman | Avg | 0.1117 | 0.1217 | 0.1024 | 0.1151 | 0.0991 | **0.0831** |
| | Std | 0.0307 | 0.0186 | 0.0310 | 0.0188 | 0.0253 | 0.0333 |
| Immunotherapy | Avg | 0.0028 | 0.0071 | **0.0014** | 0.0042 | 0.0014 | 0.0042 |
| | Std | 0.0247 | 0.0207 | 0.0157 | 0.0105 | 0.0079 | 0.0016 |
| Ionosphere | Avg | 0.0471 | 0.0488 | 0.0152 | 0.0192 | 0.0201 | **0.0146** |
| | Std | 0.0087 | 0.0127 | 0.0191 | 0.0137 | 0.0021 | 0.0105 |
| Libras movement | Avg | **0.0602** | 0.0610 | 0.0876 | 0.0744 | 0.0750 | 0.0624 |
| | Std | 0.0206 | 0.0155 | 0.0076 | 0.0104 | 0.0151 | 0.0178 |
| Leukemia | Avg | 0.005 | 0.0063 | **0.0049** | 0.0062 | 0.0060 | 0.0062 |
| | Std | 0.0002 | 0.0065 | 0.0005 | 0.0065 | 0.0004 | 0.0004 |
| Mammographic masses | Avg | 0.1291 | 0.1054 | 0.1222 | 0.1137 | 0.1137 | **0.1050** |
| | Std | 0.0127 | 0.0158 | 0.0007 | 0.0112 | 0.0093 | 0.0091 |
| Sonar | Avg | 0.0289 | 0.0300 | **0.0040** | 0.0285 | 0.0299 | 0.0229 |
| | Std | 0.0209 | 0.0130 | 0.0267 | 0.0114 | 0.0148 | 0.0101 |
| Spambase | Avg | 0.0762 | 0.0735 | 0.0752 | 0.0656 | 0.0706 | **0.0653** |
| | Std | 0.0086 | 0.0066 | 0.0059 | 0.0065 | 0.0025 | 0.0045 |
| Waveform | Avg | 0.1838 | 0.1679 | 0.1740 | 0.1720 | 0.1560 | **0.1550** |
| | Std | 0.0070 | 0.0046 | 0.0082 | 0.0029 | 0.0255 | 0.0083 |

**Table 5** Feature Selection ratio of G-BCSA with competitors

| Dataset | BPSO | BWOA | BPSO-X | BCSA | BCSA-TVFL | G-BCSA |
|---|---|---|---|---|---|---|
| Amphibians | 0.51 | 0.60 | 0.66 | 0.57 | 0.62 | **0.51** |
| Breastcancer | 0.70 | 0.7 | 0.55 | 0.65 | 0.55 | **0.50** |
| Diabetes | 0.50 | 0.64 | 0.50 | 0.81 | 0.60 | **0.50** |
| Haberman | 0.66 | 0.78 | 0.66 | **0.58** | 0.66 | 0.66 |
| Immunotherapy | **0.29** | 0.79 | 0.57 | 0.43 | 0.57 | 0.42 |
| Ionosphere | 0.58 | 0.66 | 0.44 | 0.68 | 0.61 | **0.40** |
| Libras movement | 0.50 | 0.60 | **0.46** | 0.63 | 0.67 | 0.58 |
| Leukemia | 0.51 | 0.62 | **0.49** | 0.62 | 0.60 | 0.51 |
| Mammographic masses | 0.40 | 0.72 | 0.70 | 0.70 | 0.70 | **0.31** |
| Sonar | 0.48 | 0.59 | **0.43** | 0.45 | 0.58 | 0.48 |
| Spambase | 0.56 | 0.60 | 0.50 | 0.65 | 0.50 | **0.50** |
| Waveform | 0.68 | 0.58 | 0.57 | 0.52 | 0.61 | **0.50** |




**a)** Amphibians **b)** Breastcancer **c)** Diabetes

**d)** Haberman **e)** Immunotherapy **f)** Ionosphere

**g)** Leukemia **h)** Libras Movement **i)** Mammographic Masses

**j)** Spambase **k)** Sonar **l)** Waveform

**Fig. 1 a to l** Convergence curves of G-BCSA and other state of the art methodology on Amphibians, Breastcancer, Diabetes, Haberman, Immunotherapy, Ionosphere, Leukemia, Libras movement, Mammographic masses, Spambase, Sonar and Waveform datasets

## 6 Conclusion

In this paper, the global best driven technique is applied to CSA in order to improve its performance in terms of exploitation capability. We have compared it to 5 different variations of other metaheuristics that have been recently used in the literature on 12

UCI datasets to test the efficacy of our suggested algorithm. In terms of classification accuracy and fitness, the contrast clearly indicates that the proposed approach is superior. For future work, we can apply the proposed approach on different FS methods and can also merge it with other population dependent metaheuristic algorithms.

## References


1. Mafarja M, Qasem A, Heidari AA, Aljarah I, Faris H, Mirjalili S (2019) Efficient hybrid nature-inspired binary optimizers for feature selection. Cogn Comput 12(1):150–175
2. Beheshti Z (2021) A novel x-shaped particle swarm optimization. Soft Compute 25:3013–3042
3. Hussien AG, Oliva D, Houssein EH, Juan AA, Yu X (2020) Binary whale optimization algorithm for dimensionality reduction. Mathematics 8:1821
4. Chaudhuri A, Sahu TP (2021) Binary Jaya algorithm based on binary similarity measure for feature selection. J Ambient Intell Human Comput 1–18. https://doi.org/10.1007/s12652-021-03226-5
5. Chaudhuri A, Sahu TP (2021) A hybrid feature selection method based on Binary Jaya algorithm for micro-array data classification. Comput Electr Eng 90:106963
6. Majhi SK, Sahoo M, Pradhan R (2019) Oppositional crow search algorithm with mutation operator for global optimization and application in designing FOPID controller. Evolving Syst 12(2):463–488. https://doi.org/10.1007/s12530-019-09305-5
7. Pamir JN, Mohsin SM, Iqbal A, Yasmeen A, Ali I (2019) A hybrid bat-crow search algorithm based home energy management in smart grid. In: Barolli L, Javaid N, Ikeda M, Takizawa M (eds) Complex, Intelligent, and Software Intensive Systems. CISIS 2018. AISC, vol 772. Springer, Cham. https://doi.org/10.1007/978-3-319-93659-8_7
8. Anter A, Hassenian AE, Oliva D (2019) An improved fast fuzzy c-means using crow search optimization algorithm for crop identification in agricultural. Expert Syst Appl 118:340–354
9. Arora S, Singh H, Sharma M, Sharma S, Anand P (2019) A new hybrid algorithm based on grey wolf optimization and crow search algorithm for unconstrained function optimization and feature selection. IEEE Access 7:26343–26361
10. Sayed GI, Hassanien AE, Azar AT (2019) Feature selection via a novel chaotic crow search algorithm. Neural Comput Appl 31(1):171–188
11. Anter AM, Ali M (2020) Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. Soft Comput 24:1565–1584
12. Chaudhuri A, Prasad Sahu T (2020) Feature selection using binary crow search algorithm with time varying flight length. Expert Syst Appl 168:114288
13. Das S, Sahu TP, Janghel RR (2020) PSO-based group-oriented crow search algorithm (PGCSA). Eng Comput 38(2):545–571
14. Roy R, Sahu TP, Nagwani NK, Das S (2021) Global best guided crow search algorithm for optimization problems. In: Kumar R, Singh VP, Mathur A (eds) Intelligent Algorithms for Analysis and Control of Dynamical Systems. AIS. Springer, Singapore. https://doi.org/10.1007/978-981-15-8045-1_2

# GWCM: Grid Based Weighted Clustering Method for Wireless Ad-Hoc Network


**Virendra Dani, Priyanka Kokate, and Surbhi Kushwah**

**Abstract** Technology must be able to support vast networks of consumers in order to be economically feasible. One issue is that addressing and routing in ad hoc networks do not scale up as easily as they do on the Internet. Clustering is a method for generating and managing hierarchical addresses in ad hoc networks. The Grid-based Weighted Clustering Method, or "GWCM," is proposed in this paper for use in wireless ad-hoc networks. This approach considers the number of nodes that a cluster-head can support, as well as the signal strength, mobility, and connectivity of each node. The proposed GWCM was introduced using the network simulation (NS2) environment and the AODV routing protocol to incorporate the proposed algorithm. As nodes relay high complexity, the findings show that the proposed GWCM increases network node flexibility and performance. Additionally, compared developed method to old approach of Energy-Efficient Clustering and DRL-Based Sleep Scheduling and found acceptable performance of GWCM using average comparison of End to end network delay, remain energy and packet delivery ratio.

**Keywords** Ad hoc network · Cluster-head · Clustering · Communication · Grid network · Mobility · Node · Node weight

## 1 Introduction

In mobile wireless ad-hoc networks (MANETs), wireless networking and the lack of centralized administration present various challenges [1]. Since node mobility causes frequent connection failure and activation, routing algorithms respond to topology

V. Dani (✉) · P. Kokate · S. Kushwah
Computer Science and Engineering Department, Shivajirao Kadam Institute of Technology and Management, Indore, India
e-mail: virendradani.cs@gmail.com

P. Kokate
e-mail: priyankakokate@skitm.in

S. Kushwah
e-mail: surbhikushwah@skitm.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_46

changes, increasing network control traffic [2]. Given the large number of MNs that MANETs can contain, a hierarchical structure would scale better [3]. As a result, building hierarchies among the nodes, so that the network topology can be abstracted, is one promising approach to addressing routing problems in MANET environments. Clustering is the term for this operation, and clusters are the substructures that are collapsed in higher levels [4]. Clustering in MANETs is not a new concept; several algorithms have been proposed that consider various metrics and concentrate on various objectives [5, 6].

For MANET that does not have flat topology, AODV is a very basic, reliable, and efficient routing protocol. It collects much of the beneficial principles from algorithms by DSR and DSDV [7]. The acquisition of strictly on-demand routes makes AODV a very effective and preferred MANET algorithm. The knowledge about the active neighbors for this route is preserved so that when a connection along a path to the destination splits, all active source nodes can be informed [8, 9].

## 2 Literature Review

Qi et al. [10] suggested the robust, energy-efficient weighted clustering algorithm as a new algorithm (RE2WCA). By limiting minimum iteration times, the RE2WCA takes residual energy and group mobility into account for homogeneous energy consumption. Advantage of this method is ensures even distribution of nodes and disadvantage that this not suitable for large scale networks.

Pal et al. [11] suggested EEWC a new energy efficient clustering method based on a genetic algorithm with a newly established objective function. Advantage of this method nodes equally share load up to some extent and disadvantage is that extra overheads due to dynamic clustering.

Pathak et al. [12] have suggested an improved secure clustering algorithm that will improve (advantage) network reliability by reducing clustering overhead and decreasing cluster head adjustments and disadvantage that this is not appropriate for random deployment.

Behera et al. [13] investigated an efficient CH election technique that rotates the CH position among nodes with greater energy levels than others. It improves network performance and disadvantage is that for large scale network, computational overhead will increase.

Manikanthan et al. [14] devised a protocol that incorporates grid-based mobile communication network construction, proficient path selection through cluster head selection, and data communication. This method suitable for defined small scale networks and increase life time of network as well expensive CH election process in terms of energy consumption is the main drawback.

# 3 Proposed Work

## *3.1 Methodology*

This section explains how to use the proposed method for finding cluster head using modified AODV protocol in wireless network for expanding network lifetime. The proposed GWCM works in four main phases as described in Fig. 1. This figure depicts the function of GWCM cluster head selection based on Grid environment. This method comprises 4 different phases which can be summarizing in next section. In given diagram, a first layer is the entire network area is transformed in an equally partitioned area called grid.

This method comprises 4 different phases which can be summarizing in next section. In above diagram, a first layer is the entire network area is transformed in an equally partitioned area called grid. After creation of grid, each node of this grid is evaluated for finding the efficient working nodes using weight computation. After that, calculate threshold value of all entire nodes by means of different parameters. Finally, the concluded nodes are established as the cluster head of network.

- **Parameter Definition:** The node QoS factors are chosen in this step to perform the threshold computation.
- **Compute Weight:** The weights of all nodes are computed as by their threshold value of selected parameters.
- **Proposed Algorithm for finding cluster-head:** Finally, GWCM Algorithm is prepared on the basis of decision making using node value calculation

In order to understand the entire process of the proposed algorithm the flow diagram of the algorithm is also developed. The Fig. 2 shows the description of the entire process.

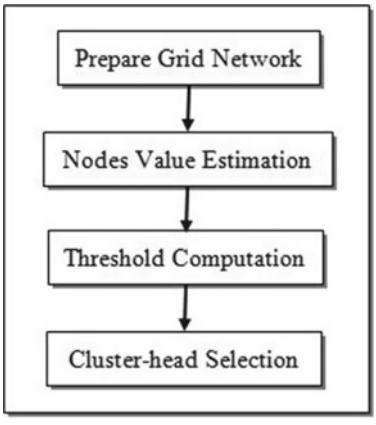



**Fig. 1** Step-by-step function of proposed work

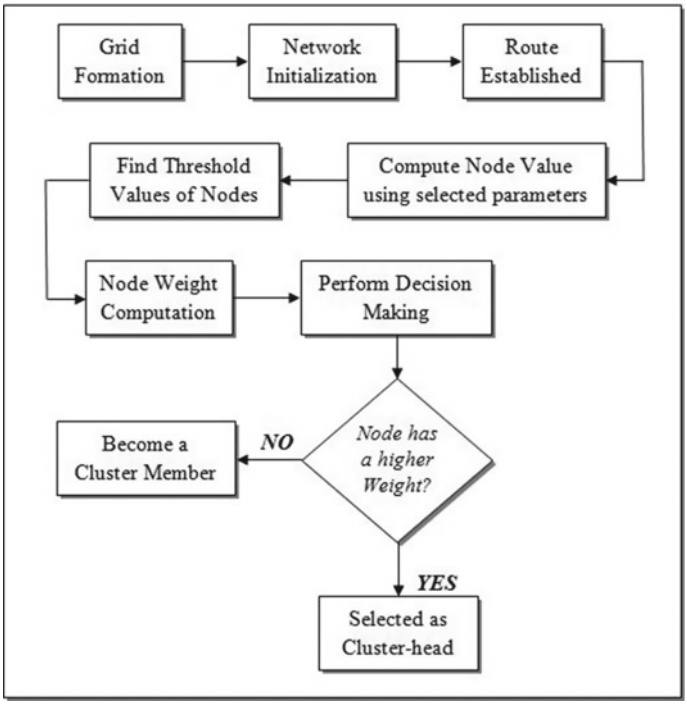



**Fig. 2** Working process of CH Selection

This process demonstrates cluster head selection in wireless as hoc network. Firstly, create Grid for node distribution of network connection. Once Grid formed then initialize network using number of nodes. After estimating value of parameter for different node than find threshold value of each node and calculate weight for all node by means of weighting factors. Consequently, final weight is computed then performs decision making for all nodes. In this, scenario, apply checks for finding cluster head by comparing and exchanging value of nodes. If got higher value of node, then the particular node selected as cluster-head and rest of the node always is a member of cluster.

**Parameter Definition.** In order to achieve the required goals, the following three main parameters is selected to propose GWCM i.e. "Grid based Weighted Clustering Method".

*Signal Strength ($S_S$).* The signal strength of a node shows the transmission power ability thus for more optimal node selection the signal strength of the node is used. The signal Strength can be denoted by $\mathbf{S_S}$.

*Node Mobility ($M_N$).* If nodes' locations shift over time, they must update their position estimates on a regular basis. To compute the relative mobility of nodes, start with the network node in place $(x_1, y_1)$ and change it to $(x_2, y_2)$ after a time $\Delta t$. As a

result, the difference in place between points $(x_1, y_1)$ and $(x_2, y_2)$ is calculated. The node mobility is characterized by $M_N$. Given Eq. (1) shows node distance from their respective place by differencing them to each other, D is the distance traveled by the node in time $\Delta t$. Therefore, rate in change in the node position called mobility is given by Eq. (2).

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

$$M_N = \frac{D}{\Delta t} \tag{2}$$

*Energy Consumption ($E_C$).* The amount of energy consumed in a given amount of time. It is represented by $E_C$.

**Weight Computation.** The normalization method is needed to calculate the weights in order to pick the most suitable cluster head from the available cluster representatives. As a result, some additional coefficients were needed for weight factor computation. In this case, it must reform the network from selecting the CH by which the clustering protocol runs and generates a response based on the chosen output parameter. After configuring the network, the selected parameters are computed for all the participating nodes. Therefore, the mean values of nodes are computed as:

$$S_{S(n)} = \frac{1}{N} \sum_{i=1}^{N} S_{S(i)} \tag{3}$$

$$M_{N(n)} = \frac{1}{N} \sum_{i=1}^{N} M_{N(i)} \tag{4}$$

$$E_{C(n)} = \frac{1}{N} \sum_{i=1}^{N} E_{C(i)} \tag{5}$$

Above Eqs. (3), (4) and (5) shows the average threshold of parameter selected for computation of average Weight. In this solution the entire simulation area is sub divided in to uniform Grid (N × N) and additionally the evaluation of node for constructing cluster head is performed on the basis of regional constraints. Such as the node locality in the specified region as Grid sub division that helps to maintain connectivity throughout the simulation area. Moreover, the CH election depends on the weight, which is computed using following function. Now, using Eq. (3), (4) and (5), calculate weight using these node's values such that

$$W_g = w1.S_S + w2.M_N + w3.E_C \tag{6}$$

Above Eq. (6) demonstrate total weight of the threshold parameter by selecting coefficients and coefficients$w1$, $w2$ and $w3$ can be chosen at random based on design

considerations, but they must be between 0 and 1 and adhere to the specified conditions. Where $W_g$ is the combined weight, $E_C$ is the consumed energy, and $M_N$ is the node mobility and $S_S$ is the node signal strength,$w1$, $w2$and $w3$ is the weight factors.

$$w1 + w2 + w3 = 1 \tag{7}$$

Equation (7), depicts that coefficient can be selected according to their normalize value and total of the weight coefficient should be 1. It must change the network by selecting the CH, which performs the clustering process and generates a response based on the output parameter chosen. According to the given condition of selection criteria less amount of weight is necessary to be a cluster head.

### *3.2 Proposed Algorithm*

For efficient use of network, here is describing algorithm formulation for finding CH to perform efficient use of network. The algorithm can be described in Table 1:

In parallel, compared this proposed approach Grid based Weighted Clustering Method i.e. *GWCM* to previously developed approach Energy-Efficient Scheduling using the Deep Reinforcement Learning i.e. *EES-DRL* which is used wireless sensor network configuration for cluster head selection proposed by Sinde et al. [15].

## 4 Simulation and Result Discussion

### *4.1 Simulation Setup*

Network Simulator 2 (NS2) [16] tool, which is part of the Ubuntu operating system, is used to implement GWCM. Because NS2 is a discrete event simulator that can simulate many types of networks, it was chosen for the suggested energy-efficient clustering strategy in the wireless ad hoc environment. In addition, the following setup Table 2 is ready for performance assessment and simulation.

A network of small size is setup that contains 30, 60, 90, 120 and 150 nodes. The constant bit rate (CBR) application creates packet through connection based on UDP, as it is connectionless protocol. CBR packet size chosen is 512 KB. Positions of the nodes are defined manually in TCL script. Topography area show that where the network nodes are moves in this area in all direction. First of all, the parameters are evaluated for the AODV protocol when it works normally.

**Table 1** Proposed GWCM algorithm

**Input: Number of Grid (g), Number of network Node (NN)**
**Output: Cluster Head**

***Process:***

**1:**[length, width] = get Area Simulation ( )

**2:**Grid length = length / N

**3:**Grid width = width/N

**4:**$Grid_k$=create Grid(Grid length, Grid width)

**5:**Populate Random mobile nodes in network;

**6:**Compute Weight Using

$$W_g = w1.S_S + w2.M_N + w3.E_C$$

**7:**Compute Parameter threshold weight;

**a:**Signal Strength ($S_S$), **b:** Node Mobility ($M_N$), **c:**Consuned Energy ($E_C$)

**8:**Comparison of individual node weight to threshold weight;

**9:**If node's $W_g$> threshold W

**a:**Node can be CH

Else

**b:**Do again from step 8

**10:** End if

**11:**Higher weight of node selected as cluster head

**12:** Return CH

**Table 2** Simulation setup definition

| Simulation parameters | Values |
|---|---|
| Antenna model | Omni directional |
| Topography area | 1000 ×1000 |
| Radio-propagation model | Two ray ground |
| Channel type | Wireless channel |
| No of mobile nodes | 30, 60, 90, 120,150 |
| Routing protocol | AODV |
| Packet size | 512 KB |

## *4.2 Result Discussion*

This section provides the detail discussion about the obtained performance. In order to compute the performance of the proposed routing technique the different experiments on 30, 60, 90, 120 and 150 nodes are performed.

**End to End Network Delay.** In network, when data packet transmits from ultimate source to ultimate destination then there is time taken for transmission by packet is called End to end day. End-to-end delay of data packets includes all possible delays caused by buffering during route discovery, queuing at interface queue, propagation and transfer time.

Figure 3 shows the E2E delay of the proposed technique and EES-DRL technique. The findings reveal that the network's end-to-end latency is higher in EES-DRL approach than in the proposed cluster-based routing GWCM. In addition, the growing number of network nodes has an effect on end-to-end latency.

**Remain Energy.** The nodes consume a portion of their original amount of energy through packet transfer and other network activities.

The amount of energy used in network nodes during the various experiments is depicted in Fig. 4. The experiments are carried out with 30, 60, 90, 120, and 150 nodes. The energy calculation is given in terms of Jules. According to the findings of the experiments, the proposed clustering strategy uses less energy than the traditional EES-DRL approach. As a result, when opposed to standard network architectures, the proposed clustering solution is network efficient.

**Compare Packet Delivery Ratio.** The ratio of data packets obtained by destinations to those provided by sources is known as the packet delivery ratio.

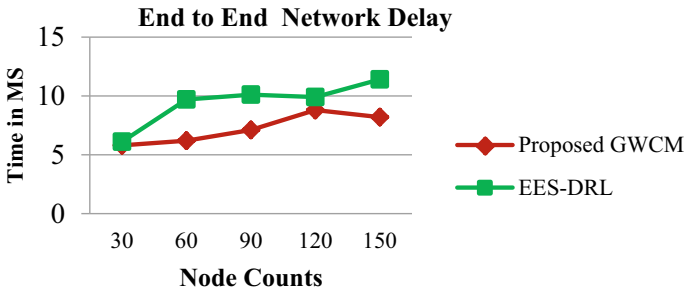



**Fig. 3** End to end network delay

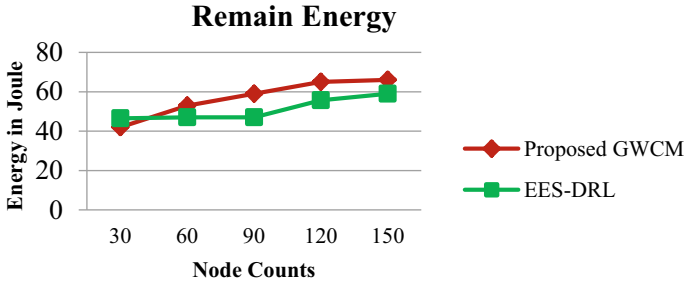



**Fig. 4** Remain energy

Figure 5 depicts the packet delivery ratio of traditional EES-DRL routing and proposed cluster-based techniques. According to experiment, the proposed approach is capable of successfully delivering more packets than the conventional EES-DRL.

## 5 Conclusion

The key properties of a number of clustering algorithms that aid in the establishment of hierarchical MANETs are presented. The cluster configuration reliability, the regulate overhead of cluster building and conservation, the energy usage of mobile nodes with dissimilar cluster-related standing, and the traffic load supply in clusters are all significant issues to consider in a cluster-based MANET.As a result, a solution is needed that confirms the selection of a consistent cluster head capable of handling high traffic while maintaining cluster head stability. As a result, the proposed research focuses on the advancement of cluster head selection. The suggested clustering strategy uses three factors to promote resource preservation in wireless ad-hoc networks: consumed energy, signal strength and node mobility. On the other hand, work demonstrates average compare performance of Proposed GWCM and EES-DRL using difference performance factor using Fig. 6.

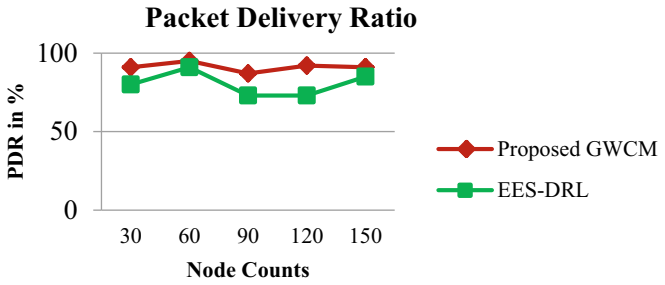



**Fig. 5** Compare PDR

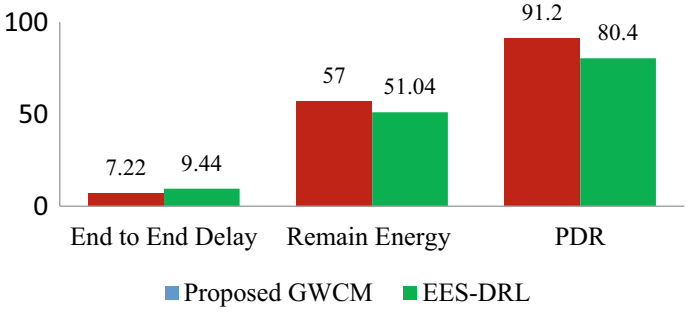



**Fig. 6** Compare average performance

## References


1. Perkins CE (2008) Ad hoc networking. Pearson Education India
2. Reham A, El-Sayed M, El-kenawy T, Saber M (2019) A proposed routing protocol for mobile ad hoc networks. Int J Comput Appl 975:8887
3. Sri TS, Prasad JR, Kumar RK (2018) Distributed and adaptive efficient energy aware routing procedure for MANETs
4. Pathak S, Jain S, Borah S (2021) Clustering algorithms for MANETs: a review on design and development. In: Soft computing techniques and applications, pp 563–578
5. Ahmad M, Hameed A, Ikram AA, Wahid I (2019) State-of-the-art clustering schemes in mobile ad hoc networks: objectives, challenges, and future directions. IEEE Access 7:17067–17081
6. Dani V, Bhati N, Bhati D (2021) EECT: energy efficient clustering technique using node probability in ad-hoc network. In: 12th international conference on innovations in bio-inspired computing and applications. Advances in intelligent systems and computing, vol 1372. Springer, Heidelberg, pp 187–195
7. Singh A, Singh G, Singh M (2018) Comparative study of OLSR, DSDV, AODV, DSR and ZRP routing protocols under black-hole attack in mobile ad hoc network. In: Intelligent communication, control and devices. Springer, Singapore, pp. 443–453
8. Fanian F, Rafsanjani MK (2019) Cluster-based routing protocols in wireless sensor networks: a survey based on methodology. J Netw Comput Appl 142:111–142
9. Jubair MA, Khaleefah SH, Budiyono A, Mostafa SA, Mustapha A (2018) Performance evaluation of AODV and OLSR routing protocols in MANET environment. Int J Adv Sci Eng Inf Technol 8(4):1277–1283
10. Qi H, Liu F, Xiao T, Su J (2018) A robust and energy-efficient weighted clustering algorithm on mobile ad hoc sensor networks. Algorithms 11(8):116
11. Pal R, Yadav S, Karnwal R (2020) EEWC: energy-efficient weighted clustering method based on genetic algorithm for HWSNs. Complex Intell Syst 6:1–10
12. Pathak S, Jain S (2017) An optimized stable clustering algorithm for mobile ad hoc networks. EURASIP J Wirel Commun Netw 2017(1):1–11
13. Behera TM, Mohapatra SK, Samal UC, Khan MS, Daneshmand M, Gandomi AH (2019) Residual energy-based cluster-head selection in WSNs for IoT application. IEEE Internet Things J 6(3):5132–5139
14. Manikanthan SV, Padmapriya T (2019) An efficient cluster head selection and routing in mobile WSN. Int J Interact Mob Technol 13(1):56–70
15. Sinde R, Begum F, Njau K, Kaijage S (2020) Refining network lifetime of wireless sensor network using energy-efficient clustering and DRL-based sleep scheduling. Sensors 20(5):1540
16. Rehmani MH, Saleem Y (2015) Network simulator NS-2. In: Encyclopedia of information science and technology, 3rd edn. IGI Global, pp 6249–6258

# Hybrid Deep Learning Approach for Brain Tumor Segmentation and Classification


**Ayalapogu Ratna Raju, Suresh Pabboju, and Ramisetty Rajeswara Rao**

**Abstract** Brain tumor classification plays a significant part in the analysis and treatment of brain tumors. This paper introduces a hybrid deep learning methodology for the classification of brain tumor. At first, the input image is pre-processed, and U-Net approach is employed for performing tumor segmentation process. In the feature extraction module, the statistical, Discrete Wavelet Transform (DWT), and the shape features are extracted from the partitioned tumor portions of the image. Meanwhile, the classification is done using the developed hybrid learning approach, which is devised by the integration of the Taylor Improved Invasive Weed Optimization-based Deep Neural Network (Taylor IIWO-based DNN), and the (HCS-based DBN). The results obtained from the Taylor IIWO-based DNN and the HCS-based DBN are fused together by the Tversky index for classifying the output. The developed hybrid deep learning technique achieved the maximal accuracy of 0.963, higher sensitivity of 0.975, and maximum specificity of 0.937.

**Keywords** Belief network · Brain tumor · Deep neural network · Deep tversky index · U-Net

## 1 Introduction

The occurrence of tumor in brain is due to the abnormal cell growth in the brain of human [1]. Based on the requirements of medical domains, and the clinical field research, segmentation process has developed into a major part in the medical fields, and has been generally concerned for an extensive time [2, 3]. Brain tumor segmentation is employed for partitioning the abnormal tissues, such as fluid-filled or solid

A. R. Raju (✉)
Mahatma Gandhi Institute of Technology, Hyderabad, India
e-mail: ratname@gmail.com

S. Pabboju
Chaitanya Bharathi Institute of Technology, Hyderabad, India

R. R. Rao
University College of Engineering, JNTUK, Vizianagaram, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_47

from the normal tissues of brain [4]. However, the brain tumor segmentation remains as a difficult procedure because of the complicated brain tumor structure, blurred image boundaries. In addition, the techniques for segmenting the brain tumor [5] can be broadly categorized into three types, such as automatic segmentation, semi-automatic and manual segmentation. An automatic classification of the brain tumor is utilized to provide guidance for the clinical analysts to follow the appropriate treatment handling choice [6]. Meanwhile, tumor region consists of non-enhancing, enhancing tumor, necrosis, and edema [7]. The brain tumor cells are segmented using the active contour model to partition the Region of Interest (ROI) from the brain tumor cells by extracting the boundary. Several techniques have been employed for classifying the brain tumor [8]. Region-based Active Contour Model (RACM) [9] was employed to segment the appropriate portions of tumor from the magnetic resonance image. The Bat optimization algorithm [10, 11], AdaBoost classifier [12], the improved tumor-cut algorithm [13], and SVM [14] have been employed for automatic classification and the segmentation of MRI image. The main contribution of this research are described as follows,

- *Proposed hybrid deep learning approach:* An effective approach for classifying the brain tumor is devised by developing a hybrid deep learning approach. Moreover, the developed hybrid deep learning technique is designed by the hybridization of the Taylor IIWO-driven DNN and the HCS-driven DBN. Moreover, the Tversky index is employed to classify the output by integrating the results of Taylor IIWO-based DNN and HCS-based DBN.

## 2 Literature Survey

In this section, various existing brain tumor classification approaches along with their disadvantages are explained**.** Jianxinzhang et al. [3] developed an Attention Gate Residual U-Net model (AGResU-Net) for partitioning the brain tumor automatically. This method effectively minimized the dataset heterogeneity. However, this method failed to employ the 3D network model for enhancing the segmentation performance of the AGResU-Net. Mzoughi et al. [15] introduced a deep multi-scale three-dimensional convolutional neural network (3D CNN) for classifying the glioma brain tumor. Here, the global and the local contextual features are extracted using the deep learning method through the 3D kernel size. This method achieved effective classification performance, but this method failed to consider a new approach, called Capsule networks (CapsNet) for effectively classifying the MRI brain tumor. Díaz-Pernas et al. [16] devised a multi-scale CNN for classifying and segmenting the brain tumor. Here, the input images were processed in three spatial scales by considering the various processing pathways. This technique achieved better brain tumor classification performance, but the major challenge lies in utilizing this method in satellite imagery fields. Krishna Kumar and Manivannan [17] designed a rough K means clustering algorithm and multi kernel Support Vector Machine (SVM) for effectively segmenting the brain tumor. Here, the Gabor wavelet transform (GWT)

was employed for extracting the necessary features from pre-processed image. This method improved accuracy of classification, but this technique failed to enhance the reliability and the optimization performance.

## 3 Proposed Hybrid Deep Learning Approach for the Classification of Brain Tumor

This section illustrates the effective brain tumor classification by designing and developing a new hybrid deep learning approach. The developed hybrid learning approach is derived by the combination of TaylorIIWO-based DNN algorithm and the HCS-DBN. Here, Taylor IIWO is formed by integrating the Taylor series [18] with the IIWO [19], whereas the HCS is formed based on the hybridization of Harmony Search (HS) algorithm [20] and Crow search algorithm (CSA) [21]. Figure 1 depicts the schematic view of the developed hybrid deep learning approach.

### *3.1 Image Acquisition*

The input image is collected from the dataset and is utilized to classify the brain tumor. Consider a dataset $D$ with $e$ brain images, which is indicated in Eq. (1),

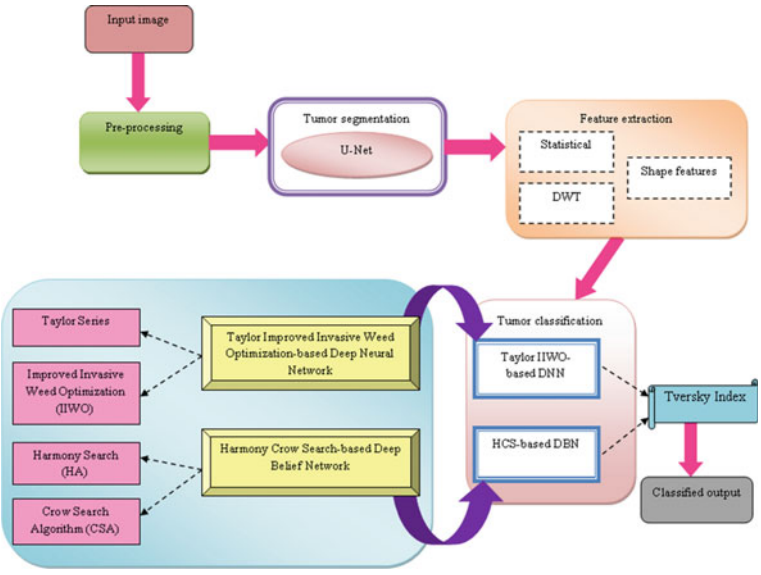



**Fig. 1** Schematic view of the developed hybrid deep learning technique

$$D = \{I_1, I_2, ...I_p, ...I_\alpha\} \tag{1}$$

where, $D$ represents the dataset, $\alpha$ signifies the total input brain images, $I_P$ indicates the image at the $p^{\text{th}}$ index.

### 3.2 Pre-Processing

The input $I_p$ is presented as an input to the pre-processing step, where the image is pre-processed for removing the noises in such a way that the quality of the image is improved. Thus, the pre-processed result is denoted as $I_k$.

### 3.3 Tumor Segmentation

The pre-processed image output $I_k$ is subjected to the tumor segmentation step for partitioning the regions of tumor into different segments. Here, the process is carried out using the U-Net [22]. Finally, the segmented image output is denoted as, $I_s$.

### 3.4 Feature Extraction

The segmented image $I_s$ is fed as an input to the feature extraction step for extracting the features, such as statistical, DWT, and shape features, and the features extracted from the $I_s$ are described below as follows,

**Statistical Features**

*Mean.* The pixels in segmented output are averaged for calculating the mean value, and is formulated as Eq. (2),

$$F_1 = \frac{1}{\left|U\left(W_q\right)\right|} \times \sum_{q=1}^{\left|U\left(W_q\right)\right|} U\left(W_q\right) \tag{2}$$

where, segments are denoted by $q$, $\left|U\left(W_q\right)\right|$ indicate the total pixels in the segmented result, and $F_1$ represent the mean feature.

*Variance.* The variance feature $F_2$ is computed by considering the mean value, and is expressed as Eq. (3),

$$F_2 = \frac{\sum_{q=1}^{|U(W_q)|} |W_q - F_1|}{U(W_q)} \tag{3}$$

*Kurtosis.* Kurtosis indicates the symmetry, and is represented as $F_3$, respectively.

*Standard Deviation.* It is a summary measure to compute the difference between every mean observation, and is represented as $F_4$.

*Energy.* The individual segment energy is computed by summing the pixel energy in the segment, and is represented as $F_5$. As a result, the extracted statistical features are indicated as, $F_\omega = \{F_1, F_2, F_3, F_4, F_5\}$.

**DWT Features.** DWT feature [23] is used to transform the image from spatial to frequency domain. The wavelet transforms extract the knowledge from the image at different scales by passing the brain image through the high pass and low pass filters. The extracted DWT features are represented as $F_\lambda$.

**Shape Features.** The segmented portions of the brain tumor are considered for extracting the shape features, which are in $[4 \times 4]$ grid size, thereby extracting 16 shape features from the grid for the further processing. The shape features are represented as $F_\rho$. Finally, the extracted features are incorporated to form a feature vector, and it is represented as $F$, such that $F$ includes $\{F_\omega, F_\lambda, F_\rho\}$.

## *3.5 Tumor Classification*

The feature vector $F$ is given as an input to tumor classification phase where the process is performed using the proposed hybrid deep learning approach, named Taylor IIWO-based DNN and the HCS-based DBN. The results obtained are incorporated using the Tversky index in order to achieve the efficient classification output.

**Taylor IIWO-Based DNN.** After the process of feature extraction, brain tumor classification is performed using DNN classifier. The feature vector $F$ is subjected to DNN classifier to classify the brain tumor. Moreover, the training procedure of DNN classifier is done by the Taylor IIWO algorithm.

*Structure of DNN.* The DNN classifier [24, 25] is used for classifying the brain tumor using extracted feature vector from the segmented tumor cell. The DNN classifier comprises of Rectified linear Unit (ReLU) layer associated among the input layer and the softmax output layer. Let us consider $b$ is the layer of network, $N$ represents the input given to the $b^{th}$ network, and $M^b$ is the output value at the $b^{th}$ layer. In addition, weight of layer $b$ is represented as $T^b$, and it linearly varies the input value

to output value, $g^b$ denotes the activation vector function, and $j$ signifies the bias. The vector input is expressed as Eq. (4),

$$V^b = N.a^b \tag{4}$$

The ReLU layer is expressed as Eq. (5),

$$U^b = T^b g^b + j^b \tag{5}$$

Moreover, the softmax layer output computed using the DNN classifier is represented as Eq. (6),

$$M^b = g\left(U^b\right) \tag{6}$$

where, the independent samples are denoted as $a$, the element wise vector product is specified as $U$, the $a$ value is re-sampled during the training update process. The final updated equation of the TaylorIIWO-based DNN is expressed as Eq. (7),

$$Y_j^{q+1} = \frac{2}{7 - 2\lambda(q)}\left[2Y_j(q-1) + \frac{Y_j(q-2)}{2}\right](\lambda(q) - 1) + \frac{5Y_{best}}{7 - 2\lambda(q)} \tag{7}$$

$Y_j(q-1)$ indicates the position of the weed at iteration $t = 1$, $Y_j(q-2)$ signifies the position of the weed at iteration $t = 2$, $Y_{best}$ specifies the best weed based on the overall population, and $Y_j^{q+1}$ specifies the new weed position at the $q^{th}$ iteration. The output of Taylor IIWO-based DNN is denoted as $C_1$.

**HCS-Based DBN.** The feature vector $F$ is given as an input for the DBN classifier network [26] for performing brain tumor classification. The DBN [26] is a kind of DNN classifier which includes several layers related to Restricted Boltzmann Machines (RBMs) and Multilayer Perceptrons (MLPs) where the RBMs include visible and hidden units that are mainly associated with weighted connections. In addition, the DBN classifier training is performed by the HCS algorithm. The final update equations of the HCS-based DBN for input and hidden layer are mentioned in Eqs. (8) and (9), respectively.

$$M_{sr}^{I\_MLP}(t+1) = \begin{cases} M_{sr(HCS)}^{I\_MLP}(t+1)\,;\ if\ R_{avg(HCS)} < R_{avg(Gr)} \\ M_{sr(Gr)}^{I\_MLP}(t+1)\,;\ otherwise \end{cases} \tag{8}$$

$$M_{rq}^{H-MLP}(t+1) = \begin{cases} M_{rq(HCS)}^{H-MLP}(t+1)\,;\ if\ R_{avg(HCS)} < R_{avg(Gr)} \\ M_{rq(Gr)}^{H-MLP}(t+1)\,;\ otherwise \end{cases} \tag{9}$$

where, the terms $M_{sr(HCS)}^{I\_MLP}$ and $M_{rq(HCS)}^{H-MLP}$ signifies the updated weights of the input and hidden layer. Thus, the HCS-based DBN output is denoted as $C_2$.

**Fusion using Tversky Index.** The output obtained from the Taylor IIWO-based DNN, and the HCS-based DBN is incorporated together using the Tversky index for classifying the output, and hence the classified output decision is expressed as Eqs. (10), (11), (12),

$$C = \beta C_1 + \gamma C_2 \tag{10}$$

$$\beta = T(C_1, E) \tag{11}$$

$$\gamma = T(C_2, E) \tag{12}$$

where, $\beta$ signifies the Tversky index between $C_1$ and class label,$\gamma$ denotes the Tversky index between $C_2$ and class label, the output of the Taylor IIWO-based DNN is indicated as $C_1$, and the HCS-based DBN output is signified as $C_2$.

## 4 Results and Discussion

The experimental results and discussion of the developed hybrid deep learning approach with respect to the performance evaluation measures is elucidated in this section.

### *4.1 Experimental Setup*

The implementation of the proposed hybrid deep learning technique is performed in MATLAB tool with respect to the BRATS-2015 dataset [27], and the fig share dataset [28].

### *4.2 Comparative Methods*

The performance assessment of developed hybrid deep learning approach is done by comparing with the existing approaches, such as Support Vector Neural Network (SVNN) [29], Harmony Crow Search (HCS)-based Multi SVNN [29], Hybrid Active Contour (HAC) + HCS-based Deep Belief Network (DBN).

### 4.3 Comparative Analysis

This section explains the comparative analysis of the proposed hybrid deep learning technique for dataset-1 and the dataset-2.

**Analysis Using Dataset-1.** Figure 2 illustrates the analysis of the developed method using training data. Figure 2a) illustrates the assessment of accuracy measure. For the training data as 80%, the performance gain measured by the developed hybrid deep learning technique in comparison with existing techniques like SVNN, HCS-based Multi SVNN, and HAC + HCS-based DBN are, 13.921%, 8.554% and 1.212%. The assessment of sensitivity measure is portrayed in Fig. 2b). For the training data 70%, the performance gain achieved by the developed hybrid deep learning approach by comparing with the existing techniques is 23.016%, 16.518% and 7.069%. Figure 2c) shows the assessment of specificity measure. By considering the training data as 60%, the specificity value achieved by the developed hybrid deep learning method is 0.899,

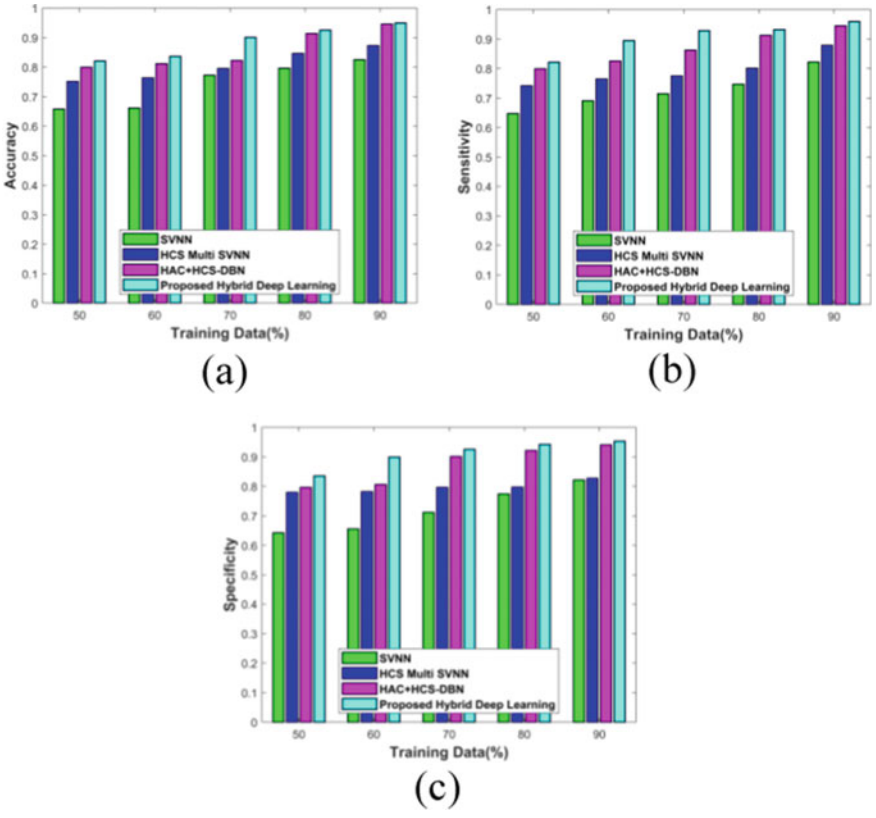



**Fig. 2** Comparative analysis based ondataset-1using training data **a** Accuracy **b** Sensitivity **c** Specificity

whereas the specificity value obtained by the existing SVNN is 0.655, HCS-based Multi SVNN is 0.781, and HAC + HCS-based DBN is 0.806.

**Analysis Using Dataset-2.** The assessment of developed method using training data is portrayed in Fig. 3. The assessment of the accuracy metric by considering the training data percentage is shown in Fig. 3a). For the training data 50%, the accuracy metric measured by SVNN is0.721, HCS-based Multi SVNN is 0.761, HAC + HCS-based DBN is 0.813, and the developed hybrid deep learning approach is 0.896, respectively. The sensitivity analysis is depicted in Fig. 3b). The sensitivity value achieved by the developed hybrid deep learning approach is 0.921, whereas, the sensitivity value computed by the existing SVNN, HCS-based Multi SVNN, and HAC + HCS-based DBN are 0.792, 0.891, 0.909 for the training data 60%. Figure 3c) illustrates the assessment of the specificity metric. For the training data 70%, the performance gain achieved by the developed hybrid deep learning approach by comparing with the existing techniques is 20.831%, 7.554% and 1.533%.

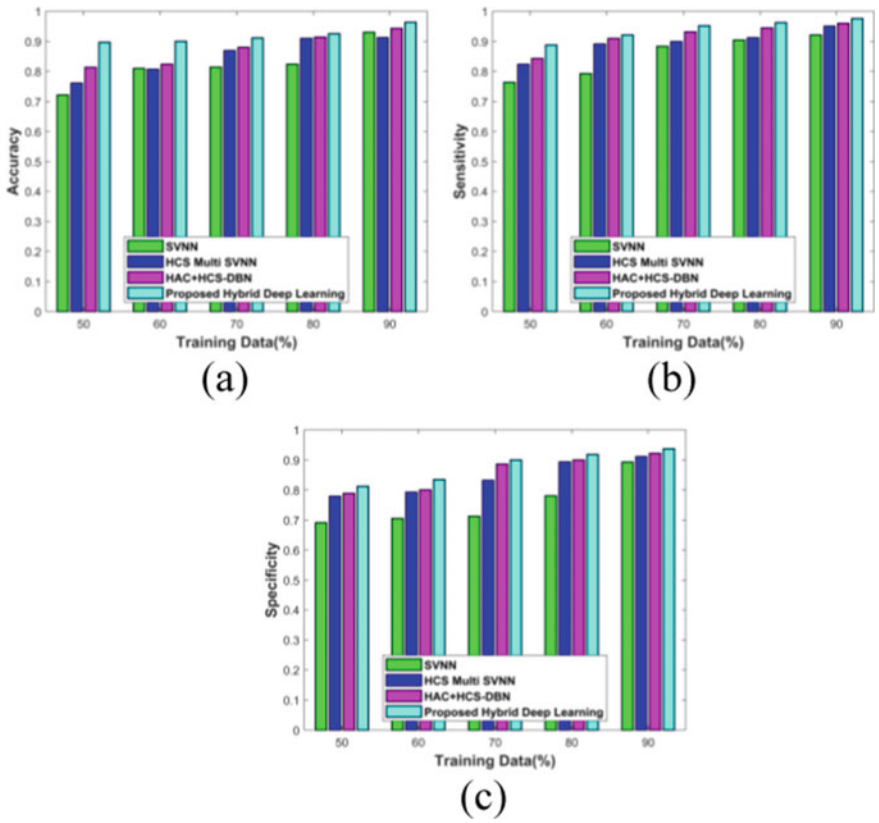



**Fig. 3** Comparative analysis based ondataset-2 by considering the training data percentage **a** Accuracy **b** Sensitivity **c** Specificity

**Table 1** Comparative discussion using training data percentage

| Dataset | Metrics | SVNN | HCS-based Multi SVNN | HAC + HCS-based DBN | Proposed Hybrid deep learning |
|---|---|---|---|---|---|
| BRATS-2015 (dataset-1) | *Accuracy* | 0.825 | 0.873 | 0.946 | **0.949** |
| | *Sensitivity* | 0.822 | 0.878 | 0.944 | **0.959** |
| | *Specificity* | 0.821 | 0.827 | 0.940 | **0.952** |
| Figshare (dataset-2) | *Accuracy* | 0.93 | 0.912 | 0.943 | **0.963** |
| | *Sensitivity* | 0.921 | 0.950 | 0.943 | **0.975** |
| | *Specificity* | 0.892 | 0.911 | 0.960 | **0.937** |

*Comparative Discussion.* The comparative discussion of the developed hybrid deep learning technique based on the best performance is illustrated in Table 1. From Table 1, it is noted that the proposed method outperforms the other methods. The existing brain tumor classification techniques failed to reduce the computational time and also failed to reduce the iteration procedures during the segmentation and classification task. Hence, it is necessary to extend the research by utilizing hybrid and fusion-based methods [30]. As a result, the proposed hybrid deep learning technique is proposed to perform efficient classification of a brain tumor by reducing the time utilization and iterations. The proposed hybrid deep learning approach also achieved accurate detection of the tumor portions due to the hybridization of deep learning techniques.

## 5 Conclusion

This research develops a hybrid deep learning technique for achieving better performance during the brain tumor classification. Here, the pre-processed outcome enhances the image quality by removing the noises. The segmentation phase transforms the images into various segments using the U-Net model. In addition, various features, like statistical, DWT, and shape features are effectively extracted in feature extraction module. Finally, the hybrid deep learning technique is used to classify the brain tumor where the Taylor IIWO based DNN and the HCS-based DBN are integrated based on the Tversky index, thereby effectively classified the output. The developed deep learning approach achieved better performance with the higher accuracy value of 0.963, higher sensitivity of 0.975, and maximal specificity of 0.937 using dataset-2. In the future, the performance efficiency during the brain tumor classification will be enhanced by employing various novel optimization algorithms.

## References


1. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L (2014) The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 34(10):1993–2024
2. Cui S, Mao L, Jiang J, Liu C, Xiong S (2018) Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. J Healthcare Eng 2018:1–15
3. Zhang J, Jiang Z, Dong J, Hou Y, Liu B (2020) Attention Gate ResU-Net for automatic MRI brain tumor segmentation. IEEE Access 8:58533–58545
4. Tong J, Zhao Y, Zhang P, Chen L, Jiang L (2019) MRI brain tumor segmentation based on texture features and kernel sparse coding. Biomed Signal Process Control 47:387–392
5. Olabarriagaa SD, Smeulders AWM (2001) Interaction in the segmentation of medical images: a survey. Med Image Anal 5(2):127–142
6. Pereira S, Pinto A, Alves V, Silva CA (2016) Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imaging 35(5):1240–1251
7. Gumaei A, Hassan MM, Hassan R, Alelaiwi A, Fortino G (2019) A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. IEEE Access 7:36266–36273
8. Sornam M, Kavitha MS, Shalini R (2016) Segmentation and classification of brain tumor using wavelet and Zernike based features on MRI. In: Proceedings of IEEE international conference on advances in computer applications (ICACA), pp 166–169
9. Ilunga-Mbuyamba E, Avina-Cervantes JG, Cepeda-Negrete J, Ibarra-Manzano MA, Chalopin C (2017) Automatic selection of localized region-based active contour models using image content analysis applied to brain tumor segmentation. Comput Biol Med 91:69–79
10. Kaur T, Saini BS, Gupta S (2017) A novel feature selection method for brain tumor MR image classification based on the Fisher criterion and parameter-free Bat optimization. Neural Comput Appl 29:1–14
11. Sauwen N, Acou M, Sima DM, Veraart J, Maes F, Himmelreich U, Achten E, Van Huffel S (2017) Semi-automated brain tumor segmentation on multi-parametric MRI using regularized non-negative matrix factorization. BMC Med Imaging 17(1):17–29
12. Sonavane R, Sonar P (2016) Classification and segmentation of brain tumor using Adaboost classifier. In: Proceedings of international conference on global trends in signal processing, information computing and communication (ICGTSPICC), pp 396–403
13. Sompong C, Wongthanavasu S (2017) An efficient brain tumor segmentation based on cellular automata and improved tumor-cut algorithm. Expert Syst Appl 72:231–244
14. Kumar TS, Rashmi K, Ramadoss S, Sandhya LK, Sangeetha TJ (2017) Brain tumor detection using SVM classifier. In: Proceedings of third international conference on sensing, signal processing and security (ICSSS), pp 318–323
15. Mzoughi H, Njeh I, Wali A, Slima MB, BenHamida A, Mhiri C, Mahfoudhe KB (2020) Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification. J Digit Imaging 33:903–915
16. Díaz-Pernas FJ, Martínez-Zarzuela M, Antón-Rodríguez M, González-Ortega D (2021) A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. In: Healthcare multidisciplinary digital publishing institute, vol 9, no 2, pp 153
17. Krishnakumar S, Manivannan K (2020) Effective segmentation and classification of brain tumor using rough K means algorithm and multi kernel SVM in MR images. J Ambient Intell Humanized Comput 12:1–10
18. Mangai SA, Sankar BR, Alagarsamy K (2014) Taylor series prediction of time series data with error propagated by artificial neural network. Int J Comput Appl 89(1):41–47
19. Misaghi M, Yaghoobi M (2019) Improved invasive weed optimization algorithm (IWO) based on chaos theory for optimal design of PID controller. J Comput Des Eng 6(3):284–295
20. Mahdavi M, Fesanghary M, Damangir E (2007) An improved harmony search algorithm for solving optimization problems. Appl Math Comput 188:1567–1579

21. Askarzade A (2016) A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm. Comput Struct 169:1–12
22. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp 234–241
23. Gupta E, Kushwah RS (2015) Combination of global and local features using DWT with SVM for CBIR. In: Proceedings of 2015 4th international conference on reliability, infocom technologies and optimization (ICRITO), pp 1–6
24. Saxe J, Berlin K (2015) Deep neural network based malware detection using two dimensional binary program features. In: IEEE 10th international conference on malicious and unwanted software (MALWARE), pp 11–20
25. Duan G, Hu W, Wang J (2016) Research on the natural image super-resolution reconstruction algorithm based on compressive perception theory and deep learning model. Neurocomputing 208:117–126
26. Hinton GE, Osindero S, Teh Y (2006) A fast learning algorithm for deep belief nets. Neural Comput 18:1527–1554
27. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). https://www.ncbi.nlm.nih.gov/pubmed/25494501, Accessed Dec 2017
28. Figshare dataset taken from. https://figshare.com/articles/brain_tumor_dataset/1512427, Accessed Mar 2021
29. Raju AR, Suresh P, Rao RR (2018) Bayesian HCS-based multi-SVNN: a classification approach for brain tumor segmentation and classification using Bayesian fuzzy clustering. Biocybern Biomed Eng 38(3):646–660
30. Babu KR, Nadipalli LS, Tejaswini CS, Kumar GB, Vasantha P (2021) CNN fusion based brain tumor detection from mri images using active contour segmentation techniques. J Phys Conf Ser 1804(1):012176

# Identification of Rice Adulteration and Bacterial Blight Using Optimized Boosting Classifier


**J. Friska, A. Rajeshwari, M. Navaneetha Velammal, and P. Hannah Blessy**

**Abstract** Ingredients added or inadvertently added for the purpose of adulteration of food are called as admixture. Food adulteration not only has serious effects for deceive consumers, but also leads to various disorders and diseases. This research study explained about adulteration in rice flour. For that purpose, spectral values are utilized to make rice flour adulteration in various mixers in spectral image form. From that image, Fast Independent Component Analysis (FICA) is used for grouping the components, where the spectrum values are having the same property. Then, the Ada-boosting classifier is used for classification purposes based on whale optimization algorithm. The optimized value address another defects name as bacterial blight. This study clearly finds the adulteration in rice flour with detection of bacterial blight. In the experimental section, the proposed FICA with Ada-boost algorithm obtained 98% of classification accuracy in identification of rice adulteration and bacterial blight. Compared to the existing algorithm, the proposed algorithm showed 0.67% improvement in identification of rice adulteration and bacterial blight.

**Keywords** Ada-boosting classifier · Bacterial blight · Fast independent component analysis · Terahertz spectroscopy · Whale optimization algorithm

## 1 Introduction

Many people who come to the hospital today suffer from food poisoning. It needs to be finding what is poison? What is food? The world is haunting us to mix indistinguishable from that. The Food Safety and Standards Commission released a statistic

J. Friska (✉) · A. Rajeshwari · M. N. Velammal · P. H. Blessy
Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli 627003, India
e-mail: friska.j@francisxavier.ac.in

A. Rajeshwari
e-mail: rajiarjun271997@gmail.com

M. N. Velammal
e-mail: navaneethavelammal.m@francisxavier.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_48

on food adulteration last year. Of that, one-third was advised that the food was adulterated. If there are stones and sand in the mixed food, it will affect the teeth and the thin flesh on the inside of the intestines. The root cause of adulteration is man's greed, to earn money crossroads. It is the desire of the merchants, as plunder is to make a profit, which paves the way for adulteration. Adulteration of food items for extra income also poses a risk to health and life. Consumers are consuming contaminants in the foods every day indescribably high. Where is the food security in this? Talc and lime powder, which carry bacteria that cause disease if they are dirty, affect the digestive system without being digested by us. Impure water can cause many stomach ailments. Our health can be affected if we continue to eat certain mixed foods [4, 5].

Some people add urea, sodium carbonate, sodium hydroxide, formaldehyde, and hydrogen peroxide to prevent milk spoilage, which can cause intestinal itching and damage. Food additives, flavorings, preservatives, antioxidants, etc., can be harmful if they are added in large quantities or are not allowed, adding counterfeit, substandard items starting from 100 g to several hundred grams per kg of food [7]. Millions of consumers are being deceived every day without even knowing it. It's is a correct time to ensure our food security. From the cereals consumed by our ancestors to the pizza and burger that today's generation loves to eat, there are a lot of innovations coming up over time. But in today's environment only the health and quality of food is in question. I do not know which of these to believe and which not to believe. It is impossible for a layman to confirm or reject these. Stone admixture made in rice today is not a problem [8]. It can be split and removed. But the real problem in rice adulteration could be done in the form of rice flour. In this study really helpful for finding where the rice flour is mixed with low quality of rice flour.

## 2 Related Works

Li et al. [1] used principal component analysis to diminish the dimension of the original spectrum information and to extract the feature vectors from the acquired images. Next, Partial Least Squares Discriminant Analysis (PLS-DA) was combined with a few conventional classification algorithms for disease classification. The support vector machine with PLS-DA obtained 97.33% of detection rate. Vlaic et al. [2] developed an algorithm to identify the rice seed cultivars using NIR spectroscopy. In this study, NIR spectral data are used as input, and then the PLS-DA algorithm was used to take out the information for classification purposes. Here, three classifiers are used and it's provide the comparable result one with another. PLS-DA algorithm with KNN classifier gives 80% accuracy. On the other hand, SIMCA with random forest provides 100% accuracy. Anyway it's not suitable for long wavelength spectrum.

Danciu et al. [3] developed an algorithm to identify the transgenic cotton seed based on GA-SVM. In this study, Terahertz spectral data are used as an input. Principle component analysis is used for extracting corresponding features. Here, the SVM classifier is used for classification purposes. By using this method, the rate of

accuracy will be 96.67%. For improving accuracy, PCA is replaced by FICA. Saritha et al. [6] identified the adulteration of papaya seed in black pepper. In this study, the real time images are captured by camera. Then, pre-processing is used to remove unwanted noise from the input image. Some features are extracted for finding accurate classification results. In this study, feature selection was accomplished on the extracted Contrast, texture, shape features. Finally, the KNN classifier was used for classification and the obtained result of adulteration in black pepper showed that the proposed algorithm obtained an accuracy rate of 90%.

## 3 Methods

In this study, four steps could be used for finding rice flour adulteration in ratio and also detection of bacterial blight, which is depicted in Fig. 1.

### *3.1 Spectral Data Collection and Pre-treatment*

In this study, standard spectral values are used some standard criteria. Before capturing the terahertz values, the different rice samples were ground into a powder form. Then various samples of rice flour are mixed with each other. In this study, there are five adulterated samples that are used in different mixed proportions. Adulterated rice flour is in a small round slide under pressure. After arranging the setup, the terahertz spectral value is calculated from that value creating the spectral image [12, 13] for finding rice flour mixture that is graphically depicted in Fig. 2. Spectral pre-treatments are used for remove the unwanted noise from the spectral image.

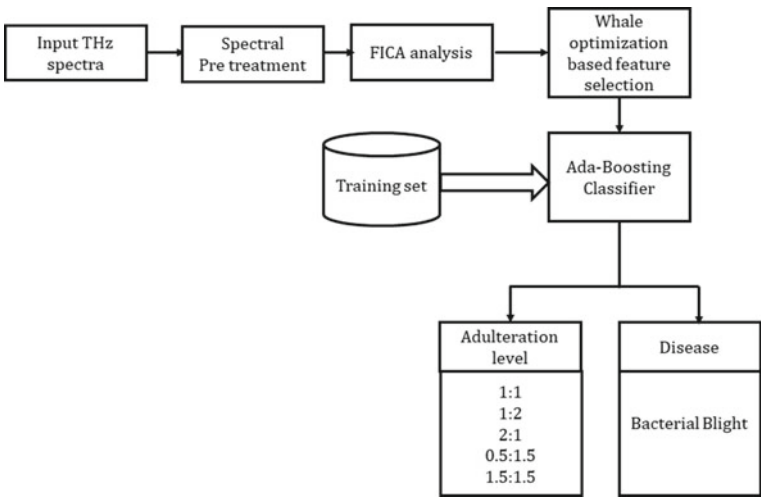
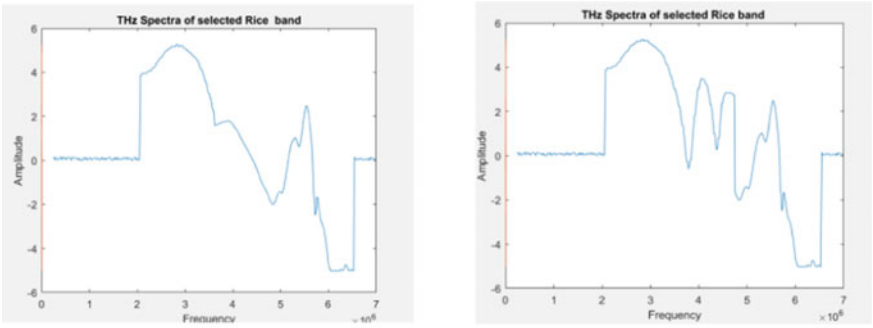



**Fig. 1** Architecture for finding rice flour mixture




a) Sample rice flour mixture in 2:0 b) Sample rice flour mixture in 2:0

**Fig. 2** Various rice flour mixtures in different proportions

**Fig. 3** Whale searching their food

## *3.2 Fast Independent Component Analysis*

In existing system Principle compound analysis used for extracting feature value with clustering based on their requirement. In existing system, PCA used for grouping the matched component. Using PCA doesn't provide the exact principle compound value. So, in this study Fast Independent compound analysis used for extracting feature. It provides the exact grouping component value for each clustering component. And also it's quite faster than PCA [9].

## *3.3 Whale Optimization*

Whale Optimization provides the optimized best value which is helpful for feature selection and optimized based classification. Whale optimization works under the principles of three steps that is graphically depicted in Fig. 3.

Step 1 - Whale searching their food.

Step 2 - Choose their food even top of the surface.
Step 3 - Eat their food. In this optimization algorithm, the best search agent is determined using the Eqs. (1) and (2).

$$\overrightarrow{D} = \left| \overrightarrow{C} \times \overrightarrow{X}(t) - \overrightarrow{X}(t) \right| \tag{1}$$

$$\overrightarrow{X}(t+1) = \overrightarrow{X}(t) - \overrightarrow{A} \times \overrightarrow{D} \tag{2}$$

where, $\overrightarrow{A}$ and $\overrightarrow{C}$ indicates coefficient vectors, $t$ denotes current iteration and $\overrightarrow{X}(t)$ states best position vector. The coefficient vectors $\overrightarrow{A}$ and $\overrightarrow{C}$ are determined using the Eqs. (3) and (4).

$$\overrightarrow{A} = 2\overrightarrow{a} \times \overrightarrow{r} - \overrightarrow{a} \tag{3}$$

$$\overrightarrow{C} = 2\overrightarrow{r} \tag{4}$$

where, $\overrightarrow{a}$ is decreased from 2 to 0 over the iterations and $r$ states random value that ranges between [0, 1].

Basically whales are making their path for eating prey is nine shape path. Whale algorithm, makes the best optimized value for making these kind of various iteration like nine shape path. Whale algorithm provides the best optimized values after much iteration.

### *3.4 Feature Extraction*

The feature extraction step function to discover the various features that represent best adulteration level of mixed rice powder. Here, various feature techniques used such as temporal and spectral.

**Temporal Feature.** In feature extraction, temporal feature is one of the important features. Most of the real time images are taken at the different time. Correlations among the images are used to calculate every change of images in dataset. Various temporal features are there for analyzing images in dataset. Temporal features are used to extract the valuable information such as amplitude, energy based on time domain analysis. This feature calculates the various images in dataset whether it's associate with time or changes over the time. In this study, the temporal features such as mean, Variance, Skewness are considered.

**Spectral Feature.** Spectral feature also very important feature for extracting valuable information from dataset. This feature provides the frequency domain metrics for each image in dataset, this feature, easy to establish whether the data value in

power spectrum value or spectral value. Here some spectral features are spectral centroid, spectral spread, spectral flux, spectral flatness measure and Chroma. In this Study, the Spectral features such as centroid, Spectral Variance considered [10, 11].

### *3.5 ADA-Boosting Classifier*

Ada boosting classifier based on optimized value feature selection which may cause accurate classification results of rice flour mixing in different proportions [14, 15]. In existing system random forest classifier is used for classification purpose. Random forest classifier is a tree like structure which may cause some delay in classifying purpose. Because of each tree iteration takes longer time more than usual. In this study, Ada boosting classifier provides the better result when compared to random forest classifier result.

**Algorithm for Ada-Boosting Classifier.**

1. Let the number of already trained with their feature selection subset A, B, C… etc., and also which is capable of making own classifier error a1, a2, a3…etc.
2. Sort the trained classifier error minimum to maximum
3. Then choose the minimum classifier error with their feature selection
4. Trained the remaining subset and also note down all the classifier error
5. Repeat the step 2–4 for all the subset
6. Now, Let's take the testing subset A1
7. Compare trained subset of minimum classifier error value with the testing subset and provides the accurate result of classifier.

## 4 Result and Discussion

The database for the proposed work is THz spectra of the rice flour mixture. The signals were downloaded from the publically available database. As a Total, 50 different samples were collected from the website and the two sample spectrums were shown in Fig. 4.

The proposed work is focusing on detecting the adulteration as well as disease identification in rice samples. For the case of adulteration detection, mixing of rice flour samples in various proportion such as 3:0, 1:2, 2:1, 0:3, 0.5:1.5, 1.5:0.5. The following Fig. 5 shows the spectra of 3:0 combination of rice flour mixture.

Figure 6 is showing the all combination of rice samples i.e. all ratios are plotted in a single Fig window.

The proposed work is introducing Fast Independent Component Analysis (FICA) for the purpose of identifying the principal component which is varying in the given spectra. The $100^{th}$ to $350^{th}$ spectra is varied, which is depicted in Fig. 7.

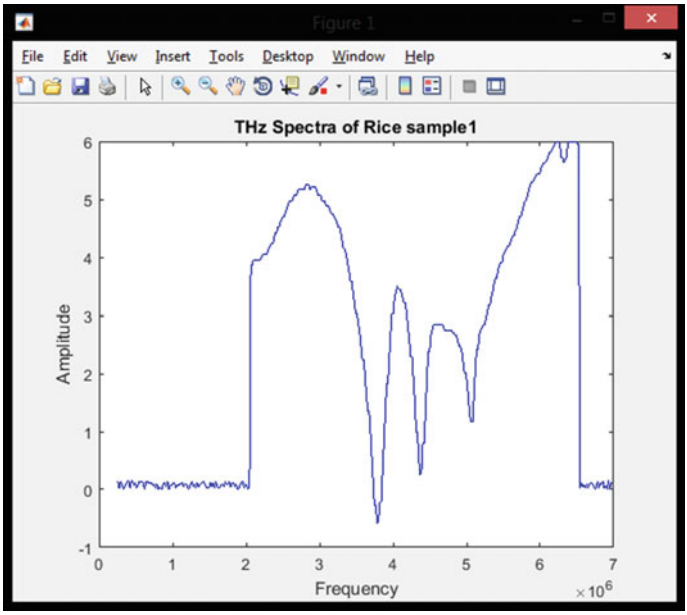



**Fig. 4** THz spectra sample1

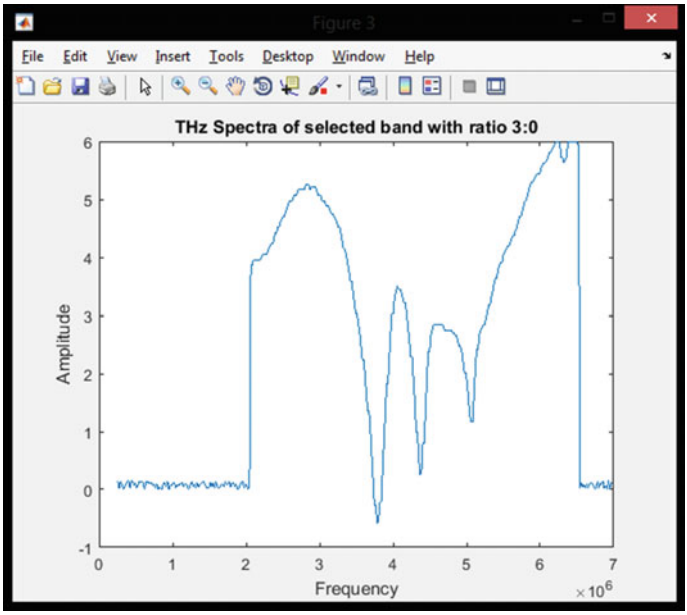



**Fig. 5** THz spectra with the ratio 3:0

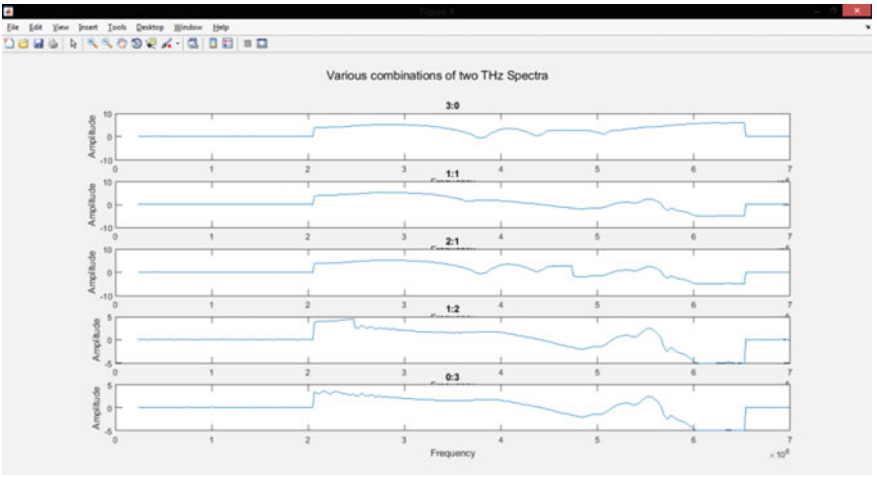



**Fig. 6** Various ratio of THz spectrum

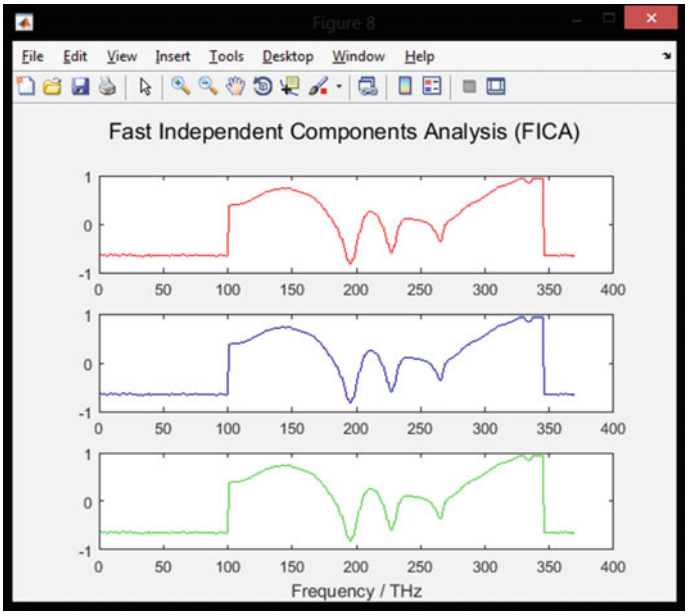



**Fig. 7** FICA response

The whale optimized Ada-boosting classifier can able to classify the exact adulteration level of the rice flour in THz spectra. Figure 8 shows the results obtained for THz spectra sample.

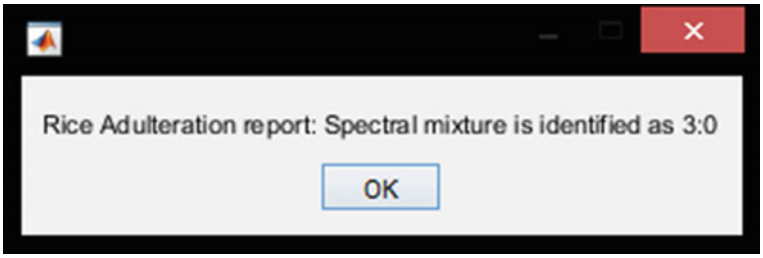



**Fig. 8** THz spectra adulteration level detection

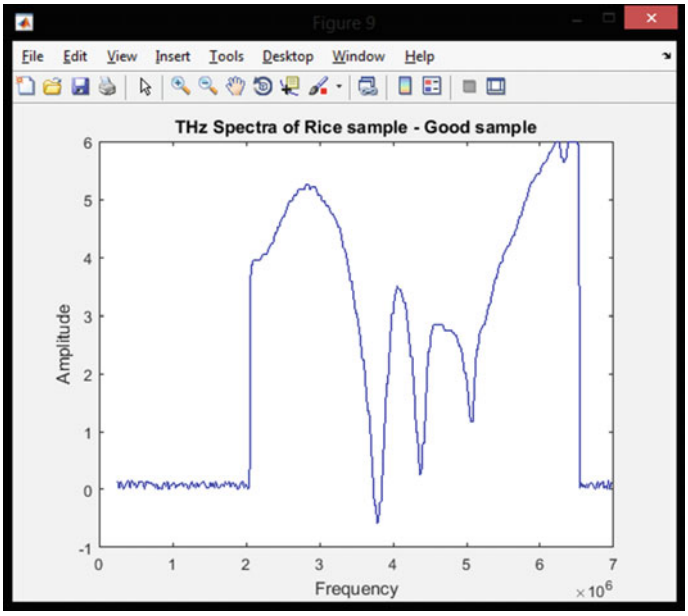



**Fig. 9** THz spectra of good rice sample

The second part of the proposed work is to identify the disease present in the rice sample and that also measured on the same THz spectra sample it, where the committed disease is Bacteria Blight. Figure 9 is showing the sample THz spectra of a good rice sample.

The diseased rice sample Fig. 10 may have undesired internal molecule structure and this can be reflected in the THz spectroscopy images. The Fig shows the affected THz spectra and also provides automated rice disease identification through the THz spectroscopy.

In Table 1 refers the various kind of clustering technique along with classifier which shows their accuracy.

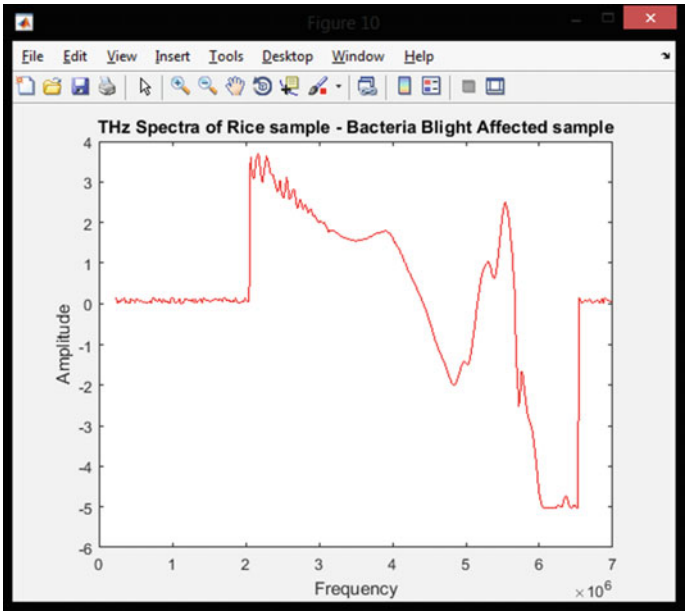



**Fig. 10** THz spectra of affected rice sample

**Table 1** Quantitative analysis with different classifiers

| Clustering technique | Classification | Total number of images | True positive images | Accuracy |
|---|---|---|---|---|
| PCA | Random forest | 30 | 23 | 73 |
| FICA | Random forest | 30 | 29 | 96 |
| Whale optimization | Optimized boosting | 50 | 49 | 98 |

**Table 2** Qualitative analysis with different dataset

| Data to be trained | Data to be tested | Accuracy |
|---|---|---|
| 50 | 50 | 98 |

In this research 50 different kind adulterated rice flour sample to be tested in Table 2 and it gives maximum accuracy. In this research, performance metrics was calculated based on accuracy. Accuracy is calculated based on error rate of test images. The performance of the proposed system was satisfactory in terms of grading of adulteration of rice flour. In this research, FICA technology with Optimized value based Ada boosting classifier is proposed. Here, the proposed algorithm achieved 98% of accuracy rate, which is better compared to the existing algorithm, as shown in Table 3.

**Table 3** Comparative analysis with recent year paper

| Methodology | Accuracy |
|---|---|
| PCA with pattern matching [1] | 97.33 |
| **FICA with Ada-boost algorithm** | **98** |

## 5 Conclusion

The technology used to grade the contaminated rice flour combination in different ratio along with disease detection. In this study, spectral time domain spectrum images were gathered from standard database. The spectral data grouped based on client requirement by using the clustering component FICA. Then, whale optimized based feature selection used for determining the correct optimized value. From that optimized value Ada-boosting classifier classify the different kind of adulterated rice flour in ratio. These highlights are utilized to review the different adulteration levels of rice flour in mixed proportions. Here, various types' rice flour images were trained, and from that 50 spectral images were tested**.** From the tested result based on Whale optimization along with random forest provides the 98% accuracy. In this result, the error rate of finding rice flour mixer provides only one negative value which means 49 true values out of testing 50 spectral images. As a future extension, (i) to speed up the process, the training of convolutional neural network (CNN) could be used and (ii) the adulterated system can be connected to make it available for users to test the adulteration level of rice powder mixture.

## References

1. Li C, Li B, Ye D (2020) Analysis and identification of rice adulteration using terahertz spectroscopy and pattern recognition algorithms. IEEE. https://doi.org/10.1109/2020.2970868
2. Vlaic I, Đurasevic M, Jakobovic D (2019) Improving genetic algorithm performance by population initialization with dispatching rules. Comput Ind Eng 137. Art. no. 106030
3. Danciu M, Alexa-Stratulat T (2019) Terahertz spectroscopy and imaging: a cutting-edge method for diagnosing digestive cancers. MDPI2019 12:1519. https://doi.org/10.3390/ma12091519
4. Anami BS, Malvade NN, Palaiah S (2018) Automated recognition and classification of adulteration levels from bulk paddy grain samples. Inf Process Agric 6(2019):47–60
5. Dhande NS, Sushir RD (2018) Detection and estimation of adulteration in oil sample using digital image processing IJSDR 4(2):ISSN: 2395-6011
6. Saritha B, Gokila G, Gokul priya A, Mahalakshmi K (2017) Adulteration identification of papaya seeds in black pepper using digital image processing. IJIRCCE 20(47):ISSN(Online): 2320-9801, ISSN (Print): 2320-9798
7. Pratibha N, Hemlata M, Krunali M, Khot ST (2017) Analysis and identification of rice granules using image processing and neural network. Int J Electron Commun Eng 10(1):0 25–33. ISSN 0974-2166
8. Deng W, Zhao H, Zou L, Li G, Yang X, Wu D (2017) A novel collaborative optimization algorithm in solving complex optimization problems. Soft Comput 21(15):43874398
9. Liu W, Zhang Y, Han D (2016) Feasibility study of determination of high-fructose syrup content of Acacia honey by terahertz technique. ProcSPIE. https://doi.org/10.1117/12.2245966

10. Herath K (2016) Rice grains classification using image processing techniques. researchgate. https://www.researchgate.net/publication/311373270
11. Liu W, Liu C, Hu X, Yang J, Zheng L (2016) Application of terahertz spectroscopy imaging for discrimination of transgenic Rice seeds with chemo metrics. Food Chem 210:415421
12. Liu W et al (2016) Discrimination of transgenic soybean seeds by terahertz spectroscopy. Sci Rep. 6:35799. https://doi.org/10.1038/srep35799.
13. Redo-Sanchez A, Laman N, Schulkin B, Tongue T (2013) Review of terahertz technology readiness assessment and applications. J Infrared Millimeter Terahertz Waves 34:500–518
14. Feng X, Zhang Q, Cong P (2013) Preliminary study on classification of Rice and detection of paraffin in the adulterated samples by Raman spectroscopy combined with multivariate analysis. Talent
15. Gowenab A, O'Sullivanc C, O'Donnell CP (2012) Terahertz time domain spectroscopy and imaging: emerging techniques for food process monitoring and quality control. Trends Food Sci Technol 25(1):40–46

# Image Classification Based on Inception-v3 and a Mixture of Handcrafted Features


**A. Shubha Rao and K. Mahantesh**

**Abstract** Annotating, retrieving and classifying images in an ever increasing large image datasets with semantic information is still an exigent task. Image pre-processing plays a vital role in partitioning an image into small meaningful regions and extracting features from these regions can be used to stimulate the learning models for better and accurate image classification. Deeplabv3+ framework based on dilated convolution is used to separate out the semantic regions and Histogram of Oriented Gradients (HOG) technique is applied to these regions and computed the distribution of gradients as features. A comparative analysis and examination of various Machine Learning classifiers using edge based, semantically segmented features are compared against the state-of-the-art deep learning techniques. The proposed deep learning based Inception-v3 architecture enables dynamic object recognition with factorized convolution and parameter reduction. The proposed model outperforms hand crafted ML classifiers, shows a significant performance improvement on much diverse dataset like Caltech-256 in comparison with Caltech 101.

**Keywords** Caltech-101 · Caltech-256 · Histogram of Oriented Gradients (HOG) · Image classification · Machine learning · Semantic segmentation

## 1 Introduction

Over the last century, there has been a major development in technology, which has changed our perspective on computer and hence there comes the entire buzz on Artificial Intelligence (AI) [1]. Machine learning and AI are so intertwined altogether—to put some sense our computer. Where instead of hard coding the algorithms, given a huge diverse set of data we expect computer to figure out the hidden logic by itself. Machine Learning was initially used to recognize pattern in our data, and to perform some task based on the gained knowledge [2]. Machine learning has wide range of

A. Shubha Rao (✉) · K. Mahantesh
Department of ECE, SJB Institute of Technology, Bangalore, India
e-mail: mail2shugar@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_49

application ranging from fraud detection to medical diagnosis to feeling analysis on twitter [3].

Machine learning can be broadly classified into four categories. Supervised learning works on labeled data. Maps the given input to output based on the learned knowledge from earlier experience using mathematical function [4]. It can be used to solve both classification and regression problems. Ex: Support Vector Machines, Random Forrest. Unsupervised learning works on unlabeled data. The function which maps the data is more complex, the output is less accurate in comparison to supervised methods [5]. Ex: K-means clustering, Principal Component Analysis. Semi-supervised learning is mainly used when the incurred cost of labeling our data is high [6]. Algorithm works with small amount of labeled data and larger unlabeled data. Ex: Generative models, Heuristic approaches. Reinforcement learning is majorly used in gaming, robotics and navigation kind of applications [7]. It motivates the working agent (model) to take decision which results in maximum profit (reward) in a given environment. Ex: Value Based, Policy Based, Model Based.

## 2 Related Works

A paper on disentanglement representation of the data in unsupervised learning, shows that without inducing bias disentangled data, unsupervised learning seems not useful. It has also been observed on different datasets that increased disentanglement does not directly imply reduced complexity of data [8]. Another approach where instead of using a dense network which will be pruned later, a "Lottery Ticket" hypothesis which says there always exists a smaller subnetwork efficient enough to achieve the same accuracy when initialized properly and with same number of iterations [9]. Instead of utilizing the representations which are produced as a side effect of generative models, a semi supervised approach which aims to produce data representations directly—meta learning, which later can be used for various subsequent tasks [10]. A supervised learning approach, ANN with feature selection and wrapper function is seen to outperform the support vector machine technique for the classifying the network traffic as malicious or benign [11]. A variance of stochastic gradient descent—RAdam which aims to reduce the initial high variance in learning rate, during the early stages of training. The method shows empirical improvement in generalization, faster convergence and soothed training of data [12].

HoG based feature extraction method which was used to detect stomach cancer from images whether it is normal, benign and malign using a specific bandwidth range was observed to improve accuracy [13]. An effective method to reduce the dimensionality, while without any loss in the accuracy is proposed for detection of pedestrian. Firstly, based on intervals of gradient orientation various HOG channels are selected independently, which are later combined depending on the statistics uniformity (CHOG-C) [14]. Human detection with increased throughput is achieved using architecture, which is a combination of HoG and SVM. The architecture achieves parallel computation using SVM along with normalized features of HOG

[15]. Another approach for fast human detection by reducing the number of features to be extracted, using a cascaded HoG-LBP classifier. SVM trained on AdaBoost is used to differentiate between the pedestrian and non-pedestrian, which are later passed on to HoG and LBP for feature extraction and recognition [16]. For the classification of glaucoma images, an extreme learning machine is proposed which preprocess the images using CLAHE contrast, later extracts the features based on HOG and wavelets [17].

Since traditional segmentation methods for object extraction are based on low level features, a new object extraction method based on semantically relevant data is proposed. The model IOEBSS, preprocess the image by using fast binary plane filtering, followed by deeplab architecture-based segmentation along with label loss, proves to be efficient [18]. An improvement in the direction of weekly supervised algorithms for segmentation with single annotation per category is proposed. The method performs segmentation based on three different perspectives—image, pixel and object while extracting common features from coarse to fine for object regions which are more accurate [19]. A novel approach for cell nuclei segmentation, which is based on U-Net inspired by encoder decoder model, the architecture adds atrous diluted convolution to the U-Net architecture along with log-dice loss, Adam optimization during training for efficient segmentation [20]. To retain the semantic relevance information of object in practical applications of style transfer, a mask R-CNN based semantic segmentation which applies segmentation based on patch level which helps in style transfer among semantically relevant objects [21]. Method to reduce the inherent compromise that exists between accuracy and inference time in segmentation algorithm, a fast spatial feature network (FSFNet) is proposed. The architecture effectively extracts the features at different resolution, which are later used to reconstruct the segmented image accurately [22].

## 3 Methods

### *3.1 Histogram of Oriented Gradients (HOG)*

Histogram of Oriented Gradients (HOG) is a method of finding distribution of gradients along with its orientation [23]. In HOG, any given image is preprocessed and reduced to a size of (64,128). Gradients are intensity change along x and y direction. Corner is when there is very large magnitude of change along both the axis. Once the gradients are found its orientation can be easily found by Pythagoras theorem. Next step, a histogram of the oriented gradients is computed and normalized to reduce the variation in intensity (please refer to Fig. 1.).

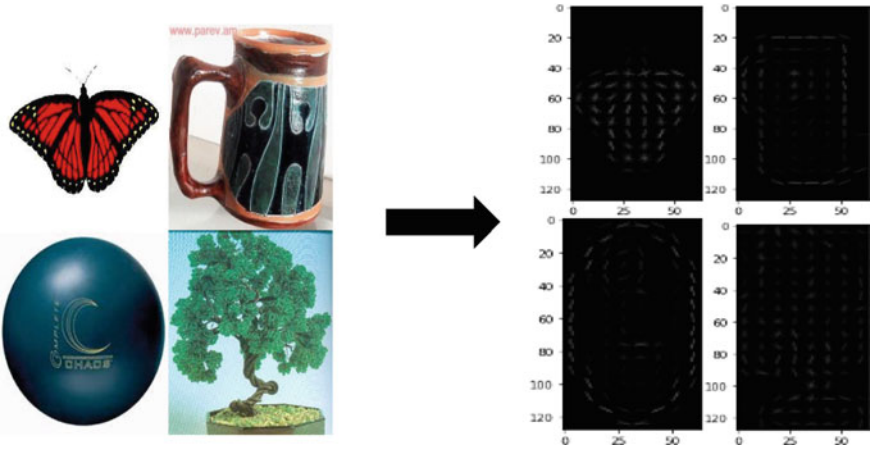


**Fig. 1** Example of applying HOG on Caltech-256, Class—butterfly, bearmug, bowling ball, bonsai

## 3.2 *Semantic Segmentation: Deeplabv3+ with Pre-trained Weights*

Semantic Segmentation is a process where each and every pixel of an image is classified and assigned a label. Deeplabv3+ framework is used to semantically segment the images inspired by encoder—decoder model. The architecture is mainly based on the Atrous depth-wise convolution (Dilated convolution) to increase the area of view which is particularly helpful in locating objects. Above which 1 * 1 convolution is being used to reduce the computation cost. Decoder reduces the channel of low-level feature by 1 * 1, so doesn't outweigh the extracted features from encoder. Output from the encoder is up sampled before merging with the low-level features before making the final pixel based prediction [24] (please refer to Fig. 2).

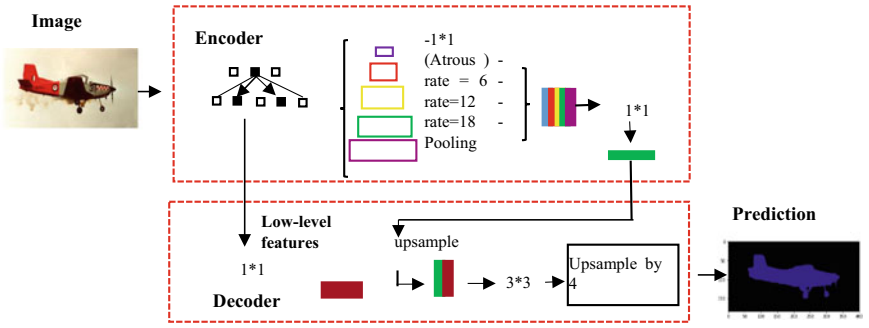



**Fig. 2** Deeplabv3+ architecture

### *3.3 Proposed Method: Inception Model with Pre-trained Weights*

Image classification performance on Caltech-101 and Caltech-256 image dataset using various machine learning classifiers is analyzed. Two different feature extraction methods are applied and results are compared. Firstly, Histogram Oriented Gradients (HOG) is used to extract the edges of an image, various classifiers are applied on HOG images. Secondly, images are segmented using Deeplabv3+, semantically segmented images are fed into algorithms. To further analyze the various classifiers HOG is applied over segmented images.

**Algorithm—Ensemble Model for ML Classifiers**

1. Load the image.
2. Preprocess the image for feature extraction.
3. Apply semantic segmentation on the images, passing the semantically segmented images for edge based feature extraction (HOG).
4. Save the extracted feature vectors.
5. Pass the features of train images to the various ML classifiers.
6. Test the algorithm on test images.

One of the major drawbacks of machine learning algorithm is static feature engineering. Since the features are manually extracted and filters are hard coded. With the growth in Artificial Intelligence, the expectation of machine to learn automatically without any human intervention has been made possible by Deep Learning. On the contrary with ML algorithms, Deep Learning adopts dynamic feature engineering with non-linear transformation, the optimal parameters (filters) for any given task, is learnt by the machine itself [25]. The effectiveness of Deep Learning algorithms in comparison to traditional Machine Learning Algorithm is clearly visible. Since deep learning algorithms take longer time to train, transfer learning is most optimal solution [26, 27].

Deep learning involves stacking the model with layers to increase the performance. Instead, an idea to make the model broader with varying size of filters spread across enables to recognize varying size of objects in image. With this approach model facilitates easy update of gradients over the layers tackling the overfitting problem. Asymmetric convolution, auxiliary classifiers and grid size dimension reduction being some of the key features of Inception model. Transfer learning allows quick learning of features by initializing the parameter with pre-trained weights on Imagenet dataset.

**Algorithm—Inception Model:**

1. Load the data, split and create train and test data frames.
2. Initialize the Inception model with pre-trained weights.
3. Freeze all the initial layers and train only the top layers.

4. Pass all the images from data frames to the model, apply preprocess.
   With data augmentation—rotation_range = 30, zoom_range = 0.3, width_shift_range = 0.2, height_shift_range = 0.2, horizontal_flip = 'true'
5. Compile the model and fit the data. (Epoch = 40)
   Adagrad Optimizer: accumulated gradients for weight update is expressed in Eqs. 1 and 2.

$$r_t = r_{t-1} + (\Delta\theta_t)^2 \tag{1}$$

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\delta + \sqrt{r_t}}\Delta\theta_t \tag{2}$$

   Cross Entropy Loss function is expressed in Eq. 3:

$$L_{CE}(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N} y_i log\hat{y}_i \tag{3}$$

6. Evaluate the model on test images.

## 4 Results

A comparison of various machine learning classifiers on Caltech-101 and Caltech-256 Image Dataset on raw image pixel data, HoG features, semantically segmented images and HoG over semantically segmented features is given in Table 1. Support Vector Machines though efficient it takes longer run time. Other hand, Decision Trees and Random Forest is the fastest with mere human intervention. It can be clearly seen from the results the HOG extracted features, and with semantically segmented the algorithms improves in performance. From the results it can be implied, edge based feature extraction methods shows a significant improvement in classification. It can be inferred that HOG applied over segmented images, does not show a consistent improvement. Although, segmented images show significant improvement in classification accuracy in Caltech-101 image dataset, fails to show improvement as the number of classes are increases (Caltech-256 image category). Hence, properly extracted features play a major role in classification. The performance of the algorithm is measured in terms of accuracy (%).

Table 2 shows the comparison of performance of various machine learning algorithms with Deep learning based Inception model. It can be clearly seen proposed dynamic feature engineering based inception model outperforms the hand crafted machine learning classifiers. Performance analysis of the earlier work done on Caltech-101, Caltecl-256 is given in the Tables 3 and 4 respectively. In comparison with various methods like shape matching, pyramid matching kernel, discriminative nearest neighbor, the deep learning model clearly outperforms. Deep learning enables self-learning, model ultimately learns the filters which extracts the best features from the image that leads to better performance

**Table 1** Performance of different classifiers on HoG, semantic segmented + HoG features

| | Metric accuracy (%) | Caltech -101 30-Train | Caltech-256 30-Train |
|---|---|---|---|
| Logistic regression | HoG | 0.61 | 0.20 |
| | Semantic segmented + HoG | 0.43 | 0.15 |
| Stochastic gradient descent | HoG | 0.55 | 0.16 |
| | Semantically segmented + HoG | 0.37 | 0.10 |
| Support vector machine | HoG | 0.41 | 0.17 |
| | Semantic segmented + HoG | 0.40 | 0.13 |
| K—Neighbors classifier | HoG | 0.41 | 0.13 |
| | Semantic segmented + HoG | 0.34 | 0.10 |
| Decision tree classifier | HoG | 0.20 | 0.08 |
| | Semantic segmented + HoG | 0.40 | 0.04 |
| Random forest | HoG | 0.51 | 0.14 |
| | Semantic segmented + HoG | 0.46 | 0.16 |
| Multi-layer perceptron | HoG | 0.02 | 0.13 |
| | Semantic segmented + HoG | 0.36 | 0.10 |

**Table 2** Qualitative analysis of ML classifiers with deep learning technique

| Methods | Caltech 101(%) 30—Train | Caltech 256 (%) 30—Train |
|---|---|---|
| Image pixels with logistic regression | 0.42 | 0.12 |
| HoG features with logistic regression | 0.61 | 0.20 |
| Semantic segmentation with random forest | 0.44 | 0.15 |
| Semantic segmentation + HoG with random forest | 0.46 | 0.13 |
| **Proposed model-Inceptionv3 with Top** | **0.67** | **0.59** |

**Table 3** Comparative analysis of Caltech-101 result with previous work

| Methods | Caltech101(%) 15—Train | Caltech101(%) 30—Train |
|---|---|---|
| Shape matching [28] | 45 | – |
| Pyramid match kernels [29] | 49.5 | 58.2 |
| Discriminative nearest neighbour [30] | 59 | 66 |
| Local naïve bayes nearest neighbour [31] | 47.8 | 55.2 |
| Sparse localized features [32] | 33 | 41 |
| Relevance based classification [33] | – | 43.8 |
| Gaussian mixture models [34] | – | 72.3 |
| VGG-16 [25] | 66 | 78.42 |
| **Proposed model—Inceptionv3 with Top** | **64** | **67** |

**Table 4** Comparative analysis of Caltech-256 result with previous work

| Methods | Caltech 256(%) 15—Train | Caltech 256(%) 30—Train |
|---|---|---|
| Learning dictionary [35] | 30.35 | 36.22 |
| Sparse spatial coding [36] | 30.59 | 37.08 |
| Discriminative coding for object classification [37] | 28 | 30 |
| Local naïve bayes classifier [31] | 33.5 | 40.1 |
| Caltech Institute classification [38] | 28.3 | 34.1 |
| Combined image descriptors [39] | – | 33.6 |
| VGG-16 with Top [25] | 51 | 57.57 |
| **Proposed model—Inceptionv3 with Top** | **58** | **59** |

## 5 Discussion and Conclusion

Though performance of machine learning algorithms can be improved with HoG extracted features. HOG contains only edge related information of an image. Semantically segmentation classifies each pixel of an images, change its color so that it can be easily identified, and turns the background to black. In the procedure the algorithm losses all the finest details of an object—color, texture, edge, corner, intensity variation, which are essential feature for any classification algorithm identity and correlate. Although further improvement can be witnessed with semantically segmented images, deep learning clearly outperforms old-style machine learning techniques.

Proposed transfer learning based Inceptionv3 model improves the efficiency with much diverse dataset Caltech-256 in comparison to Caltech-101. Also it can be clearly inferred the proposed model outperforms even with much smaller sample data (15-train). The choice of whether to go for deeper model (vgg16) or wider model (inceptionv3) depends greatly on your choice of dataset.

Furthermore, an efficient segmentation method which retains all the necessary unique feature of an object is most evident. Although Convolutional Neural Network (CNNs) outperform, it comes with its own set of drawbacks—tuning of parameters (hyper parameter) is crucial, requires huge amount of training data, memory requirement, adequate training time from scratch. The problem of image classification with diverse dataset like Caltech-256 with 256 object categories is still a challenge.

## References


1. Mahantesh K, Shubha Rao A (2019) Content based image retrieval - inspired by computer vision & deep learning techniques. In: 2019 4th international conference on electrical, electronics, communication, computer technologies and optimization techniques (ICEECCOT), pp 371–377. https://doi.org/10.1109/ICEECCOT46775.2019.9114610
2. Ray S (2019) A quick review of machine learning algorithms. In: 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon), pp 35–39. https://doi.org/10.1109/COMITCon.2019.8862451
3. Obulesu O, Mahendra M, ThrilokReddy M (2018) Machine learning techniques and tools: a survey. In: 2018 international conference on inventive research in computing applications (ICIRCA), pp 605–611. https://doi.org/10.1109/ICIRCA.2018.8597302
4. Shrestha A, Mahmood A (2019) Review of deep learning algorithms and architectures. IEEE Access 7:53040–53065. https://doi.org/10.1109/ACCESS.2019.2912200
5. Chowdhury S, Schoen MP (2020) Research paper classification using supervised machine learning techniques. In: 2020 intermountain engineering, technology and computing (IETC), pp 1–6. https://doi.org/10.1109/IETC47856.2020.9249211
6. Tamilselvi P, Kumar KA (2017) Unsupervised machine learning for clustering the infected leaves based on the leaf-colours. In: 2017 third international conference on science technology engineering & management (ICONSTEM), pp 106–110. https://doi.org/10.1109/ICONSTEM.2017.8261265
7. Le Nguyen MH, Gomes HM, Bifet A (2019) Semi-supervised learning over streaming data using MOA. In: 2019 IEEE international conference on big data (Big Data), pp 553–562. https://doi.org/10.1109/BigData47090.2019.9006217
8. Locatello F, Bauer S, Lucic M, Raetsch G, Gelly S, Schölkopf B, Bachem O (2019) Challenging common assumptions in the unsupervised learning of disentangled representations. In: Proceedings of the 36th international conference on machine learning. PMLR, vol 97, pp 4114–4124
9. Frankle J, Carbin M (2019) The lottery ticket hypothesis: finding sparse, trainable neural networks. In: International conference on learning representations
10. Metz L, Maheswaranathan N, Cheung B, Sohl-Dickstein J (2018) Meta-learning update rules for unsupervised representation learning. CoRR
11. Taher KA, Mohammed Yasin Jisan B, Rahman MM (2019) Network intrusion detection using supervised machine learning technique with feature selection. In: 2019 international conference on robotics, electrical and signal processing techniques (ICREST), pp 643–646. https://doi.org/10.1109/ICREST.2019.8644161
12. Liu L, Jiang H, He P, Chen W, Liu X, Gao J, Han J (2019) On the variance of the adaptive learning rate and beyond. CoRR
13. Korkmaz SA, Akçiçek A, Bínol H, Korkmaz MF (2017) Recognition of the stomach cancer images with probabilistic HOG feature vector histograms by using HOG features. In: 2017 IEEE 15th international symposium on intelligent systems and informatics (SISY), pp 000339–000342. https://doi.org/10.1109/SISY.2017.8080578
14. Weixing L, Haijun S, Feng P, Qi G, Bin Q (2015) A fast pedestrian detection via modified HOG feature. In: 2015 34th Chinese control conference (CCC), pp 3870–3873. https://doi.org/10.1109/ChiCC.2015.7260236
15. Nguyen N, Bui D, Tran X (2019) A novel hardware architecture for human detection using HOG-SVM co-optimization. In: 2019 IEEE Asia Pacific conference on circuits and systems (APCCAS), pp 33–36. https://doi.org/10.1109/APCCAS47518.2019.8953123
16. Park W, Kim D, Suryanto CL, Roh TM, Ko S (2012) Fast human detection using selective block-based HOG-LBP. In: 2012 19th IEEE international conference on image processing, pp 601–604. https://doi.org/10.1109/ICIP.2012.6466931

17. Shyla NSJ, Emmanuel WRS (2021) Automated classification of glaucoma using DWT and HOG features with extreme learning machine. In: Third international conference on intelligent communication technologies and virtual mobile networks (ICICV), pp 725–730. https://doi.org/10.1109/ICICV50876.2021.9388376
18. Wang X, Xu P, Yu Z, Li F (2020) Image object extraction based on semantic segmentation and label loss. IEEE Access 8:109325–109334. https://doi.org/10.1109/ACCESS.2020.2999942
19. Li X, Ma H, Luo X (2020) Weaklier supervised semantic segmentation with only one image level annotation per category. IEEE Trans Image Process 29:128–141. https://doi.org/10.1109/TIP.2019.2930874
20. Pan X, Li L, Yang D, He Y, Liu Z, Yang H (2019) An accurate nuclei segmentation algorithm in pathological image based on deep semantic network. IEEE Access 7:110674–110686. https://doi.org/10.1109/ACCESS.2019.2934486
21. Xie C et al (2021) Image style transfer algorithm based on semantic segmentation. IEEE Access 9:54518–54529. https://doi.org/10.1109/ACCESS.2021.3054969
22. Kim M, Park B, Chi S (2020) Accelerator-aware fast spatial feature network for real-time semantic segmentation. IEEE Access 8:226524–226537. https://doi.org/10.1109/ACCESS.2020.3045147
23. Zhang L, Zhou W, Li J, Li J, Lou X (2020) Histogram of oriented gradients feature extraction without normalization. In: IEEE Asia Pacific conference on circuits and systems (APCCAS), pp 252–255. https://doi.org/10.1109/APCCAS50809.2020.9301715
24. Libiao J, Wenchao Z, Changyu L, Zheng W (2021) Semantic segmentation based on DeeplabV3+ with multiple fusions of low-level features. In: IEEE 5th advanced information technology. Electronic and automation control conference (IAEAC), pp 1957–1963. https://doi.org/10.1109/IAEAC50856.2021.9390753
25. Rao AS, Mahantesh K (2021) Learning semantic features for classifying very large image datasets using convolution neural network. SN Comput. Sci. 2:187. https://doi.org/10.1007/s42979-021-00589-6
26. Bamne B, Shrivastava N, Parashar L, Singh U (2020) Transfer learning-based object detection by using convolutional neural networks. In: International conference on electronics and sustainable communication systems (ICESC), pp 328–332. https://doi.org/10.1109/ICESC48915.2020.9156060
27. Wu X, Liu R, Yang H, Chen Z (2020) An xception based convolutional neural network for scene image classification with transfer learning. In: 2020 2nd international conference on information technology and computer application (ITCA), pp 262–267. https://doi.org/10.1109/ITCA52113.2020.00063.
28. Berg TL, Berg AC, Malik J (2005) Shape matching and object recognition using low distortion correspondence. In: IEEE CVPR, vol 1, pp 26–33
29. Grauman K, Darell (2006) Pyramid match kernels: discriminative classification with sets of image features. Technical report MIT-CSAIL-TR-2006-020
30. Maire M, Malik J, Zhang H, Berg AC (2006) SVM-KNN: discriminative nearest neighbor classification for visual category recognition. In: IEEE-CVPR, vol 2, pp 2126–2136
31. Mc Cann S, Lowe DG (2012) Local naive bayes nearest neighbor for image classification. In: IEEE-CVPR, pp 3650–3656
32. Mutch J, Lowe DG (2006) Multiclass object recognition with sparse, localized features. In: IEEE CVPR, vol 1, pp 11–18
33. González G, Türetken E, Benmansour F, Rigamonti R, Lepetit V (2014) On the relevance of sparsity for image classification. Comput Vis Image Underst 125:115127
34. Mahantesh K, Aradhya VNM, Niranjan SK (2014) An impact of complex hybrid color space in image segmentation. In: Recent advances in intelligent informatics. Advances in intelligent systems and computing, vol 235. Springer, pp 73–83. https://doi.org/10.1007/978-3-319-01778-58
35. Liu Y-J-D, Wang Y-X (2012) Learning dictionary on manifolds for image classification. Pattern Recogn 46:1879–1890

36. Vieira AW, Campos MF, Oliveira GL, Nascimento ER (2012) Sparse spatial coding: a novel approach for efficient and accurate object recognition. In: IEEE international conference on robotics and automation (ICRA), pp 2592–2598
37. Zhang Y, Zheng Y, Liu B, Wang Y (2012) Discriminant sparse coding for image classification. In: Proceedings of the 37th international conference on acoustics, speech and signal processing, pp 2193–2196
38. Holub A, Griffin G, Perona P (2007) Caltech 256 object category dataset. Technical report, California Institute of Technology
39. Banerji CLS, Sinha A (2013) New image descriptors based on color, texture, shape, and wavelets for object and scene image classification. Neurocomputing 117:173–185

# Image Process Based Plant Diagnostic System


**Naga Raju Jangam, Archish Amar Ringangonkar, Battula Mohan Kumar, Linga Vishal, and Kalal Hanush Goud**

**Abstract** Plant disease diagnosis is critical for reducing crop losses, and increasing agricultural growth. This research paper discussed methods used to diagnose plant diseases using leaf pictures and discussed specific classifications and the algorithms to extract a factor used in the diagnosis of plant disease. Then, plant specific diseases plays an essential role in the agricultural sectors, because plant-borne diseases are quite nature. If proper care is not taken in plant disease, it causes adverse effects on the plants and consequently affects the quality of product, quantity or product. Plant diseases cause's outbreaks that results in production loss, and this problem need to be addressed in the first phase, saving lives and money. Detection of diseases in plants is critical research area, which shows benefits in monitoring large plant fields. Farm owners, and caregivers of plants (say, kindergarten) can benefits greatly from early detection of diseases, to prevent the bad from reaching their plants and to inform the person what needs to be done in advance for the similar to work properly, to prevent the worse.

**Keywords** Classification · Feature extraction process · Image processing application · Plant disease · Symptom

## 1 Introduction

India is a cultivated country, with agriculture employing over 70% of the people directly or indirectly. Farmers have a wide variety of options to choose from and find the right pesticides [1, 2]. Therefore, crop damage will lead to significant losses in productivity and ultimately affect the economy. The leaves are a very sensitive part of plants that show signs of disease at an early age. Plants must be checked for disease from the start of their life cycle until they are ready to be harvested [3–5]; formerly, disease monitoring was done in a different method, with traditional naked eye tracking being a time-consuming process that required experts to watch crop

N. R. Jangam (✉) · A. A. Ringangonkar · B. M. Kumar · L. Vishal · K. H. Goud
Department of Electronics and Communication, MLR Institute of Technology, Hyderabad, India
e-mail: nagaraju.jangam@mlrinstitutions.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_50

fields [6–8]. In recent years, a number of techniques have been developed to improve the diagnosis of spontaneous and autoimmune diseases by simply recognizing the symptoms on the leaves of plants and making them easier and cheaper [9, 10]. These programs have so far led to them being faster, less expensive and more accurate than the traditional way of looking at farmers [11].

This paper is organized into the following sections. The first section provides a brief introduction to the importance of diagnosing plant diseases. The second section discusses the work that has been done recently in this area and reviews the methods used. The third part lists the methods used in this research paper. Lastly, section four concludes this paper with future indicators. The identification of plant disease is most important aspect in modern agriculture system. Determining the health of a plant or a particular leaf in this plant is very important and plays a big role because a bad plant can damage the plants and plants around it and as a result, it can give a bad harvest.

## 2 Related Works

Kaur and Devendran [13] introduced a new optimization based segmentation and law mask system for leaf disease detection and classification. In this literature study, Support Vector Machine (SVM) classifier was applied for classifying the plant leaves. In addition, Verma and Dubey [14] developed a novel model; Long Short Term Memory (LSTM) and Simple Recurrent Neural Network (SRNN) for detecting the disinfected or disease plants with dynamic learning capability. Peker [15] has introduced new deep learning model on the basis of ensemble learning and capsule networks for plant disease detection. Further Vasumathi and Kamarasan [16] implemented convolutional neural network based LSTM model for classifying 6519 fruits into abnormal or normal classes. In this literature study, the CNN model was used for feature extraction and then the LSTM model was applied for classifying the fruits classes.

## 3 Description and Methodology

Plant diseases are usually caused by infectious substances such as fungi, bacteria and viruses [12]. Symptoms of plant diseases are visible evidence of infection and symptoms are a visible consequence of these types of diseases. Fungal infections cause symptoms such as blisters, mildew, or mildew and the basic symptoms are leaf spot and yellowing [13]. Fungal infections are fungal infections caused by mold. The fungus can be solitary or multi-cellular, but in any case, infects plants by stealing nutrients and breaking down tissues. Fungal infections are the most common plant disease. There are some symptoms of the condition, or visible effects of the disease, on the plants. Fungal infections can be identified by symptoms such as spots on plant

leaves, yellow leaves, and bird spots on the berries. With some fungal diseases, the body itself can be viewed on the leaves and appear as a growth and as a fungus. The leaf affected by fungal infection is represented in Fig. 1.

This may be irregular with the stems or under the leaves. This specific target of a pathogen is called symptoms of a single-cell, prokaryotic virus. Bacteria are everywhere and many can be beneficial, but some can cause disease in humans and plants. Symptoms of bacteria are often harder to spot than mold, because they are smaller bacteria. When you cut an infected stem, a white milk-colored substance, called bacterial ooze, may appear. Other symptoms include water-soaked sores, which are wet areas on the leaves that bloom. Eventually, as the disease progresses, the lesions grow and develop reddish-brown spots on the leaves. The most common sign of infection is leaf spot or fruit spots. Unlike fungi, these are usually composed of arteries. The leaf affected by bacterial is indicated in Fig. 2.

The process of the plant disease process consists of four stages. The first stage involves obtaining photos with a digital camera and mobile phone or on the web. The second stage divides the image into different number groups for groups where different strategies can be used. The next section contains methods of removal of the feature and the final section is about the classification of diseases, and the work follow of the proposed model is given in Fig. 3.

**Fig.1** Leaf affected by fungal infection

**Fig. 2** Leaf affected by bacteria

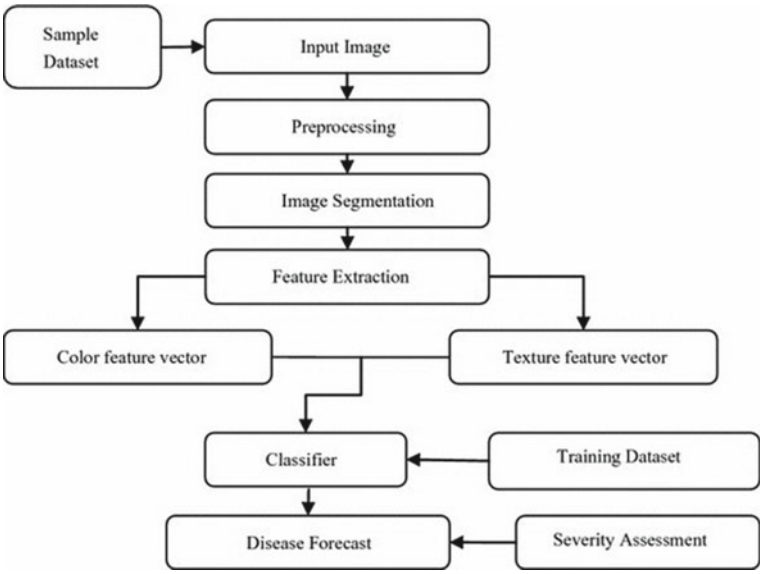



**Fig. 3** Workflow of the proposed model

## *3.1 Image Acquisition*

In this section, photos of the leaves of the plants are collected using digital media such as camera, mobile phones etc. with the adjustment and size you want. Photos can also be taken from the web. Image data formatting depends entirely on the application development tool. To get best functioning of the partition in the final phase of the acquisition process the data base is playing crucial role.

## *3.2 Image Segmentation*

It primarily focuses on simplifying the image's representation so that it sounds better and is easier to process. As a basis for element rendering, this section is also a basic method of image processing. There are a variety of methods for categorizing photos, including k-means clustering, Otsu's algorithm, and adhesion, among others. Clustering based on a set of elements in the $K$ number of classes divides objects or pixels. The division is performed by lowering the total number of distances between items and the groups to which they belong.

### 3.3 Feature Extraction and Classification

Therefore, at this stage we only focus on extraction of image and its related parameters. The parameters are required to obtain a description of the sample image. Features can be based on color, shape, and texture. Recently, most researchers intend to use texture features to diagnose plant diseases. There are a variety of feature extraction methods that can be used to create a system such as the gray-level co-occurrence matrix (GLCM), color occurrence method, spatial gray-level dependence matrix, and histogram-based feature extraction. The GLCM features includes cluster shade, contrast, maximum probability, difference variance and sum entropy, which are stated in the Eqs. (1) to (5). These methods are mathematical methodology of texture separation, and classification is performed by artificial neural network.

$$Clustershade = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i + j - \mu_i - \mu_j)^3 P_{ij} \tag{1}$$

$$Contrast = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} P_{ij} (i - j)^2 \tag{2}$$

$$Maximum\ probability = max(p_{ij}) \tag{3}$$

$$Difference\ variance = variance\ of\ p_{i-j} \tag{4}$$

$$Sum\ entropy = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} P_{ij} \log P_{ij} \tag{5}$$

where, $\mu$ denotes mean and $P_{ij}$ indicates normalized co-occurrence matrix.

## 4 Experimental Results

In the research paper, we are trying to design a way in which we can detect and test whether the affected plant is infected. We do this experiment by using MATLAB and image processing algorithms.

### 4.1 Discussion

In the existing systems, the deep learning algorithm training is not very efficient and it results in many false disease detections. Hence, the removal of good crops with an impression of being bad [14, 15]. In the performance analysis section, the proposed model obtained 97.8% of classification, and the existing models SVM [14]

**Table 1** Comparative analysis between the proposed and existing method

| Method | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|
| SVM [14] | 89 | 76 | 89 |
| LSTM [15] | 90.27 | 85 | 89.83 |
| **Proposed** | **97.80** | **96.05** | **97** |

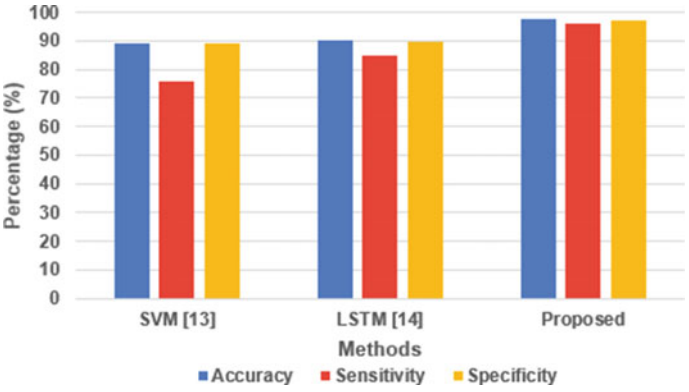



**Fig. 4** Graphical analysis between the proposed and existing method

and LSTM [15] obtained limited performance of 89 and 90.27% of classification accuracy, and also showed better performance in terms of sensitivity and specificity. Comparative analysis is given in Table 1, and Fig. 4.

## *4.2 Outcome of Proposed System*

In the proposed system, we try to design and implement a system with proper amounts of training required by the system in less time before deploying it, and the output of the proposed model is represented in the Figs. 5 and 6.

**Fig. 5** Classification of leaf diseases

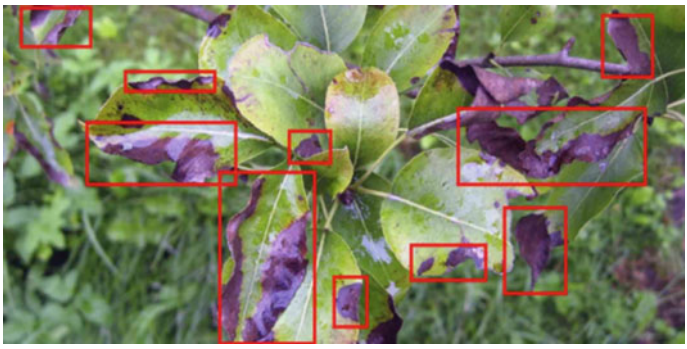


**Fig. 6** Identification of leaf disease

## 5 Conclusion

This paper descries a research of various disease classification methods used for plant disease detection, as well as a classification technique algorithm that may be utilized for later automatic detection and classification of plant leaf diseases. The 10 most extensively tested algorithms are bananas, beans, jackfruit, lemons, mangoes, potatoes, tomatoes, and sapota. As a result, diseases connected with these plants were collected for analysis. The findings were produced with relatively little effort, demonstrating the efficiency of the suggested algorithm in the acceptance and classification of leaf diseases. Another advantage of this technology is that it can detect plant diseases early or in their early stages. Bayes division, Fuzzy Logic, and hybrid algorithms can be utilized to improve the recognition rate of the Artificial Neural Network separation process.

## References

1. Ghaiwat SN, Arora P (2014) Diagnosis and classification of plant leaf diseases using imaging techniques: a review. Int J Recent Adv Eng Technol 2(3):2347–2812. ISSN (online) is Google Scholar
2. Dhaygude SB, Nitin PK (2013) Agricultural plant leaf disease detection using image processing. Int J Adv Res Electr Electron Instrum Eng 2(1)
3. Badnakhe Mrunalini R, Prashant RD (2011) An application of K-means clustering and artificial intelligence in pattern recognition for crop diseases. Int Conf Adv Inf Technol 20. 2011 IPCSIT
4. Arivazhagan S, Newlin Shebiah R, Ananthi S, Vishnu Varthini S (2013) Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agric Eng Int CIGR 15(1):211–217
5. Kulkarni AH, Ashwin Patil RK (2012) Applying image processing technique to detect plant diseases. Int J Mod Eng Res 2(5):3661–3664
6. Bashir S, Sharma N (2012) Remote area plant disease detection using image processing. IOSR J Electron Commun Eng 2(6):31–34. ISSN 2278–2834

7. Naikwadi S, Amoda N (2013) Advances in image processing for detection of plant diseases. Int J Appl Innov Eng Manage 2(11)
8. Patil SB et al (2011) Leaf disease severity measurement using image processing. Int J Eng Technol 3(5):297–301
9. Chaudhary P et al (2012) Color transform based approach for disease spot detection on plant leaf Int Comput Sci Telecommun 3(6)
10. Rathod AN, Tanawal B, Shah V (2013) Image processing techniques for detection of leaf disease. Int J Adv Res Comput Sci Softw Eng 3(11)
11. Vijayaraghavan V, Garg A, Wong CH, Tail K, Bhalerao Y (2013) Predicting the mechanical characteristics of hydrogen functionalized graphene sheets using artificial neural network approach. J Nanostruct Chem 3:83
12. Garg A, Garg A, Tai K (2014) A multi-gene genetic programming model for estimating stress-dependent soil water retention curves. Comput Geosci 1–12
13. Kaur N, Devendran V (2020) Novel plant leaf disease detection based on optimize segmentation and law mask feature extraction with SVM classifier. Mater Today Proc
14. Verma T, Dubey S (2021) Prediction of diseased rice plant using video processing and LSTM-simple recurrent neural network with comparative study. Multimedia Tools Appl 1–32
15. Peker M (2021) Multi-channel capsule network ensemble for plant disease detection. SN Appl Sci 3(7):1–10
16. Vasumathi MT, Kamarasan M (2021) An effective pomegranate fruit classification based on CNN-LSTM deep learning models. Indian J Sci Technol 14(16):1310–1319

# Instance Based Authorship Attribution for Kannada Text Using Amalgamation of Character and Word N-grams Technique


**C. P. Chandrika and Jagadish S. Kallimani**

**Abstract** Authorship Attribution is the task of identifying a true author of a given text from a set of suspected authors stylometry features play a vital role in recognizing the right author, it includes lexical and syntactic features. N-gram is one of the popular techniques used to extract syntactic features from the text. The main objective of this work is to use both lexical and syntactic features on a Kannada text and compare the performance of both approaches using different machine learning algorithms. The Kannada language is spoken by the Indian southern state Karnataka. Even though we can see major works in text processing, Authorship Attribution is in a tender state. Researches have been carried out on handwritten Kannada documents but not on digital text. Char n-gram, word n-gram and the combination of these two known as Amalgamation technique are used as syntactic features to extract the writing style of an author. The results show that Support Vector Machine algorithm outperform with 94% and 60% accuracy using N-grams and lexical features respectively.

**Keywords** Authorship attribution · Decision tree · Instance based approach · Machine learning algorithms · Naïve bayes · Profile based approach · Random forest and support vector machine

## 1 Introduction

Authorship Attribution is a process of predicting a true author from a candidate set of authors. It is one of the applications of Natural language processing, this domain also requires the knowledge of machine learning algorithms and information retrieval. This is one of the emerging fields and its applications can be found useful in

C. P. Chandrika (✉) · J. S. Kallimani
M S Ramaiah Institute of Technology, Bangalore 560054, India
e-mail: chandrika@msrit.edu

J. S. Kallimani
Visvesvaraya Technological University, Belagavi, Karnataka, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_51

Literature, Digital forensics, Plagiarism check, etc. There are mainly two approaches used for predicting the author:

## *1.1 Profile Based Approach*

All available text samples by a candidate author are treated cumulatively, that is all text samples of an author are concatenated as one big document and then a single representation is extracted to create a profile of the author. All the authors' profiles are trained and used for testing an anonymous text. Profile based approach is the best choice for short articles.

## *1.2 Instance Based Approach*

In this approach, all available samples by an author are treated individually. Each text sample has its own representation. In this way, all samples have individual features set which is used for training and testing the anonymous text. It is better to use Instance based approach for lengthy articles.

Every writer has his own style of using a vocabulary set, stylometry features are used for extracting these writing styles. Lexical features and Syntactic features are extracted in the proposed work and are listed in the feature extraction task under Method section.

**N-gram Technique.** N-gram is a popular technique mainly used to extract syntactic and semantic features, also sometimes called as SN gram. The well-known N-gram approaches are character and word N-grams. N-gram mainly used to predict the next possible character and a word used by an author.

**Character N-gram.** Given a set of characters in a word, it will predict the next possible character in a word by modeling a sequence of characters. Consider an example, for the following sentence,

"ಬದಲಾವಣೆ ಜಗದ ನಿಯಮ"

The result of the character n-gram where N value is 3 is as follows:

["ಬದಲಾ", "ಲಾವಣೆ", "ದಲಾವ", "ವಣೆ", "ಣೆ ಜ" ,'ಗದ', 'ಜಗ', 'ಜಗದ', 'ಣೆ', 'ದಲ', 'ದಲಾ', 'ನಿ', 'ನಿಯ', 'ಬದ', 'ಬದಲ', 'ಯಮ', 'ಲಾ', 'ಲಾವ', 'ವಣ', 'ವಣೆ', ] so on.

To train the model using character N-gram, the input that is the sequence of articles written by different authors should be represented as a sequence of characters.

**Word N-gram.** Similar to characters, words can also be modeled so that the system will predict the next possible word in a sentence. For example, consider the statement:

" ಇರುವುದೆಲ್ಲವ ಬಿಟ್ಟು ಇರದುದರೆಡೆಗೆ ತುಡಿವುದೇ ಜೀವನ."

The result of the word n-gram where N value is 3 is as follows:

["ಇರುವುದೆಲವ ಬಿಟ್ಟು ಇರದುದರೆಡೆಗೆ", "ಬಿಟ್ಟುಇರದುದರೆಡೆಗೆ ತುಡಿವುದೇ", "ಇರದುದರೆಡೆಗೆ ತುಡಿವುದೇ ಜೀವನ"] and so on.

Using character and word N-gram, we can extract the vocabulary feature like how an author picks the words, it indicates how much knowledge an author has, about a language. Using the same pair of words indicates low vocabulary and different words indicate high vocabulary knowledge.

## 2 Related Works

The techniques or methodologies used by different researchers on other languages is discussed in this section:

Authors in [1] have tried cross domain authorship using an efficient document vector based on N-gram technique. They have focused on character, word and POS patterns to build the vectors and the highest average accuracy of 87.32% using the POS feature is obtained. The Ensemble technique using variable character and word N-gram model is discussed in [2], authors have employed multinomial logistic regression. The Ensemble model outperforms with the highest accuracy of 82.3%. Authorship attribution using NN model for the short article is demonstrated in [3], authors have compared SRI Language model with the proposed model using word N-gram as a baseline method and achieved 95% accuracy. Authorship profiling is also a method of identifying the true author, using N-gram and stylistic features authors [4] this work focus on how these features are effectively used as vectors. Hinglish is a combination of Hindi and English languages. Authorship prediction on Hinglish WhatsApp messages is illustrated in [5]. The designed model extracts text style features using the character and word N-gram technique, using SVM with character N-gram and they obtained an accuracy of 95.7%. Performance analysis of authorship attribution model with and without using Neural network is presented in [6], the author has demonstrated with multiple datasets using N-gram with NN. The proposed model achieves an accuracy of 80% by extracting valuable features using NN model.

Instance based approach for AA outperforms profile-based method using LDA and N-gram which is discussed in [7]. Instance- based technique achieved 93.17% accuracy. LDA with cosine similarity is found to be more useful for the Urdu language. Cross -domain authorship model on four different foreign languages is developed by authors in [8]. They have applied N-gram technique to extract features from raw text, processed, POS tagging and stemming text. Authors have used various classifiers with ensemble technique which combines the accuracy of all the classifiers and used soft voting to predict the accuracy. The Overall accuracy of 87% is obtained. Authors in [9] have tried to extract stylometry features of Arabic text and trained and tested with fifteen different classifiers and multilayer feed-forward neural network. With ANN, a maximum of 78% accuracy is obtained compared to other ML algorithms. AA using

syntax, N-gram and POS tagging as features is shown in [10], authors have shown the accuracy obtained by classifying the authors using different classification algorithms, with these features separately and combining all three features known as a fusion of features outperform the state of art and top accuracy of 96% on IMDb62 dataset is obtained. Another useful technique SABA [11] based on Stylometric Authorship Balanced Attribution is used for extracting the best features like non-word attribute set that includes sentence length, special characters and punctuation symbol and obtained 87% accuracy. The authors in [12] have tried to predict author of an anonymous Arabic language text by extracting lexical and N-gram features and trained using different classifiers and deep learning technique and they achieved 75.46% accuracy. AA on the Russian language using deep learning networks is discussed in [13]. Authors have Extracted 33 to 5000 features and maximum accuracy above 90% using Deep neural networks is obtained. Applying N-gram with the techniques of Natural language processing and ML algorithms [14] proved to get good accuracy of 90% in predicting the authors of blogs and short messages from various social websites based on uni and bigrams techniques. Binary N-gram is a variant of N-gram [15] which treats texts as binary, it is effective as character N-gram which extracts a more useful frequent pattern of characters/ words in extracting the authors writing style.

## 3 Methods

The dataset collection is a challenging task in Kannada, not many articles are available online freely. We require each author's several articles for training and testing, so it is hard to obtain a huge dataset for the specific domain, here we have considered philosophy related documents. Totally 18 authors of 100 instance documents are collected, out of which, for N-gram technique: 65% is considered for training and 35% for testing. For the lexical model, 75% of the dataset is reserved for training and the remaining 25% is for testing. Figure 1 shows the detailed steps involved in developing the model.

### *3.1 Data Preprocessing and Encoding*

In order to retain the writing style of an author not much preprocessing is done on the text. The extra characters which don't contribute any meaning to this work were removed. To extract few features, punctuations and special characters are not required, in that case, it is removed only during the extraction of that particular feature. Encoding is an important task that represents raw text data into the machine understandable format. For this proposed work, TF-IDF encoding is used for train and test the documents to indicate how pertinent the word is. The TF and IDF are calculated as Eq. (1):

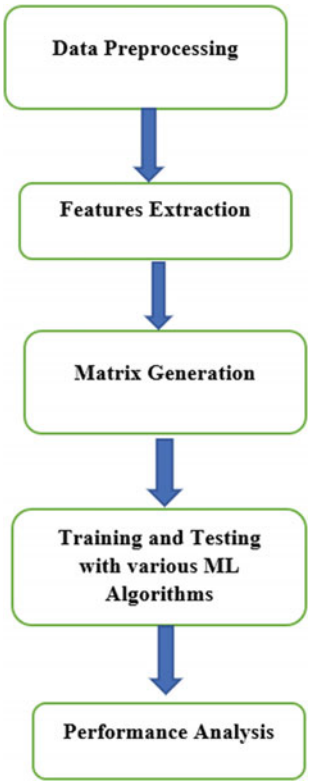



**Fig. 1** The overall workflow of the proposed model

$$TF = \frac{Frequency\ of\ Kannada\ words\ in\ an\ Instance\ i}{Total\ Number\ of\ Kannadawords\ in\ a\ sentence} \tag{1}$$

and IDF is calculated as Eq. (2):

$$IDF = \log\left(\frac{Total\ number\ of\ sentences\ in\ an\ instance\ i}{Total\ number\ of\ sentences\ has\ that\ particular\ word}\right) \tag{2}$$

Finally, the TF and the IDF values are multiplied to get the TF-IDF value for each word from all the documents. The values are represented as a sparse matrix. It indicates the TF-IDF score for all non-zero values in the word vector for each document. The below Fig. 2 shows the sample TF-IDF values obtained for an instance.

The general form: (A, B) C where A: Document index B: Specific word-vector and Index C: TFIDF score for word B in document A. This is a sparse matrix. It indicates the TF-IDF score for all non-zero values in the word vector for each document.

```
instance : 5
  (0, 9)        0.025383205103290507
  (0, 23)       0.0168180044642011
  (0, 25)       0.028412176572961974
  (0, 31)       0.016552252878002365
  (0, 32)       0.044199836789645355
  (0, 33)       0.013953638297030276
  (0, 43)       0.11638637135543511
  (0, 50)       0.020553280881049543
  (0, 64)       0.030135265844385112
  (0, 82)       0.05210188101376039
  (0, 86)       0.02877452154088751
```

**Fig. 2** Sample of Tf-Idf values for Instance 5

## *3.2 Feature Extraction*

Features are obviously more important for this AA model. It is a classification problem of predicting the true author, quality features influence the accuracy of the model. The following is a list of different stylometry features extracted for our work:

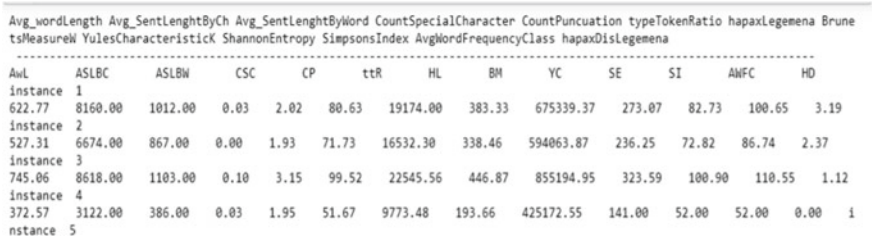
1. Sample Lexical features extracted:

   i. Average Sentence Length By character
   ii. Average word length
   iii. Average Sentence Length By word
   iv. Count Special Character
   v. Count Punctuations
   vi. Type Token Ratio

2. Syntactic features:

   i. Character n-grams (1–5 chars)
   ii. Word n-grams (1–5 words)
   iii. Union of both (1–5)

The TF-IDF values obtained after encoding all the instances are an unbalanced vector that can't be directly used for ML algorithms. This will be converted into two balanced sparse matrix one for N-gram and another for the lexical model, in which each row represents instances of an author and in N-gram column represents the features char, word and union of both. In lexical, each column represents different stylometry features like average sentence length by character/word, count special characters, punctuations, etc. as mentioned under the lexical feature set. After generating matrix, now the dataset is ready for training and testing, here based on different features two sub models are identified: Lexical and N-gram models which extract the writing styles of an author. The details of the implementation are given below:



```
Avg_wordLength Avg_SentLenghtByCh Avg_SentLenghtByWord CountSpecialCharacter CountPuncuation typeTokenRatio hapaxLegemena Brune
tsMeasureW YulesCharacteristicK ShannonEntropy SimpsonsIndex AvgWordFrequencyClass hapaxDisLegemena
 ---------------------------------------------------------------------------------------------------------------------------
AwL       ASLBC       ASLBW       CSC       CP       ttR       HL       BM       YC       SE       SI       AWFC       HD
instance  1
622.77    8160.00     1012.00     0.03      2.02     80.63     19174.00     383.33     675339.37     273.07     82.73     100.65     3.19
instance  2
527.31    6674.00     867.00      0.00      1.93     71.73     16532.30     338.46     594063.87     236.25     72.82     86.74      2.37
instance  3
745.06    8618.00     1103.00     0.10      3.15     99.52     22545.56     446.87     855194.95     323.59     100.90    110.55     1.12
instance  4
372.57    3122.00     386.00      0.03      1.95     51.67     9773.48      193.66     425172.55     141.00     52.00     52.00      0.00     i
nstance  5
```

**Fig. 3** Sample lexical features extracted with values

## *3.3 Lexical Model*

As already mentioned, the dataset has several articles from different authors, each article is treated as an instance so out of articles of 18 authors, we created 100 instances. The above mentioned lexical features are extracted from each author to analyze their writing style. This model focuses on the primary writing skills of an author, it provides basic information about usual writing styles like average word length/average sentence length by character/ word, frequency of nouns. Some authors may have a habit of using more adjective related words which are usually quoted and some special symbols to express their feelings. The extracted features are stored separately as instances. In Fig. 3, the top row demonstrates the lexical features extracted from a training dataset.

The system is trained using several ML algorithms like SVM, Decision Tree, Random Forest and NB to learn these necessary features to predict the probability of the true author. When a new anonymous text document is given as an input for authorship prediction, its lexical features are extracted and now a model will be called to compare the extracted features with trained features to classify and predict the author. SVM outperformed with 60% accuracy compared to other ML algorithms.

## *3.4 N-gram Model*

The lemmatization and stop words removal have been done on the text as one of the steps in preprocessing and then the cleaned data which is in a vector format is passed as an input to this model. N-gram mainly helps to remember the usage of the words of an author, each author has a unique style of using words pattern in a sentence and characters pattern in a word. Using the python library functions, character N-gram with N values as 1, 2, 3, 4 and 5, character sequences are extracted. Similarly, the word patterns are also drawn out from the training set using word N-gram with N values ranging from 1, 2, 3, 4 and 5. These extracted characters/words features are trained and tested using several machine learning algorithms like Decision tree, Random Forest, NB, and SVM. In this work, character and word N-gram has been tried separately and found that, SVM performed well with character N gram with N

**Table 1** Performance of N-gram model

| Algorithms | N = 1 | | N = 2 | | N = 3 | | N = 4 | | N = 5 | | Union |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Char | word | char | Word | Char | Word | Char | Word | Char | Word | Char & word |
| SVM | 60 | 80 | 60 | 82 | 82 | 60 | 91 | 31 | 94 | 14 | 94 |
| NB | 85 | 54 | 62 | 51 | 54 | 60 | 54 | 37 | 57 | 14 | 65 |
| DT | 97 | 60 | 62 | 54 | 60 | 30 | 60 | 17 | 71 | 28 | 68 |
| RF | 68 | 54 | 74 | 40 | 68 | 30 | 71 | 11 | 60 | 28 | 68 |

= 5. To improve the accuracy of the N-gram, we propose an idea of amalgamation of both character and word N-grams features. This union helped to extract more writing patterns and is labeled as Union as shown in the last column in Table 1. In this, first the features of character N-gram are drawn out and fed as an input to word N-gram module and more word features are extracted, these features are trained and tested with different ML algorithms. The average accuracy of each ML algorithm is tabulated in Table 1. N-gram model outperforms the Lexical model and also to generalize the accuracy rate, cross validation with K = 10 is done.

## 4 Results

The Table 1 shows the accuracy obtained after implementing the N-gram technique for both character and word, the graph in Fig. 4 shows SVM outperforms well with maximum accuracy of 94% with char N-gram with N = 5 and for word N-gram feature, it works better with N = 1 and 2 but gradually drops down for the testing dataset. NB, RF and DT performed well with initial values of N but accuracy drops out in between with the increasing value of N. When the character and word N-gram features are combined, other algorithms have also progressed by improving the accuracy reaching to 70%.

### *4.1 Discussion*

The proposed model's performance can be compared with the lexical model, which identifies the author based on the lexical features and Table 2 shows the difference in the accuracy of ML algorithms after training N-gram and lexical model. The N-gram performs well compared to the lexical model. In the previous work AA using Profile based approach based on the lexical model has been carried out and we obtained 85% accuracy for four authors dataset. Surprisingly with the increase in the dataset we got less accuracy when tried with Instance based approach for lexical model but improved accuracy of 94% with the N-gram model. Based on the previous work, it is proved that the N-gram method is best suited for Instance based approach for our dataset.

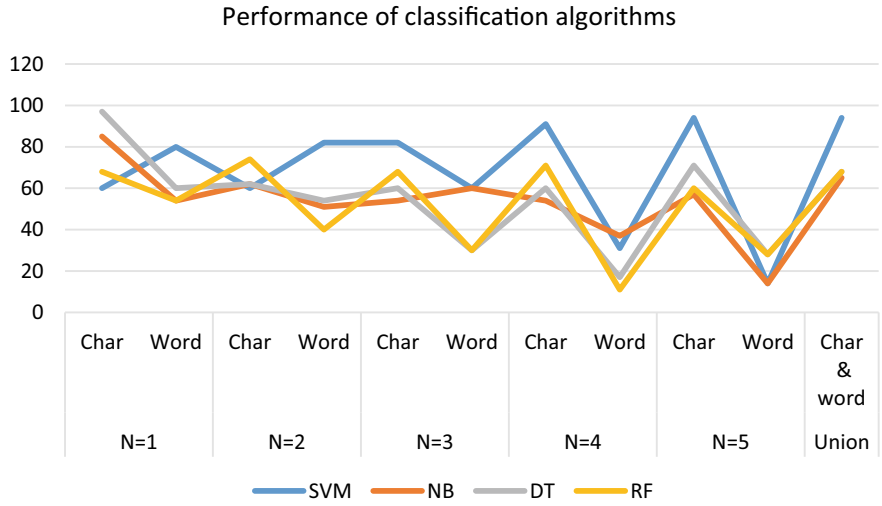



**Fig. 4** Accuracy distribution of N-gram model

**Table 2** Performance comparison

| Algorithms | Feature's type: N-Grams/Lexical | Precession | Recall | F1-score | Accuracy |
|---|---|---|---|---|---|
| SVM | N-Grams | 0.96 | 0.94 | 0.94 | 0.94 |
| | Lexical | 0.6 | 0.6 | 0.57 | 0.6 |
| NB | N-Grams | 0.56 | 0.66 | 0.57 | 0.65 |
| | Lexical | 0.22 | 0.36 | 0.27 | 0.36 |
| DT | N-Grams | 0.68 | 0.66 | 0.62 | 0.68 |
| | Lexical | 0.38 | 0.40 | 0.37 | 0.4 |
| RF | N-Grams | 0.63 | 0.69 | 0.63 | 0.68 |
| | Lexical | 0.25 | 0.32 | 0.26 | 0.32 |

Comparing to the State-of-the-Art: The proposed model is compared with the few other language models found in literature survey shown in the Table 3. The proposed model performed well with the available dataset. Here only character and word N-gram are considered, accuracy may still improve if we focus on other features also.

**Table 3** Performance comparison of the proposed model with State-of-the-Art

| Dataset | State-of-the-Art | Accuracy/F1 Score in % |
|---|---|---|
| Politics | [1] | 87.32 |
| Multiple domains on five different languages | [2] | 71.4 |
| Twitter dataset | [3] | 85.3 |
| Religious text | [9] | 78.4 |
| IMDb62 dataset | [10] | 96.3 |
| Philosophy articles | Proposed method | 94 |

## 5 Conclusions and Future Works

The proposed model helps to predict the authorship of an anonymous text based on lexical, character, word N-gram and the combination of character and word N-gram techniques. Usually AA based on stylometry features gives better accuracy for other languages. In our previous work we have tried stylometry features using profile based approach on Kannada dataset and obtained around 85% accuracy but it didn't perform well for the instance based approach. The N-gram technique using both character and word patterns with SVM algorithm outperformed with 94% accuracy. In future, we try to test this model with a huge dataset by applying N-gram on POS tagging of Kannada words which can be used as a feature. Profile based approach can also be implemented using N-gram model and an accuracy of this can be compared with the Instance based approach.

## References

1. Gomez Adorno H, Posadas Durán J, Sidorov G, Pinto D (2018) Document embeddings learned on various types of n-grams for cross-topic authorship attribution. Computing 100:741–756. https://doi.org/10.1007/s00607-018-0587-8
2. Custódio JE, Paraboni I (2018) EACH-USP ensemble cross-domain authorship attribution. Notebook for PAN CLEF. https://pan.webis.de/clef18/pan18-web/author-identification.html
3. Ge Z, Sun Y, Smith M (2016) Authorship attribution using a neural network language model. In: Thirtieth AAAI conference on artificial intelligence, vol 30, pp 4212–4213
4. Radha D, Sekhar PC (2019) Author profiling using stylistic and n-gram features. Int J Eng Adv Technol 9(1). ISSN 2249–8958
5. Sharma A, Nandan A, Ralhan R (2018) An investigation of supervised learning methods for authorship attribution in short hinglish texts using char & word n-grams. ACM Trans ALRL Inf Process 1(1)
6. Sari Y (2018) Neural and Non-neural Approaches to Authorship Attribution. Psychology, Computer Science, Corpus ID 106407420
7. Anwar W, Bajwa IS, Ramzan S (2019) Design and implementation of a machine learning-based authorship identification model. Hindawi Sci Program 2019:1–14, Article ID 9431073. https://doi.org/10.1155/2019/9431073

8. Bacciu A, La Morgia M, Mei A, Nemmi E, Neri V, Stefa J (2019) Cross-domain authorship attribution combining instance based and profile-based features. In: Proceedings of bacciu 2019 cross domain AA, Corpus ID 198489778, CLEF
9. Al-Sarem M, Alsaeedi A, Saeed F (2020) A deep learning-based artificial neural network method for instance-based arabic language authorship attribution. Int J Adv Soft Comput Appl 12(2):1–15. ISSN 2074–8523
10. Fourkioti O, Symeonidis S, Arampatzis A (2019) Language models and fusion for authorship attribution. Inf Process Manag 56(6):1–13. ISSN 0306–4573, https://doi.org/10.1016/j.ipm.2019.102061
11. Tareef KM (2019) Non-word attributes' efficiency in text mining authorship prediction. J Intell Syst 29(1):1408–1415. https://doi.org/10.1515/jisys-2019-0068
12. Hossain AS, Akter N, Islam MS (2020) A stylometric approach for author attribution system using neural network and machine learning classifiers. In: 2020: proceedings of the international conference on computing advancements, Article no 22, pp 1–7. https://doi.org/10.1145/3377049.3377079
13. Romanov A, Shelupanov A, Fedotova A, Goncharov A (2021) Authorship identification of a russian-language text using support vector machine and deep neural networks. Fut Internet 13:3. https://doi.org/10.3390/fi13010003
14. Vijayakumara B, Fuad MMM (2019) A new method to identify short-text authors using combinations of machine learning and natural language processing techniques. Procedia Comput Sci 159:428–436. https://doi.org/10.1016/j.procs.2019.09.197
15. Carman M, Ashman H (2019) Evaluating binary n-gram analysis for authorship attribution. Int J Comput Linguist 10(4):60–69 (2019). ISSN 2180–1266

# IoT Enabled Virtual Home Assistant Using Raspberry Pi


**Md. Tarequl Islam, Md. Selim Azad, Md. Sobuj Ahammed, Md. Wahidur Rahman, Mir Mohammad Azad, and Mostofa Kamal Nasir**

**Abstract** This manuscript represents a virtual assistant developed with modern computation named after "JERRY". The proposed model is entirely based on the new Internet of things (IoT) mechanism and Natural Language Processing (NLP). Experimental data analysis on the virtual assistant is also enumerated and interpreted to find the proposed system's feasibility. The accomplished experiments are the practicability testing of the virtual assistant, how the users will be able to interact with home automation in AC supply, a comparison on response time among existing solutions and System Usability Scale (SUS) to check the random user's satisfaction level. Average response time of 1.93 is tracked out, which is more impressive than other available intelligent assistants in the global market. SUS score of 98% is also found. Though the proposed model has some limitations, a reliable smart assistant for regular activities and an interactive home automation system is developed. However, the proposed architecture can be adjustable in the regular activities of the general users.

**Keywords** Internet of Things · Machine learning · Smart home · Virtual assistant · Voice recognition

## 1 Introduction

Virtual Assistant (also known as AI Assistant or Digital Assistant) is an application that can understand voice commands in natural language and complete tasks for

Md. T. Islam (✉) · Md. S. Ahammed · Md. W. Rahman · M. M. Azad
Department of Computer Science and Engineering, Khwaja Yunus Ali University, Enayetpur, Sirajganj 6751, Bangladesh
e-mail: tareq.cse@kyau.edu.bd

Md. S. Azad
Department of CSE and CSIT, Shanto Mariam University of Creative Technology, Dhaka, Bangladesh

M. K. Nasir
Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail 1902, Bangladesh



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_52

users. The virtual assistant has gained tremendous popularity in the contemporary world. According to STATISTA [1], the number of voice assistant users in the United States is expected to reach 123 million by 2021. In 2015–2021 there were around 1.6 million users worldwide, and by the end of 2021, it reached 1.8 billion. By 2021, business growth is expected to increase from 155 million in 2015 to 843 million in 2021. Interfaces with virtual assistive devices can support largely in daily activities of humankind such as browsing messages, searching for anything in Google or Wikipedia, or questioning temperature and other data. The voice-activated virtual assistant makes it easier for the blind or disabled person to access the internet and support regular household activities [2]. The statistics on IoT with home automation [3], the number of IoT devices in homes rose up to 12.86 billion in the year 2020. This statistic remarkably indicates the importance of IoT devices in the home automation or smart home sectors.

Therefore, the proposed system implements an integrated system for home automation with a virtual assistant. The proposed system plays a vital role in interactive home automation and is also used for a complete personal virtual assistant. Though there are some existing solutions in virtual assistant, no solution has yet been met to control home appliances in 220 V AC supply efficiently. Thus, the contribution of this article is as follows:

- To control any home appliances (connected with relay) through voice command like "turn on/off light."
- To assist the user by performing math calculations, searching Wikipedia, checking emails, current time, date, etc.
- To provide you information from the internet about "any person/place organization you want to know by the command of "tell me about/tell me/tell me/say about/the name of that person/place/institution" or speaking "information from Wikipedia in command".

## 2 Related Work

This section presents previous contributions in artificial intelligent virtual assistant. Numerous major contributors are making a massive impact in the field of intelligent assistant development and helping in the deep understanding of home automation-based IoT. The author of this paper [4] had proposed a virtual assistant that can communicate with the stranger at the door from a specific distance. It also could notify the owner through e-mail and SMS with the captured photos of stranger. The authors in the manuscript [5] proposed a python task that was able to accomplish some features like climate news, playing song from playlist, information about any film, Wikipedia, and controlling some electronic devices. The authors in the paper [6] had proposed a system that is able to control some home appliances at the same time. In the paper [7], This paper proposes an efficient home automation system using Node Microcontroller Unit (Node MCU) which focuses on sensing and maintaining suitable indoor climate conditions like temperature and humidity. The author

proposed a method [8] where the system aims at processing human natural voice and gives a meaningful response to the user. In the article, [9] shows the working of a device based on implementation of a voice command system as an intelligent personal assistant which have many features such as weather, telling time or accessing online applications to listen to music. Authors of that articles [10] had proposed a system where a micro-controller of Raspberry pi takes user inputs from a created website access through a username and password. Where the Raspberry pi can be controlled from any place with the help of weaved cloud service, in this papers, [11] the authors represents a system called Intelligent Personal Assistant (IPA), which is a software agent that is able to assist the general users in their daily life. A wireless sensor is needed with proper internet connections to provide information about any detecting objects autonomously through IoT vision. The authors in this article [12] proposed a home automation system using Raspberry pi. A smart phone needs to connect with Raspberry pi using IP address to accessing the system through Wi-Fi. The authors of this paper introduce [13] with a personal assistant based on python that can control home appliances using voice and gesture command. In this article, the authors represent [14] an Artificial Intelligent based personal assistant which is having a virtual personality of its own, just like a particular man has in a specific profession. In this paper, the authors present [15] a speech-enabled, customizable personal assistant for the innovative environment. To overcome some virtual assistant issues, they were used computer vision, deep learning, speech generation, and recognition with artificial intelligence as a personal voice assistant for developing smart home automation.

Moreover, the paper [1–12] have only focused on developing virtual assistant and innovative home concept. But the proposed model has presented the integrated architecture for virtual assistant and IoT based system, which has been tested in a 220-V AC supply. In addition, the proposed system's response time is faster compared to the existing techniques which can give immense pleasure to the customer to use it.

## 3 Proposed Methodology and Working Principle

In this section broadly describes the working procedure of the proposed system. Overall activity diagram of the proposed system is depicted in the Fig. 1. The virtual assistant named "Jerry" will make user tasks over internet easier. In this framework, Jerry is associated with a Raspberry pi. The relay module is similarly associated with the Raspberry Pi through a GPIO pin. When the user gives any command through the microphone, it sends the user voice to the microprocessor of Raspberry pi. The microprocessor takes the voice data and going to perform that task. The output of that task will be announced by the system through a speaker. Jerry is also able to handle any browsing task or replying any type of questions. It can also be able to control home appliances.

After running the system, Jerry wishes the users good morning/good evening/good evening depends on the time. While the system completes the wishing step, now it is

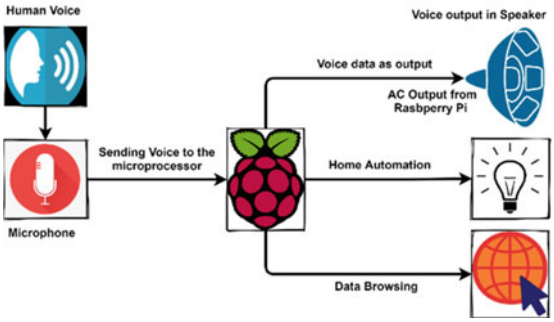



**Fig. 1** Overall System illustration

ready to take voice command from the user and speak, "I'm listening, how can I help you? After receiving command, Jerry will recognize the command what the user send and display the user command. After recognizing the command, it checks whether the user's command is matched or not. If the command is matched, the user's voice will be sent to Jerry as text in Fig. 2. In case of unmatched, it will recognize the instructions performing a google search of that corresponding command. When it detects the user's voice, it goes through Google text to speech API and converts commands into text. After getting the command text, it will check whether the command text is matched or not with cloud data. If the command text is matched, it will go to the execution step of that corresponding command. If the command text is not matched, it will go to the initial step and notify to user that "I didn't get it, can you repeat it again?

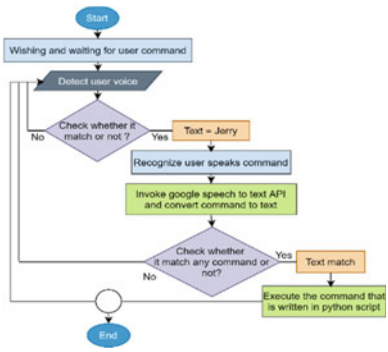



**Fig. 2** Activity diagram of the proposed model

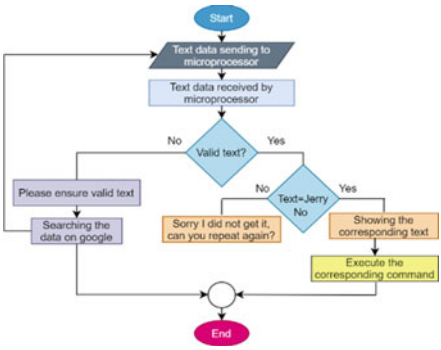



**Fig. 3** Working principles of voice detection

## *3.1 Working Principles of Voice Detection and Recognition*

Figure 3 depicts the workflow diagram of the system detected the voice command from the corresponding users. In this section, user voice text data sends to the microprocessor of Raspberry pi. Once the voice as the text is received, it goes to the next step of verification whether the text is valid or not. If the voice is valid, it will check whether the text is matched with any command with Jerry or not? In case of mismatch, it will show on the screen as text and speak. "Sorry, I didn't get it. Can you repeat that again? If the command is matched, the text will show on the screen and will tell to the user the same information by voice command. If the voice data is not valid, it will speak "Please ensure valid text" and perform a google search of that corresponding text. On the other hand, If the voice data is received, the data is shown in the screen as "user said: text" by recognizing the user voice module. If the voice data is not received, it will speak and show text as "Sorry; I didn't get it, I am opening google for you". After that, google is opened for the user and go back to the listening step again and speak, "How can I help you? Fig. 4 shows the respective workflow diagram of the user-provided command to the proposed model. Figure 5 illustrate a voice recognition procedure and its corresponding working principles.

## *3.2 Working Principles of Voice Command Execution*

In this section, Jerry waits for the corresponding user's command and says, "I am listening; how can I help you?" When the user gives any command, it will detect the user's voice. While getting voice data from the user, will it match the data whether it matches or not? If the voice data is matched, it will show the command as a text, and after that, it will perform that specific task as the user said. If the voice data is not matched, it will show the text on the screen and also speak, "Sorry, I didn't get it. Can you repeat that again?" After that, it will go back to the listening step and again

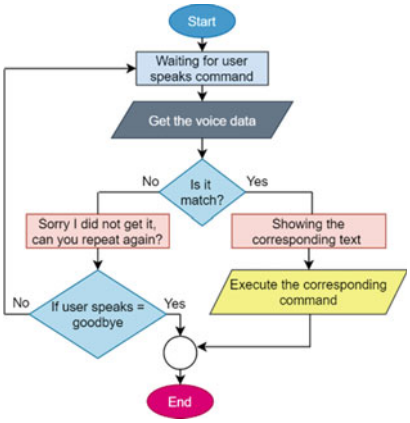



**Fig. 4** Working principles of user voice command


**Fig. 5** Working principles of user voice recognition

ready to take voice. If the user says "goodbye" in command, then it will terminate the system. Figure 6 shows the corresponding workflow diagram of how the google speech recognized API [16] will be performed in the proposed model.

After taking the voice data from user by the microprocessor, it will go to the Google speech recognition step, where it will check whether the internet service is available or not? If the internet service is available, then it will send the voice data [17] to the cloud. If the voice data is matched with any pre stored command in the cloud, it will execute that particular command. It will show the text and speak "use proper text for executing the command" if not match. While voice data from the users is taken, it will send to the Google speech recognition section. After recognizing the voice, google text to speech API will go through where the command data [18] is converted into text. Here the user's voice is checked whether it matches with any


**Fig. 6** Working principles of google speech API

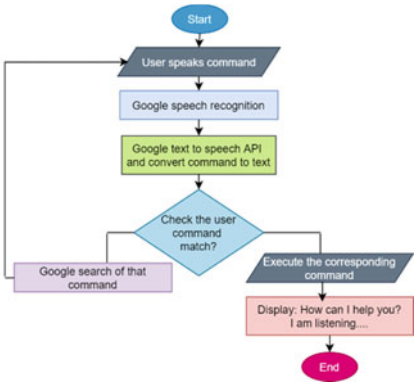



**Fig. 7** Working principles of user command execution state

declared command or not. If the command is matched, it will go to the execution step to perform that task. And again, go back to the listening step and ready to take voice input from the user. If the voice is not matched, it will perform a google search of that corresponding command. Figure 7 present the respective workflow diagram of how the user's command will be extracted and executed.

## 4 Result and Discussion

Here experimental data analysis is presented for the intelligent home assistant. and comparison among Google assistant, Amazon Alexa, Apple Siri and this proposed intelligent home assistant. In addition, a result of the system usability scale (SUS) in shown. and finally, present some snapshots of the proposed developing system.

### *4.1 Experiment on Developed Smart Assistant*

Some experiments on smart home assistant are made for analysis. The experiment had some subsection like sending data voice, recognizing the voice, response time, connection stability, idle time, etc. are depicted in Table 1.

From this experiment, among 15 times, around 13 times of voice is being send and recognized successfully. The response time is also calculated, which was very first. On the other hand, without sending any voice data, the smart assistant will go on to an idle state, which is about 3 or 4 s. Figure 8 shows comparison between response time and idle time. This figure clearly remarks how the proposed system takes input from the current user and how much time the model must process and perform the

**Table 1** Experimental Data analysis of the proposed solution

| Serial no | Sending voice data | Recognize voice | Response time (ms) | Connection stability | Idle time (ms) |
|---|---|---|---|---|---|
| 1 | Yes | Yes | 1000 | Yes | 1500 |
| 2 | Yes | Yes | 1500 | Yes | 1700 |
| 3 | Yes | Yes | 1200 | Yes | 1500 |
| 4 | Yes | Yes | 1500 | Yes | 1600 |
| 5 | No | No | 3000 | No | 3000 |
| 6 | Yes | No | 3000 | Yes | 1500 |
| 7 | Yes | Yes | 1200 | Yes | 2000 |
| 8 | Yes | Yes | 1500 | No | 2500 |
| 9 | Yes | Yes | 1500 | Yes | 1500 |
| 10 | Yes | No | 3000 | No | 2000 |
| 11 | No | No | 3000 | No | 3000 |
| 12 | Yes | Yes | 1500 | Yes | 1200 |
| 13 | Yes | Yes | 1500 | Yes | 1500 |
| 14 | Yes | No | 3000 | Yes | 1500 |
| 15 | Yes | Yes | 1500 | Yes | 1200 |

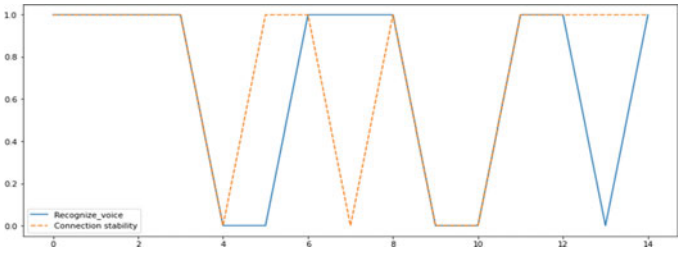



**Fig. 8** A box plot of response vs. idle time of the proposed model

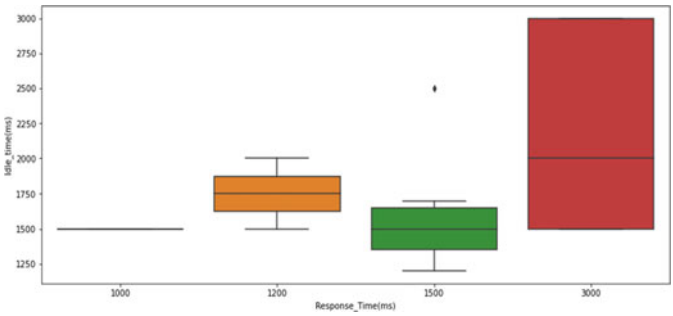



**Fig. 9** A-line chart diagram of recognizing users vs. connection integrity

corresponding command. In Fig. 9 presents a line chart diagram of binary response, how the system recognized upcoming voice data received from the users to developed model, and how internet connection integrity affects the reliability of the proposed solution.

## *4.2 Smart Home Activity Analysis with the Developed Smart Assistant*

Table 2 shows the experimental data analysis of the proposed system. When the relay module and internet connection are in the active state, the user's command will execute. According to the internet speed, the time delay can be varied. While the internet availability is high [19], the delay can minor and execute command very fast.

**Table 2** Response analysis and status monitoring of the IoT components

| Experiment no | Relay activated | Internet availability | Executed command | Time delay (sec) | Results |
|---|---|---|---|---|---|
| 1 | Yes | Yes | Turn on | 2 | Turn on AC bulb |
| 2 | Yes | Yes | Turn on | 1.5 | Turn on AC bulb |
| 3 | Yes | Yes | Turn on | 1.2 | Turn on AC bulb |
| 4 | No | No | Turn off | 2 | Turn off AC bulb |
| 5 | No | No | Turn off | 1 | Turn off AC bulb |
| 6 | Yes | Yes | Turn on | 2.5 | Turn on AC bulb |

**Table 3** A comparison among existing works

| Serial no | Existing approach | Existing response time | Proposed work response time | Comments on comparison |
|---|---|---|---|---|
| 1 | Google [21] | 1.9 | 1.93 | Good |
| 2 | Alexa [21] | 2.3 | | Better |
| 3 | Siri [22] | 3.5 | | Excellent |

### *4.3 Comparison Among Existing Work Among Response Time*

The Table 3 compares average response time among some developed system like google assistant, apple Siri and amazon Alexa. The study shows that google assistant's average response time of 1.9 s, amazon Alexa's response time of 2.3 s and apple Siri's response time of 3.5 s [20]. By calculating the proposed system's average response time is 1.93 which is very nearly google assistant and much faster than Alexa and Siri. As Google speech recognized API have been utilized in this prototype, the proposed work's response time is almost equal to Google. The prototype experienced a slight bit of delay due to the internet connection. Besides, the experimental data have been enumerated and compared from claimed response time available on the website.

### *4.4 Snapshots of the Developed System*

This section ensures the external and internal architectural design of the proposed system model. Figure 10(a) presents architectural orientation, Fig. 10(b) shows the development modules with a relay circuit required for the smart home model in the AC connection. Figure 10(c) shows the overall smart home assistant with a speaker, a microphone, and a computer desktop monitor. Figure 10(d) presents how the proposed system can be interactive for conventional home automation in AC supply.

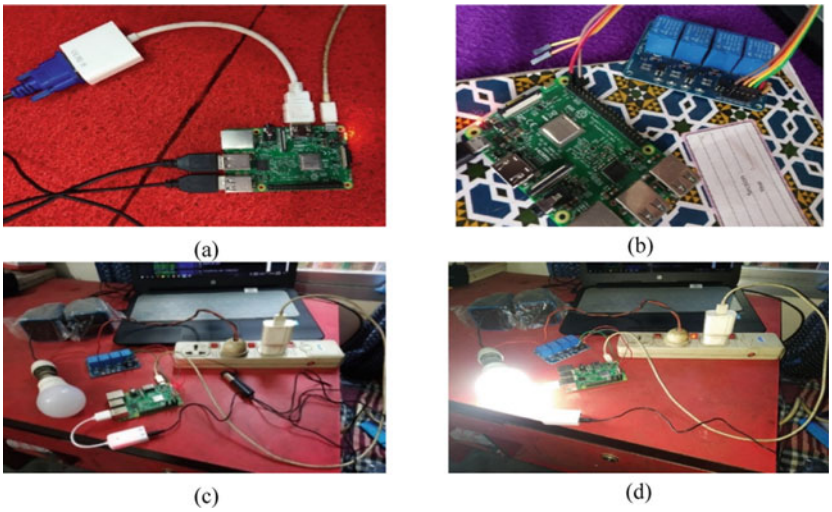



**Fig. 10** Snapshots of the developed model

## 5 Conclusion

This system will be more effective and usable for the users. The users will get both the service of virtual assistant and intelligent home assistant at a time. The virtual assistant home automation is developed after an in-depth understanding of python languages- its module packages and libraries, the working procedure and its offence. Mainly the working principles of Raspberry Pi and its requirements to implement a virtual assistant on the Linux platform. And all of these resources are chosen to keep in mind the expenses and resource availability. Though the system has some limitations, like this system is not workable without internet connection, and the other is the operating system as the system is developed on Linux platform, that's why it is not suitable for another operating system. The limitations will be resolved in future to fix that issues.

## References

1. Statista (2020) Number of voice assistant users in the United States from 2017 to 2022. Accessed 10 Apr 2021, https://www.statista.com/statistics/1029573/us-voice-assistant-users/
2. Rahman MW et al (2021) The architectural design of smart blind assistant using IoT with deep learning paradigm. Internet of Things 13:100344
3. Milenkovic J (2019) Mind-Twisting IoT Statistics for 2020 – Your Home Is Now Online. Accessed 10 Apr 2021, https://kommandotech.com/statistics/iot-statistics/#:~:text=Top%20IoT%20Statistics%2C%20editor's%20pick,to%20the%20Internet%20every%20second.&text=By%202020%2C%20the%20number%20of,the%20healthcare%20industry%20by%202020.

4. Khattar S, et al (2019) Smart home with virtual assistant using raspberry Pi. In: 2019 9th international conference on cloud computing, data science & engineering (Confluence). IEEE
5. Srinivas P et al (2020) Raspberry pi based personal voice assistant using python. Int J Eng Appl Sci Technol 4(11):105–108
6. Kulkarni B et al (2017) IoT based home automation using Raspberry PI. Int J Innov Stud Sci Eng Technol (IJISSET) 3(4):13–16
7. Garg S et al (2020) IoT based home automation. J Inf Optim Sci 41(1):261–271
8. Bajaj D, Pramod D (2020) Conversational system, intelligent virtual assistant (IVA) named DIVA using raspberry Pi. Int J Secur Priv Perv Comput (IJSPPC) 12(4):38–52
9. Singh P, et al (2019) Voice control device using raspberry pi. In: 2019 Amity international conference on artificial intelligence (AICAI). IEEE
10. Sandeep V, et al (2015) Globally accessible machine automation using Raspberry pi based on Internet of Things. In: 2015 international conference on advances in computing, communications and informatics (ICACCI). IEEE.
11. Santos J et al (2016) Intelligent personal assistants based on internet of things approaches. IEEE Syst J 12(2):1793–1802
12. Venkatesh K et al (2018) IoT based home automation using raspberry pi. J Adv Res Dyn Control Syst 10:1721–1728
13. Chayapathy V, Anitha G, Sharath B (2017) IOT based home automation by using personal assistant. In: 2017 international conference on smart technologies for smart nation (SmartTechCon). IEEE
14. Nair G, Johnson S, Sathya V (2018) Chatbot as a personal assistant. Int J Appl Eng Res 13(20):14644–14649
15. Iannizzotto G, et al (2018) A vision and speech enabled, customizable, virtual assistant for smart environments. In: 2018 11th international conference on human system interaction (HSI). IEEE
16. Siegert I, et al (2020) Recognition performance of selected speech recognition APIs–a longitudinal study. In: International conference on speech and computer. Springer, Heidelberg
17. Ermolina A, Tiberius V (2021) Voice-controlled intelligent personal assistants in health care: international delphi study. J Med Internet Res 23(4):e25312
18. Jia Y, et al (2019) Leveraging weakly supervised data to improve end-to-end speech-to-text translation. In: ICASSP 2019–2019 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE
19. Sultan MR, et al (2012) Adrisya sahayak: a bangla virtual assistant for visually impaired. In: 2021 2nd international conference on robotics, electrical and signal processing techniques (ICREST). IEEE
20. Palanica A, Fossat Y (2021) Medication name comprehension of intelligent virtual assistants: a comparison of amazon alexa, google assistant, and apple siri between 2019 and 2021. Front. Digital Health 3:48
21. DGIT (2018) What's the best smart speaker? Apple HomePod vs Google. Accessed 10 Apr 2021, https://dgit.com/apple-homepod-vs-google-home-vs-amazon-echo-53296/
22. Medium (2020) Hey Siri, you're too slow!. Siri is too slow in responding. Accessed 10 Apr 2021, https://medium.com/@marktraphagen/hey-siri-youre-too-slow-c26742a12fc8

# Load Balanced Content Prefetching Model for MANET-CLOUD Environment


**Shashidhara Doddamane Nagendrappa, Chandrappa Dasanapura Nanjundaiah, and Puttamandappa Chaluve Gowda**

**Abstract** With evolution of wireless communication and Internet of things has resulted in generation, collection, and processing of huge amount of data for provisioning various reliable and low-latencies services. MANET (Mobile Adhoc Networks) connected with cloud enabled future generation communication network for storing, processing, and communication of Big Data. Offering uninterrupted admission to data with less latency, cost and better resource utilization is difficult. For meeting such requirement content prefetching considering multiple copy has been envisioned in few existing models. However, these models induce high operation cost as content prefetching are done without considering user request pattern and its location. In addressing above defined problems, load balanced content prefetching (LBCP) model with multiple copy (i.e., replication) with multi-objective parameter using Bi-graph is proposed. The LBCP model minimize query access processing time with better cache consistency, energy efficiency, and throughput when compared with existing content prefetching model.

**Keywords** Bi-Graph · Caching · Content prefetching · MANETs

S. Doddamane Nagendrappa (✉)
SJB Institute of Technology, BGS Health and Education City, Bangalore, India
e-mail: shashidharadn@gmail.com

C. Dasanapura Nanjundaiah
Department of Electronics and Communication Engineering, SJB Institute of Technology, BGS Health and Education City, Kengeri, Bangalore, India

P. Chaluve Gowda
Department of Electronics and Communication Engineering, Dayananda Sagar University, Kumaraswamy Layout, Bangalore, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_53

## 1 Introduction

With rapid development of IoT and communication technologies, MANET have progressed from a research subject to actual provisioning of various real-time applications. MANET provides Device to device (D2D) directly and through intermediate association for addressing the ever-growing application traffic loads needs for MANET device-based services [1–3]. Extensive work has been done for exploiting benefit of utilizing MANET infrastructures to improve communication efficacy, quality of service (QoS) and also for proving quality of experiences (QoE) for its subscribed users. The application traffic pattern of MANET significantly varies when compared with traditional wireless-based applications; thus it is important to study the traffic prerequisite of different mobile devices. Recent work has showed the significance of using cooperative caching mechanism for meeting low-latencies and dynamic mobile nature characteristic of MANETs.

However, existing content prefetching model performs very poor as data placement is done using single-objective function. In order to reduce most of existing model place data in multiple location; however, will result in high cost of service provisioning. Thus, it is important to balance load and design content prefetching model considering multi-objective function. In order to address the aforementioned problems in this work we present load balanced content prefetching (LBCP) model for provisioning multimedia application [13, 18] in MANET-Cloud environment [7, 8], and [14].

*The significance of using LBCP model for MANET-Cloud is described below*

The LBCP model employs multi-objective parameter such as latency and cost for placing data considering user-data relation employing Bi-graph. Thus, aiding in achieving high cache consistency performance with minimal query access delay.

Experiment outcome shows the LBCP model is energy efficient considering varied densities and also archives high throughput considering varied speed in comparison with recent prefetching and routing models.

The paper is arranged as follows. The Sect. 2, discusses about various existing content prefetching models and highlight the limitation of standard models. In Sect. 3, load balanced content prefetching model for MANET-Cloud environment. In Sect. 4, LBCP performance is validated with state-of-art models. Finally, the research is concluded with future research direction.

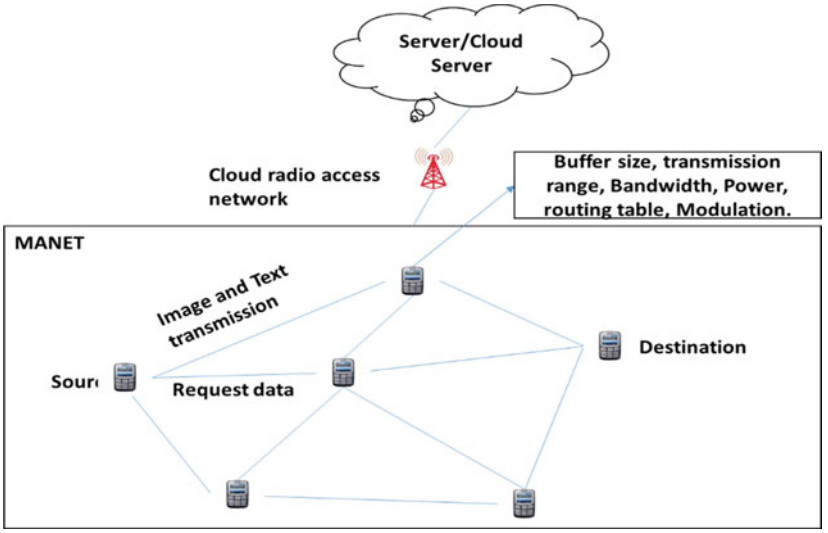



**Fig. 1** Architecture of MANET-Cloud for provisioning multimedia applications

## 2 Related Works

This section study some of recent work presented to improve performance of MNAET for provisioning diverse applications. The traditional MANETs generally allow the connected MANET devices to down-load data through interconnected communication device or directly from datacenter while being mobile in nature. However, for provisioning multimedia applications [4] requires low-latencies, reliability, and seamless connectivity which the traditional MANETs fails to provide.

In order ensure less latencies, QoE, and providing seamless multimedia data delivery, a good content prefetching mechanism of needed [5, 6]. However, content prefetching is a challenging task; because a poor prefetching model will result in high latencies, bandwidth wastage, and I/O overhead. For building good prefetching model it is important to predict the future location of mobile MANET device.

The characteristic of MANETs such as dynamic mobility, rigid connection, unpredictable channel condition, and shorts association time makes location prediction and content prefetching even more a challenging task.

Existing content prefetching model [9–11] for providing reliability and reduce latencies guarantees the content are placed in multiple location; However, the placement is done considering single objective function [12]; resulting in high I/O overhead and bandwidth wastage [15] especially for resource starved devices. Thus, for overcoming above discussed research problems in next section this work presents load balanced content prefetching model for Cloud-MANET.

## 3 Methods: Load Balanced Content Prefetching Method for Cloud Based MANET

In the proposed methodology, content prefetching to multiple node for improving performance of MANET-Cloud is proposed. The architecture of MANET-Cloud is shown in Fig. 1. Here we present multi-objective based content prefetching methodology for minimizing query latency with better bandwidth utilization for achieving fine-grained load balanced content prefetching model considering inter domain mobility of MANET user for mobile adhoc network. For addressing the problem of content prefetching in minimizing content access latency, and balance load among MANET nodes the data are pre fetched across different MANET node. Further, prefetching data on multiple nodes induces higher cost of service provisioning. This work adopts a $Bi - Graph$ based data prefetching method for solving the unfamiliarity of the variance between locations and its association between several contents. Let $L(K, H)$ define a $Bi - Graph$, where $H$ and $K$ defines edges and vertices, respectively. The $Bi - Graph$ provision manifold vertices for every edge sets; and the same-time only two vertices are given for each edges. The LBCP uses vertices set with respect to entire gateway server and content identifier is described as Eq. 1

$$K = IUJ \tag{1}$$

The request characteristics among all the connected pairs and MANET gateway server is defined using $H$ and is established as Eq. 2

$$H = \{ h_a | a \in A \} \cup \left\{ h_{ij} | i \in I, \ \ j \in J \right\} \tag{2}$$

There exist several content information for different request configuration of edge due to adoption of $Bi - Graph$. Then weights are assigned to each edges $h \in H$ for assuring desired level of qualities of services prerequisite of content prefetching and reduce latencies. Here we give different weight for each edges of $Bi - Graph$ due to adoption of multi-objective placement strategy as described in below Eq. 3

$$S = M.\left(e^{[X]}, Q^{[Q]}, Q^{[S]}\right)^{U} \tag{3}$$

where M$M$ represent multi-objective performance vector weights, where $e^{[X]}$ depicts the cost of data processing, and $Q^{[Q]}$ depicts the constant latency $S_{xy}^{Q}$ for obtaining data from MANET datacenter$x$ to MANET datacenter y$y$. Therefore, the latency can be minimized using following Eq. 4 and 5

$$Q^{[Q]} = \sum_{i \in I} \sum_{j \in J} 1_{ij} O_{ij}^{[Q]}, \tag{4}$$

where

$$O_{ij}^{[Q]} = \sum_{S \in J} S_{xy}^{[Q]} G_{ix}. \tag{5}$$

Similarly, the cost of keeping data object i$i$ at location $y$ by $S_{ij}^{[s]}$. Therefore, for minimizing the storage cost can be estimated using following Eq. 6 and 7

$$Q^{[s]} = \sum_{i \in I} \sum_{j \in J} 1_{ij} O_{ij}^{[s]}, \tag{6}$$

where

$$O_{iy}^{[s]} = S_{iy}^{[s]} \tag{7}$$

In LBCP both content information and its replicated content across node as replication (i.e., prefetching multiple copy of same content (PMCSC)) is considered. The content prefetching is very difficult when PMCSC of content information's are permitted. The PMCSC cost relies PMCSC's size and PMCSC content information location put forth. In this work, the PMCSC's of different content information is represented through parameter $z$. Then, the location of PMCSC for mapping content prefetching is determined through following Eq. 8

$$y : i \rightarrow \{j_1, j_2, \ldots, j_z\}. \tag{8}$$

Then, the content information $i$ for respective location of PMCSC with $i$, the routing strategy is mapped as follows Eq. 9

$$K : (i, a, j) \rightarrow j_z, \tag{9}$$

The Eq. (9) aid in providing content routing objective $jy \in J$ for different content information i$i$ in a configuration a$a$ from MANET gateway serverj$j$. The routing objective relies on information of both $Y$ and K$K$ in doing PMCSC. The content routing must be done in accordance with respective PMCSC prefetching. After finishing of content routing strategy, the content prefetching gained priory might not be ideal. Thus, content routing with PMCSC is very problematic. For addressing, this work presents an low query latency, energy efficient with better bandwidth utilization with better load balanced content prefetching model considering inter domain mobility of MANET user for mobile adhoc network. The proposed LBCP model is collected of three phases.

First, for solving initial content prefetching on multiple MANET nodes greedy algorithm is employed. Using greedy method for each content $i \in J$, LBCP model collects $M_i$ as described below Eq. 10

$$M_i = \left\{ G_{ij} | j \in J \right\} \tag{10}$$

The $M_i$ defines the bandwidth request for $i$ from each MANETs locations, and arrange it with high request rat to low request rate. The PMCSC is set to $z$ for different content and MANET gateway server max request rate within $M_i$ is chosen for storing PMCSC of respective information $i$. The initial content prefetching on multiple node help for assurance of reduction of cumulated traffic load. Our placement strategy is much more superior than existing random-based placement strategy.

Second, is to find an effective routing strategy to cater request configuration of each MANET gateway server with multiple prefetched copies. Every device will have different bandwidth request rate $G_{aj}$ associated with request configuration $a$; in this work we present a multi-objective strategy for reducing cost (i.e., bandwidth usage) of MANETs device to find ideal content routing model considering present PMCSC location of gateway server datacenter $v$ using following Eq. 11

$$\begin{gathered} \min \delta^e \sum_{j \in J} J_j + \sum_{i \in a} \sum_{j \in J} J_{ij} \delta_{xj}^{Q} \\ such\ that \sum_{j \in J} J_{ij}.1\left(i \in Y_j\right) = 1, \forall i \in a \\ J_{ij} \leq J_j, \forall i \in a, \forall j \in J \\ J_{ij} \in \{0, 1\}, \forall i \in a, \forall j \in J \\ J_j \in \{0.1\}, \forall j \in J \end{gathered} \tag{11}$$

where $\delta^e$ is a constant $M.(\alpha, \varphi, 0,)^u$ under the present content prefetching and $\delta_{xj}^{Q}$ is also a constant $M.\left(\delta_{xj}^{[u]}, \delta_{xj}^{[Q]}, 0\right)^u$. The optimal solution of Eq. (11) assures the minimization of multi-objective function defined in Eq. (3) considering different PMCSC prefetching location.

In stage 3, using improved request rate with respect to multiple pre fetched copies, content prefetching on multiple node solution is performed in the space of multiple pre fetched copies considering geographically distributed data using Bi-Graph partitioning. The algorithm of Load balanced content prefetching (LBCP) technique for MANETs is shown in Algorithm 1. In stage (3), we consider the content prefetching of multiple pre fetched copies in the space of multiple pre fetched copies in accordance to improved request rates with respect to multiple pre fetched copies. The second and third stage are executed in iterative manner till desired performance i.e., below threshold parameter $\tau$. This iterative process assures to satisfy any given request to completed with less latencies which is very key factor for future MANET based application services. In next section the performance evaluation of proposed method over existing method is presented.

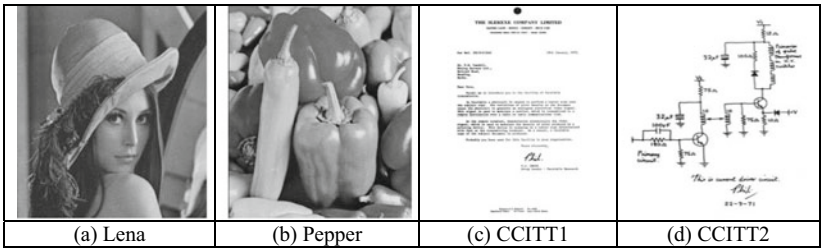
## 4 Result Analysis

The proposed Load balanced content prefetching (LBCP) technique for MANETs is modelled using MATLAB. The experimental evaluation is done under windows OS,

Intel I7, and 8 GB memory. The LTE uplink simulator [7, 8] is used for validating content LBCP and existing prefetching models [9, 13]. Cache consistence's and query access time is the metrics used for validating models. We used standard multimedia image such as Lena, pepper and CCITT1 and CCITT2 image as shown in Fig. 2.

## *4.1 Query Access Time Performance*

Figure 3, shows query access time performance evaluation of LBCP model with respect with standard-CP model with respect to different MANET device size. The MANET device are varied from 50 to 200. From graphical representation outcomes, LBCP minimizes content access time from cache by 29.29, 29.09, and 37.536% with respect to standard CP model for 50, 100, and 200 MANET device, respectively. An average content access time reduction of 31.97% is attained by LBCP over CP model.



(a) Lena (b) Pepper (c) CCITT1 (d) CCITT2

**Fig. 2** Standard image used for carrying of content prefetching

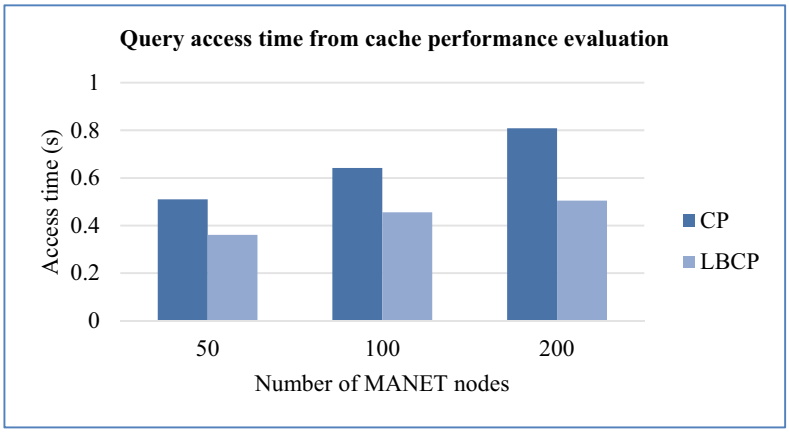



**Fig. 3** Query access time from cache performance evaluation considering varied MANET nodes

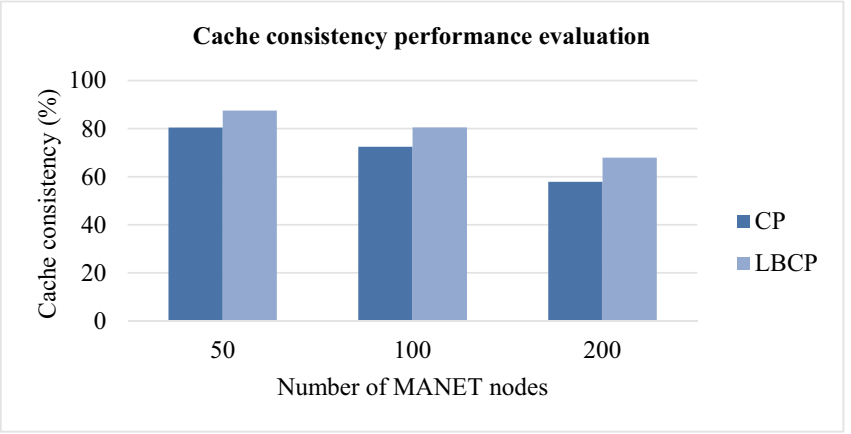



**Fig. 4** Cache consistency performance evaluation considering varied MANET nodes

## *4.2 Cache Consistency Performance*

Figure 4, shows cache consistency performance evaluation of LBCP model with respect with standard-CP model with respect to different MANET device size. The MANET device are varied from 50 to 200. From graphical representation outcomes, LBCP improves cache consistency performance by 8.011%, 10.058%, and 14.81% with respect to standard CP model for 50, 100, and 200 MANET device, respectively. An average cache consistency performance of 10.96% is attained by LBCP over CP model.

## *4.3 Energy Consumption Performance*

Figure 5, shows energy consumption performance evaluation of LBCP model with respect with standard-CP model for different MANET device size. The MANET device are varied from 50 to 200. From graphical representation outcomes, LBCP improves energy consumption performance by 4.01, 18.06, and 19.74% with respect to standard CP model for 50, 100, and 200 MANET device, respectively. An average energy consumption reduction of 31.97% is attained by LBCP over CP model.

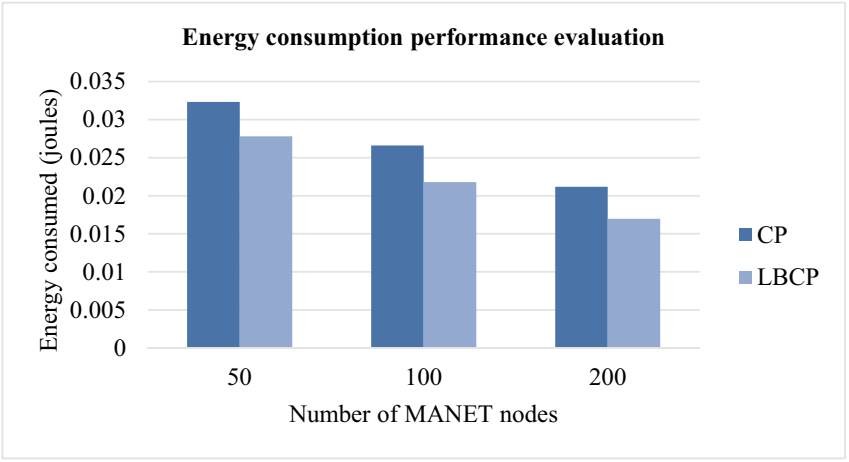



**Fig. 5** Energy consumption performance evaluation considering varied MANET nodes

**Table 1** Comparative analysis of throughput performance achieved using LBCP over existing prefetching models

| Speed (m/s) | LBCP [Proposed] | TA-AOMDV [17] 2021 | QMR [19] 2021 | QSMVM [18] 2020 |
|---|---|---|---|---|
| 10 | 43.8 | 36.9 | 19.7 | 10.5 |
| 20 | 44.56 | 34.7 | 19.7 | 10.5 |
| 30 | 41.9 | 27.3 | 19.8 | 10.5 |
| 40 | 39.56 | 19.3 | 19.8 | 10.6 |
| 50 | 38.56 | 33.9 | 19.8 | 10.6 |

## *4.4 Comparative Analysis*

This section presents performance analysis of proposed LBCP model with respect to various existing models. In [13] presented a prefetching model namely PrefCache for high speed MANET environment [14, 16, 17] which consider user preference and congestion overhead [15] for performing caching. The model attained a maximum hit ratio of 52.4%. Similarly, [9] achieved a maximum hit ratio of 80.4%; on the other side the proposed LBCP model achieved a maximum hit rate of 87.46%. Similarly, in Table 1, here we evaluate the throughput performance achieved using proposed LBCP over existing models such as TA-AOMDV (Topological change Adaptive Ad hoc On-demand Multipath Distance Vector) [17] model, QSMVM (QoS-Aware and Social-Aware Multimeric Routing) model [18], and QMR (QoS aware weight based on demand Multipath Routing) [19] model. From result achieved in Table 1, LBCP outperforms existing method TA-AOMDV [17], QSMVM [18], and QMR [19] in

terms of throughput efficiency considering varied speed. Thus, LBCP can be adopted in highly dynamic environment such as VANET.

## 5 Discussion and Conclusion

Content prefetching technique aid in improving performance of adhoc network such as MANETs and VANETs. Unlike, VANET, the MANET imposes certain battery constraint. Further, similar to VANET, the nodes in MANET are mobile in nature (i.e., the position/location of nodes keeps changing with respect time). These characteristic of MANETs makes content prefetching a challenging thing. In MANETs caching and prefetching technique can be used to upgrade the system performance in MANETs. Further, for provisioning multimedia service for healthcare sector through MANET-Cloud it is important to provide fault tolerance. Existing model always route packet through shortest or best path leading to bottleneck and degrades performance. These problems are addressed in load balanced content prefetching for MANET-Cloud. The Load balanced content prefetching model improves cache consistency performances with less minimal query processing time and high energy efficiency by better balancing load through prefetching the data to multiple node. Further, LBCP achieves much better throughput than recent work such as TA-AOMDV, QSMVM, and QMR. Thus, from overall result attained it can be seen the proposed content prefetching model is robust and scalable considering MANET-Cloud environment. Future work would further consider content prefetching considering minimizing energy efficiency and maintain reliability performance requirement.

## References

1. Abdelatif S, Derdour M, Ghoualmi-Zine N, Marzak B (2020) VANET: a novel service for predicting and disseminating vehicle traffic information. Int J Commun Syst 33(6):e4288. https://doi.org/10.1002/dac.4288
2. Pande SK, Panda SK, Das S, Alazab M, Sahoo KS, Luhach AK, Nayyar A (2020) A smart cloud service management algorithm for vehicular clouds. IEEE Trans Intell Transp Syst 22(8):5329–5340. https://doi.org/10.1109/TITS.2020.3021075
3. Parajuli N, Alsadoon A, Prasad PWC, Ali RS, Alsadoon OH (2020) A recent review and a taxonomy for multimedia application in Mobile cloud computing based energy efficient transmission. Multimedia Tools Appl 79(41):31567–31594
4. Vigneri L, Spyropoulos T, Barakat C (2019) Low cost video streaming through mobile edge caching: modelling and optimization. IEEE Trans Mob Comput 18(6):1302–1315
5. Ainsworth S, Jones TM (2018) An event-triggered programmable prefetcher for irregular workloads. In: Proceedings of the twenty-third international conference on architectural support for programming languages and operating systems, Williamsburg, VA, USA, 24–28 March 2018
6. Ji K, Ling M, Shi L, Pan J (2018) An Analytical cache performance evaluation framework for embedded out-of-order processors using software characteristics. ACM Trans Embed Comput Syst (TECS) 17(4):1–25

7. Ramalho L, Fonseca MN, Klautau A, Lu C, Berg M, Trojer E, Höst S (2016) An LPC-based fronthaul compression scheme. IEEE Commun Lett 21(2):318–321
8. Ramalho L, Freire I, Lu C, Berg M, Klautau A (2019) Improved LPC-based fronthaul compression with high rate adaptation resolution. IEEE Commun Lett 22(3):458–461
9. Zhao Z, Guardalben L, Karimzadeh M, Silva J, Braun T, Sargento S (2020) Mobility prediction-assisted over-the-top edge prefetching for hierarchical VANETs. IEEE J Sel Areas Commun 36(8):1786–1801
10. Beckmann N, Sanchez D (2018) Maximizing cache performance under uncertainty. In: Proceedings of the 23rd international symposium on high performance computer computer architecture, Austin, TX, February 2017, pp 109–120
11. Beckmann N, Sanchez D (2018) Modeling cache performance beyond LRU. In: Proceedings of the 22nd international symposium on high performance computer architecture (HPCA 2016), Barcelona, Spain, pp 225–236
12. Wan L, Cao Q, Wang F, Oral S (2019) Optimizing checkpoint data placement with guaranteed burst buffer endurance in large-scale hierarchical storage systems. J Parallel Distrib Comput (JPDC) 100:16–29. https://doi.org/10.1016/j.jpdc.2016.10.002
13. Guan Y, Zhang X, Guo Z (2021) PrefCache: edge cache admission with user preference learning for video content distribution. IEEE Trans Circuits Syst Video Technol 31(4):1618–1631. https://doi.org/10.1109/TCSVT.2020.3006388
14. Ali M, Anjum A, Rana O, Zamani AR, Balouek-Thomert D, Parashar M (2022) RES: real-time video stream analytics using edge enhanced clouds. IEEE Trans Cloud Comput 10(2):792–804. https://doi.org/10.1109/TCC.2020.2991748
15. Akhtar N, Khan MA, Ullah A, Javed MY (2019) Congestion avoidance for smart devices by caching Information in MANETS and IoT. IEEE Access 7:71459–71471. https://doi.org/10.1109/ACCESS.2019.2918990
16. Vegni AM, Souza C, Loscrí V, Hernández-Orallo E, Manzoni P (2020) Data transmissions using hub nodes in vehicular social networks. IEEE Trans Mob Comput 19(7):1570–1585. https://doi.org/10.1109/TMC.2019.2928803
17. Chen Z, Zhou W, Wu S, Cheng L (2020) An adaptive on-demand multipath routing protocol with QoS support for high-speed MANET. IEEE Access 8:44760–44773. https://doi.org/10.1109/ACCESS.2020.2978582
18. Jara EP, Mezher AM, Aguilar-Igartua M, Redondo R, Vilas AF (2021) QSMVM: QoS-aware and social-aware multimetric routing protocol for video-streaming services over MANETS. Sensors (Basel, Switzerland) 21:901
19. Priyambodo T, Wijayanto D, Gitakarma M (2020) Performance optimization of MANET networks through routing protocol analysis. Computers 10(2):2021. https://doi.org/10.3390/computers10010002

# Low Energy Reduction Technique via Memristor for Wireless Body Sensors


**K. Ramesh, S. Parasuraman, G. P. Ramesh, and P. Rachana**

**Abstract** Wireless Body Sensors (WBS) is the most supporting and detectable health care device, individuals can extent their life time as well as need to maintain with good health condition. In emerging trends, there is a growing interest in compulsion and necessary needs to tele-monitoring of bio signals and maintaining their healthy body condition. In future trend, it is necessary thing to implant the WBS in its own body and recording internal changes of health. The emerging application of WBS requires low energy consumption memory cell. The existing volatile type of Static Random Access Memory (SRAM) cell is used in the WBS device. The SRAM cell has high energy consumption during data storage and transition period. The simple battery is to supply the energy in the entire WBS system, the life cycle of battery is certain limited level of operation and it is difficult to replace WBS battery after implanted the system. Hence, the new memory element of memristor is embedded in the device is known as Memristive Random Access Memory (MRAM). The main property of memristor is a Non-Volatile Memory (NVM) through its own resistance. The device of memrsistor is capable of storing data and its own parameter of resistance as the features including small area, no leakage current and excellent energy efficiency as compared with existing SRAM cell. The main focus of this paper is to investigate different voltage level of energy consumption, and comparison between SRAM cells with MRAM cell.

**Keywords** Doped region · Injected flux · Ionic transport · Memristive random access memory · Non-volatile memory · Wireless body sensors

K. Ramesh (✉) · S. Parasuraman
Karpaga Vinayaga Engineering and Technology, Chennai, India
e-mail: nishusishu@gmail.com

G. P. Ramesh
St. Peter's Institute of Higher Education and Research, Chennai, India
e-mail: gpramesh@stpetersuniversity.com

P. Rachana
New Horizon College of Engineering, Bangalore, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_54

## 1 Introduction

The rapidly growing population (aged 60 years) up to 962 million in 2017, more than double as compared with the year 1980 and approximately 382 million old aged peoples living in worldwide. There is a probability that it could reach an approximation of 2.1 billion older-aged populations at the end of 2050 [1]. The 2/3 of world population facing chronicle diseases such as diabetes mellitus, hypertension, cancer and intruding bio virus. The emerging device is available in health care monitoring via WBS. It is possible to measure promising psychological/bio-medical storage data via remotely available base station. The next decade, growing interest of people requisite to tele-monitoring of bio signals by independently. Hence, the most necessary thing to implant the WBS in human body and recording internal changes of their health condition. However, the simple battery is used to supply the energy in the entire WBS device. The well-known fact, the battery life is certain limited cycle of operation. The WBS battery is drained, redeploying and replacing process of the battery can be expensive in terms of time and money [2]. The WBS device contains existing conventional memory cell of SRAM and DRAM. These memory cells are unable to keep up with low energy operation in WBS device. In this paper focus novel idea, and addressed by emerging new element like memristor. The physical structure of memristor is based on MRAM and consumes low energy as compared with SRAM cell used in WBS device. The resistor, capacitor and inductor all these phenomena are built in single element known as memristor in [3–5] and has presented the IV characteristics of memristor. The energy consumption, transition delay and data storing management of the existing WBS are the higher rate end. Memristors utilized as storage element with N-type MOSFET for row access. Moreover, achieving READ and WRITE operation for Look–up table (LUT) and thus saving 25% of total energy [6]. In this paper, these issues have been addressed via memristor, empowering the NVM and low energy consumption. The memory element of memristor is embedded with WBS for overcomes high energy consumption. The following sections of the paper are organized as follows. In Sect. 2 describes the related works. Section 3 comprises of low energy reduction method of SRAM and MRAM. In Sect. 4 discusses validation results and discussions of low energy consumption. Finally, we conclude this work in Sect. 5.

## 2 Related Works

SRAM cell performance analysis is demonstrated in the 6 T-SRAM cell and implemented in 90 nm technology [7]. Memory cells like Random Access Memory and Resistive Random Access Memory have proposed and estimated NVM technology described in [8]. The authors of [9] have presented the energy consumption in the WSNs through the Dynamic Power Management (DPM) technique. Memristive switching characteristics described in [10], applied for biomaterial solid polymer

electrolyte (SPE) chitosan. The biologically plausible models were presented in [11], implementation of memristive computing technologies in neuromorphic wearable sensors. The authors in [12] described memristor based scroll-controllable hyper-chaotic system adapted in operational amplifiers. Hardware security applications based on memristor physical unclonable functions (PUFs) and Random Number Generators (RNGs) is presented in [13]. The authors in [14] described bridge rectifier in terms of multi-dimensional polynomials of split signals with memristors. The different applications based on memristors implemented in the artificial pathway of synapses in neuromorphic networks applied in terms of Short Term Memory (STM) to Long Term Memory (LTM) are described in the authors [15–19].

## 3 Methodology

In emerging memory technology requires the most important characterizes are low energy consumption. The memristor memory chip, i.e. MRAM cell stores under non-volatile memory technology. There is no loss of data when power becomes off and saves its previous state after power removed from the source.

### *3.1 Energy Dissipation of SRAM Cell*

The conventional 6 T SRAM cell structure is proposed that the energy spends in the SRAM cell includes two state like dynamic and static energy consumption. Dynamic energy consumption of SRAM cell, charging and discharging of capacitors in terms of DC supply during the period of read/write operation. The static energy consumption caused due to current leakage in SRAM. Dynamic energy spends during read/write operation can be obtained from the following derivation. Where, the bit line voltage $V_{bl}$ and $V_d$ are the main supply voltages. The interval time of δt, during transistor cell is in only turn-on period. Where, $I_{st}$ is the current at saturation region of accessing transistor cell as given in the Eq. (1).

$$E = V_d^2(C_{wl} + C_{dl} - C_{bl}) - \frac{1}{2}V_{bl}^2(C_{bl} + C_{dl}) + \left(2^{column} - 1\right) \times \left[C_{bl}V_d^2 - \frac{1}{2}\frac{(C_{bl}V_d - I_{st}\delta t)^2}{C_{bl}}\right] \tag{1}$$

where, $\mu_n$ is the electron mobility and $C_g$ is the gate capacitance per unit area, W/L is the ratio between width and length of the transistor. $V_{gs}$ is the gate to source voltage of the transistor cell, and $V_{th}$ is the knee voltage of the access transistor. There are three main capacitances in the SRAM that includes $C_{bl}$, $C_{wl}$, and $C_{dl}$. Where, $C_{bl}$ is the bit-line capacitance composed from the drain junction capacitance $C_{dj}$ of access transistor and metal capacitance $C_m$ of bit-line. The row is the number of address line in row decoder, and bit-line capacitance are defined in the Eqs. (2), (3), (4) and

(5). The metal capacitance of bit line is assumed to be 10% of the metal capacitance of bit-line obtained from Eq. (6). The word-line capacitance $C_{wl}$ is composed of gate capacitance of access transistor of the memory cell.

$$I_{st} = \frac{\mu_n C_g W}{2L}\left(V_{gs} - V_{th}\right)^2 \tag{2}$$

$$C_{bl} = C_{dj} \times 2^{Row} + C_m \tag{3}$$

$$C_m = C_{dj} \times 2^{Row} \times 10\% \tag{4}$$

$$C_{wl} = 2 \times C_g \times 2^{Column} \tag{5}$$

$$C_{dl} = 2 \times C_{jc} \times 2^{Column} \tag{6}$$

The column is the number of address line in column decoder, and the large capacitance of SRAM is data-line capacitance $C_{dl}$ and this capacitance has composed from junction capacitance of column access transistor $C_{jc}$.

### *3.2 Memory Element of Memristor*

A memristor is a non-volatile two terminal nanoscale memory element and having dynamic resistance. The feature of memristor includes small area, low leakage current and excellent energy efficiency. The memristor possesses the ability to retain a resistance value even after the power source is removed from the device. The functional relationship of voltage V(t) and current I(t) are described in Eq. (7). The resistance $R_{on}$ and $R_{off}$ represents the minimum and maximum memristance of the element, respectively. The actual resistance of the device is dependent on the ratio between the thickness of conductive doped region $w$(t), and thickness (D). The value of $w$(t) increases with leads to lowering the memresistance value ($R_{off} > R_{on}$), as stated in Eq. (8).

$$V(t) = \left[R_{on}\frac{w(t)}{D} + R_{off}\left(1 - \frac{w(t)}{D}\right)\right] I(t) \tag{7}$$

$$w(t) = \frac{\mu_v R_{on}}{D} q(t) \tag{8}$$

where, $\mu_v$ is the mobility of dopants in thin film is equal to the value of $10^{-14}$ $m^2 s^{-1}$ $V^{-1}$. The value of $w$(t) is calculated by integrating in Eq. (8), and the result is proportional to the charge of the memristor. The charge is integral part of the current

and charge also constant in the absence of current. Same manner the resistance remains unchanged; this phenomenon makes the non-volatile effect of the memristor.

### *3.3 Memristive Random Access Memory*

The characteristics of memristor is going to off-state to on-state and storing data values via MRAM. The diagram MRAM cell consists of one MOS transistor T and one memristor R required for one memory cell as shown in the Fig. 1. During read/write operation the row and column cells are selected through Word Line (WL) and Bit Line Bar (BLB). Memristor (M) has memory cell is connected to the WL and the BLB and made interface with other cells. The voltage $V_d$ is applied to the MRAM word line path. The voltage drop across the memristor is achieved by charging the bitline to $V_d$. The data access time depends on voltage drop across the memristor. The two different levels of inputs of threshold voltage $V_{th}$ and the bit-line voltage $V_{bl}$ are applied to sense amplifier. In particular, smaller $R_{on}$ is chosen in the 1T1R which decreases the sense amplifier gain. The ($R_{off}$ / $R_{on}$) ratio of memristor leads to higher and maintains reliabilities and lead to low power consumption. During read/write operation, the discharging of bitline to zero volts is the additional source of energy saving as against the recharging operation. The selected wordline and bitline in read and write mode that activated for all the cells in the row. The CMOS access transistor is used to isolate the unselected cells. The resistance and capacitance value of bitline is calculated to be 6 KΩ and 150 f F for 1 mm bitline length.

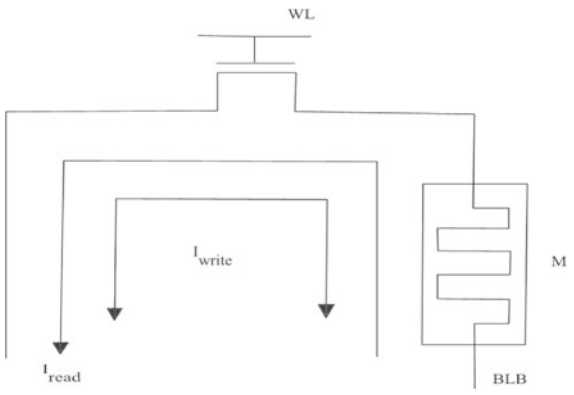



**Fig. 1** MRAM memory cell (1T1R)

### 3.4 Energy Dissipation of MRAM Cell

We drive the explicit expressions for low energy reduction of MRAM. The state of the memristor is to change to a feasible state with applied voltage source. Where, $V_A$ is the amplitude of voltage source and required time width T is stated in Eq. (9).

$$T = \frac{D^2(1+\beta)}{2\mu_v V_A} \tag{9}$$

The ratio between ($R_{off}/R_{on}$) represents as β. The parameters of the device increases with increase in β and D. As well as decreases with increase in applied voltage due to increasing injected flux. The nonlinear ion drift take place in the memristor due to massive electric field. The ionic transport speed slowly decreases to zero between doped to undoped regions of boundary line. The control parameter P is a positive integer then the existing window function provides a harder boundary effect and it gets stuck state. An additional parameter of current I added with the non-linear window function in Eq. (10). Memristor doped layer width increases, assume that current I in the expression of f'(x) is positive in Eq. (11). The proposed window function of f'(x) is helpful for energy model of MRAM cell. The function f'(x) is zero at both the edges. The instantaneous current through the memristor carries the state information of logic 1 or 0. The resistor in series which depends on instance of time is determined in Eq. (12).

$$f(x) = 1 - [2x(t) - 1]^{2P} \tag{10}$$

$$f'(x) = 1 - [x(t) - stp(I(t))]^{2P} \tag{11}$$

$$i(t) = \frac{V_{\mathrm{d}}}{2\left(\mathrm{R_{ch}} + R_{tg} + R_{bl} + R_m(t)\right)} \tag{12}$$

$$E = \int_0^{\mathrm{T}} \frac{V_{\mathrm{d}}}{2} i(t)\mathrm{dt} \tag{13}$$

where, $R_m$ is the equivalent resistance of the memristor. During read/write mode, $R_{ch}$ is the channel resistance of triode region access transistor. Transmission gate switch resistance $R_{tg}$ is used to pre-charging and discharging the bitline capacitor. Where, $R_{bl}$ is the bitline resistor. The low power energy of E is dissipated in the MRAM cell with respect of dc analysis is derived in Eq. (13). Different window function of β values in the MRAM cell can be calculated via MATLAB tool.

## 4 Results and Discussions

To estimate the precision of our proposed analytical framework, testbed experiments and simulations are conducted. The experiment domain is a computer with an Intel (R) 9300H CPU working at 2.40 GHz and 8 GB RAM, 512 GB SSD and 4 GB. The conventional SRAM cell consists of 6 T (transistor) structure and MRAM device used in 1T1R, 1 MOS transistor and 1 memristor serve as an access switch for read/write operation. The derived model of both cells is executed by MATLAB tool and simulation is conducted through HSPICE software. HSPICE simulations in standard of 250 nm CMOS technology confirms accuracy for performance of analytical expressions derived from this paper. The given dc supply of 0.5 V for NMOS and negative 0.5 V for PMOS threshold de voltage of MOS transistors. Summarized the parameters list of both SRAM and MRAM cells indicated in the Table 1.

To keep the bitline dc voltage at 0.2 V at all the times since the target is to achieve low power consumption. Hence, implementations of both read and write mode, precharge-free pulling operation and one-side driving operation in respectively. We assume that memristor resistance ratio of $\beta = 800$ for MRAM cell model. The memristor has the potential to enhance Non-Volatile MRAM cell structure. Highly pervasive area with low power consumption is given in the Table 2. The proposed model obtained better energy consumption performance compared to the

**Table 1** Simulation parameters

| Parameters | Specification values |
|---|---|
| Time interval, $\delta t$ | 0.105–0.120 ns |
| Gate capacitor, $C_g$ | $4 \times 10^{-4}$pF/$\mu m^2$ |
| Drain junction, $C_{dj}$ | 2.8 pF/$\mu m^2$ |
| Junction capacitor, $C_{jc}$ | 1.9 pF/$\mu m^2$ |
| Oxide capacitor, $C_{ox}$ | 6 pF/$\mu m^2$ |
| Gate switch resistance, $R_{tg}$, | 500 Ω |
| Memristor thickness, D | 10 nm |
| Minimum Resistance, $R_{on}$ | 10 Ω |
| Amplitude voltage, $V_A$ | 1 V |

**Table 2** Power dissipation

| Method | Memory cell | Power (μW) |
|---|---|---|
| **Proposed** | **SRAM** | **1.4571** |
| | **MRAM** | **0.0678** |
| | SRAM | 1.8979 |
| Marković et al. [4] | MRAM | 0.1657 |
| | SRAM | 1.7291 |
| Min and Cho [10] | MRAM | 0.2190 |

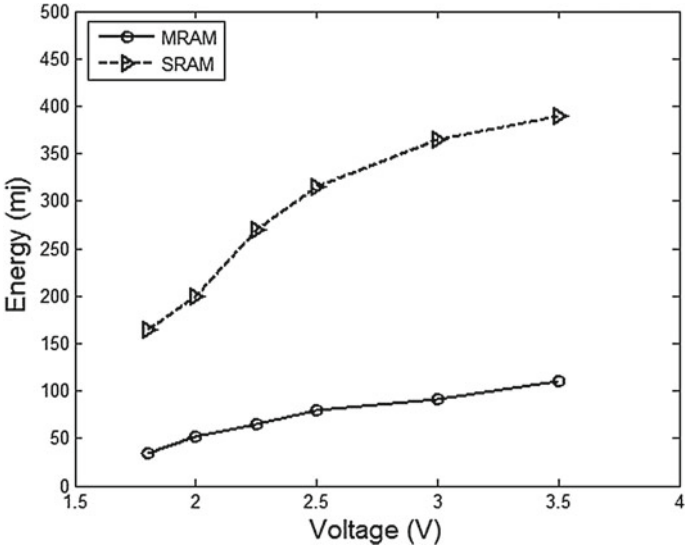



**Fig. 2** Average energy consumption

existing models developed by I. Marković et al. [4] and Min and Cho [10]. The average energy consumption of SRAM cell versus MRAM cell as shown in the Fig. 2. Average energy consumption increases with the increasing voltages of SRAM cell.

The applied dc voltage of 2.5 V that the average energy consumption is 80 mj for memristor MRAM cell whereas 325 mj for SRAM cell. Evident from the results that the MRAM cell consumes four times lesser energy than SRAM. The peak energy consumption of SRAM versus MRAM cells as shown in the Fig. 3. The line of

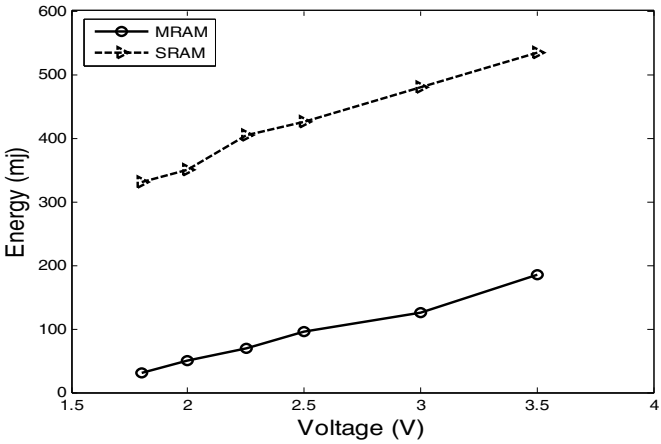



**Fig. 3** Peak energy consumption

SRAM energy consumption increases for the rising in the applied external voltage. The result shows that the applied dc voltage of 2.5 V and hence the peak energy consumption is nearly 100 mj for memristor MRAM cell and 410 mj for SRAM cell.

## 5 Conclusion and Future Work

The main highlight of this paper, we systematically derived a set of comprehensive equations. Characterizing the SRAM cell and MRAM cell device and we investigated the energy consumption of both memory devices. We proved the results that MRAM cell has lesser energy consumption than SRAM cell. The target is achieved via this paper that it made on comparison of energy between SRAM and MRAM cells. Energy consumption of MRAM cell is four times lesser than the SRAM cell. The MRAM cell architecture has been extended to fabricate with the applications where there is a need for low energy consumption. Wireless Body Sensor dominated energy consumption and hence increasing communication costs for bio-signal transmission and lowering the battery life time. Further, we concluded that MRAM architecture suits better for the WBS for minimizing transmission costs and prolonged life of battery. The size of the memristor is extremely small and by applying mimicking the functions of a brain. Moreover, future enhancement work of memristor technology is used to apply neuro-biological (Neuromorphic) architectures present in the nervous system.

## References

1. Department of Economic and Social Affairs (2017) World survey of growing population from 1980–2050, United Nations, pp 1–40
2. Kasan S, Gaber J, Lorenz P (2019) A survey of energy efficiency in wireless human body sensors lifetime for healthcare applications. Telemed e-Health 2(1):1–5
3. Vaidyanathan S, Pham V-T, Tlelo-Cuautle E, Kapitaniak T (2019) Study memory circuit element of Memristor with application oriented. Wiley Hindawi Complexity, pp 1–4
4. Potrebic M, Markovi I, Tosic D (2021) Memristors as candidates for replacing digital potentiometers in electric circuits. Electronics 10(181):1–18
5. Apollos E (2019) Memristor Theory and Mathematical Modelling. Int J Comput Appl 178(27):1–8
6. Kumar TN, Almurib AF, Lombardi F (2016) Design and evaluation of a memristor-based look-up table for non-volatile field programmable gate arrays. IET Cric Dev Syst 10(4):292–300
7. Shivaprakash G, Suresh DS (2016) Design of low power 6T SRAM cell and analysis for high speed application. Indian J Sci Technol 9(46):1–10
8. Budhathoki RK, Parajuli S, Kim H (2020) Nonvolatile memory cell based on memristor. Univ J Electr Electron Eng 7(2):110–117
9. Pughat A, Sharma V (2016) Performance analysis of an improved dynamic power management model in wireless sensor node. Dig Commun Netw 3:19–29
10. Min SY, Cho WJ (2021) Memristive switching characteristics in biomaterial chitosan-based solid polymer electrolyte for artificial synapse. Int J Molec Sci 22(773):1–13

11. Donati E, Covil E, Payvand M, Heidari H, Liang X, Kappel D, Wang W (2020) Comparison of related work and implementation of memristor technology with neuromorphic sensors, pp 1–29. arXiv:2012.14937v1 [cs.ET]
12. Jin J, Cui L (2019) Fully integrated memristor and its application on the scroll-controllable hyperchaotic system. Complexity 2019:1–8
13. Geng Z, Lv S, Liu J (2021) Application of memristors in hardware security: a current state-of-the-art technology. Adv Intell Syst 3(2000127):1–21
14. Harchuk H, Solovyeva E, Schulze S (2021) Behavioral modelling of memristor-based rectifier bridge. Appl Sci 11(2908):1–15
15. Mahata C, Rahmani MK, Ismail M, Kim S (2020) Study the concept of $SnO_2$ – memristor for synaptic application. Results Phys 18:1–8
16. Jang HW, Kim SB, Kim SJ (2021) Competing memristors for brain-inspired computing. iScience 24(101889):1–21
17. Guo T, Sun B, Ranjan S, Zhou G et al (2020) Multistate resistive switching behaviours for neuromorphic computing in memristor. Mater Today Adv 9(2021):1–8
18. Kuang R, He YH, Pan WQ, Duan N et al (2020) Implementation on accuracy of memristor based neural networks. IEEE Trans Electron Dev 67:1–7. https://doi.org/10.1109/TED.2019.2963323
19. James A, Chua L (2021) Study the related work of memristor based crossbars applications, pp 1–13. arXiv:2105.04614v1 [cs.ET]

# Machine Translation for Indian Languages Utilizing Recurrent Neural Networks and Attention


**Sonali Sharma and Manoj Diwakar**

**Abstract** Neural Machine Translation better known as NMT is an end-to-end approach for autonomous language translation that utilizes neural models. This is an effort to bridge the gap between the multinational and multilingual people to understand their views. The NMT systems involves models to learn directly through mapping of input–output which has proven to generate increased accuracy outputs. This technique has made remarkable accomplishments and has overcome the weakness of the conventional translations models. The paper implements the RNN, Attention based mechanism and transformer on Indian-English language pairs. So far there are no specific benchmarks for Indian languages. There are companies such as Facebook, Bing, Google whose translators supports few Indian languages. In this research work models have been trained on two set of Indian language pairs which have been retrieved from open source platform Tatoeba.

**Keywords** Attention mechanism · Encoder-Decoder · Machine translation · Recurrent Neural Network (RNN) · Transformers

## 1 Introduction

India is a diverse nation of people belonging to different culture, religion, race, ethnicity and language as well. Across the country there are multilingual speaking people and so communication is necessary not only for trading purpose but also to bind people of the country together. Machine translation is the one of the task of natural language processing which facilitates automatic translation of language that is required to have effective communication among the multilingual people.

Machine translation task can be described as the translating the text or speech from source language to target language. The task is to achieve resultant language which is syntactically equal to source language and semantically equal to target. The work in machine translation dates back in early 1950s and since that time to current period

S. Sharma (✉) · M. Diwakar
Department of CSE, Graphic Era Deemed to Be University, Dehradun, India
e-mail: soniesharma227@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_55

machine translation has massively progressed because of various factors: new & improved techniques, the increased computational power, high memory availability and large monolingual, bilingual & multi-lingual text corpora available over internet [1].

Various kinds of approaches have been proposed for performing machine translation such as rule based translation, hybrid translation, knowledge based translation and corpus based translation. Each of the kinds has its own merits and demerits. Statistical machine translation (SMT) produced far much better results in comparison to previous techniques and requires much little human intervening [2]. Neural machine translation approach became more popular when deep learning architecture got incorporated and showed promising results.

A lot of work by research scholars has been done in machine translation but that is limited to English and European language. Not much attention has been given to the Indian languages. This work focuses on Indian Languages and so picked up two Indian languages Marathi-English, Hindi-English on the datasets obtained from http://www.manythings.org/anki/.

## 2 Related Work

In the last couple of decade neural machine translation displayed a great deal of improvement in translation task. In 1987 the translation from English to Spanish Language was conducted using backpropagation with the vocabulary being a limited constrained. Since the neural machine architecture has been subjected to multiple alterations. For translation RNN was considered to a deFacto approach [3]. It employed to encode the variable length sequence to fixed context understanding leading to loss of information. Lately GRUs, LSTMs came which tended to replace typical RNN cells and offer better translation quality and propagation.

The sequence to sequence model and its varied types such as RNN (Sutskever, Bahadanau and Chen) have been heavily adopted by many techno which gained massive success in global market. The model can be used to perform multiple task parallelly which is showcased by multilingual machine translation such as Google neural machine translation (GNMT) to translate between multilingual languages [4, 5]. The more recent model is Transformer by Vaswani et al. introduced in the paper "Attention is all you need" completely revolutionized the NLP field.

The machine translation task is sensitive to the quality and the quantity of the dataset applied. During research it was witnessed that massive work has been done by the researchers for European languages but little attention has been payed to the Indian languages. So, this paper, primarily focuses is on the Indian language pairs.

## 3 Neural Machine Translation Models

NMT is an end-to-end framework whose basic structure comprise two core components i.e. encoder and decoder. Encoder takes in the variable length sequence of input tokens, encodes them and captures the information in vector representation called as "context vector" which is served as the conditioning to the decoder (input) in the next step (see Fig. 1) [6].

The decoder is trained to predict the next token in the sequence given the previous tokens in the target language. It takes a bit-shifted version of vector representation from encoder and takes masked null values and produces the tokens in different language. In machine translation domain there is decorrelation in the length of the input and the length of the output and structures as well. The length of the output and structures as well [7].

NMT is a sequence to sequence model where $c = (c_1, c_2 \cdots\cdots c)$ is a source sequence and $e = (e_1, e_2 \cdots e_m)$ is a target sequence. NMT aims to maximize the conditional probability which is expressed in Eq. 1.

$$(e|c; \theta) = i = 1NP(eN|E < n\,C; \theta) \tag{1}$$

where $\theta$ is a model parameter set. The network includes an encoder which encodes the source sequence into a hidden states representation $h_c = h_1, \ h_{2\ldots\ldots\ldots}h_m$. The hidden representation is updated at each time step and the decoder produces the nth sequence as a resultant language [8].

### *3.1 Long Short Term Memory*

Long short term memory a variant of RNN is known to resolve the long range dependency issue and vanishing gradients which persists in the conventional recurrent neural networks. The RNN cells in the network are replaced by LSTM cells. On

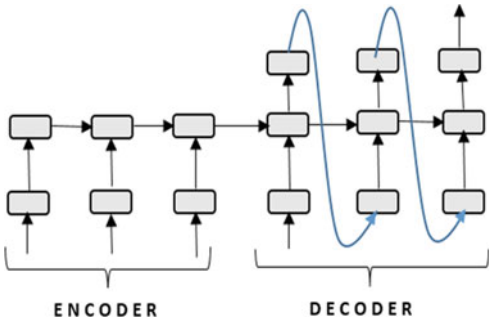



**Fig. 1** Encoder-Decoder framework

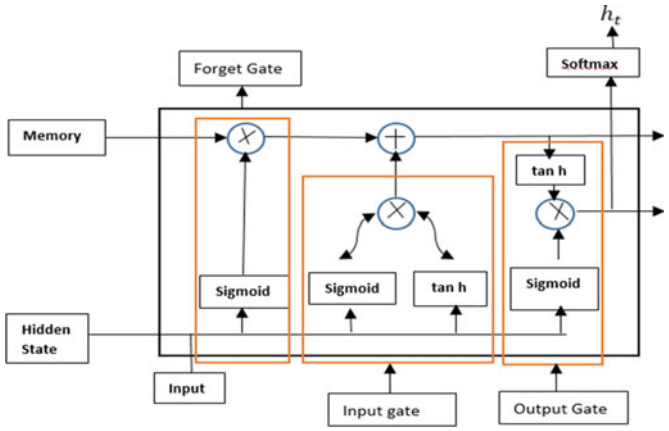



**Fig. 2** LSTM cell unit

the other hand, both Suskever et al. (2014) and Luong et al. (2015) stacked multiple layers of an RNN with a Long Short-Term Memory (LSTM) hidden unit for both the encoder and the decoder. These cell have the ability to retain the necessary information by regulating the gates in the structure. The below Fig. 2 shows the diagram of the basic LSTM cell.

Input Gate – Allows which input value should be utilized to change the memory which are expressed in Eq. 2 and 3.

$$i_t = \sigma\left(W_i.\left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

$$C_t = \tanh(W_c.\left[h_{t-1}, x_t\right] + b_c \tag{3}$$

Forget Gate – functions to discard the irrelevant details from the block is expressed in Eq. 4.

$$f_t = \sigma(W_f.\left[h_{t-1}, x_t\right] + b_f) \tag{4}$$

Output gate – Memory and input block both decides the output which are expressed in Eq. 5 and 6.

$$o_t = \sigma(W_f.\left[h_{t-1}, x_t\right] + b_o) \tag{5}$$

$$h_t = o_t * \tanh(c_t) \tag{6}$$

where,
$\sigma$ = sigmoid function

$h_{t-1}$ = previous state
$x_t$ = content input
$C_t$ = cell state
$f_t$ = forget gate
$W_i and W_f$ = Weight matrices

### 3.2 *Transformer Neural Machine Translation:*

Transformer a more recent developed architecture for machine translation which is composed of two major constructs: encoder and Decoder which can take the input sequence in one language and produces the output in another language see Fig. 3. Original encoder- decoder components are 6 layered where encoder blocks and decoder blocks at high level are stacked on the top of each other [9]. Each block consists feed-forward and self-attention layers but decoder consists an additional layer i.e. encoder-decoder attention layer whose purpose is to map relevant tokens to encoder. Self-attention layer looks for remaining input words at different positions to derive the relevance with the current word which is in translation process.

The output generated by the top layer of the encoder which is a set of attention vectors k and v is fed to the decoder as an input. These values are used by the encoder-decoder attention layer which allows the decoder to concentrate at the relevant places of the input sequences. The steps are repeated until the decoder confronts with the end symbol signalling the decoder finished its output [10].

Final Linear Layer: The function of the linear layer is to transform the word vector into a word [14]. This layer is simply a fully connected neural network that projects the vector produced by the decoder as logistics vectors.

The softmax layer transform the scores into probabilities and the word with the highest probability is chosen and then the lower ones are discarded. Positional encoding: It is use to represent the order of the sequence. The transformer adds a vector to each input embedding [11]. The vector follows a pattern that helps to determine the distance between different words and the position of each word which is mentioned in Eq. 7.

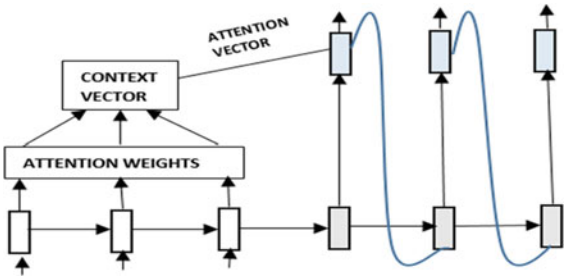



**Fig. 3** Attention Mechanism: attention vector is produced by consuming context vector [8]

$$attention(q, k, v) = softmax(qk^t/\frac{qk}{\sqrt{d^k}})v \tag{7}$$

In the encoder component each encoder block has sub layers which has residual connections around it, followed by layer normalization [12].

### *3.3 Attention Mechanism*

The drawback of the conventional model such recurrent neural network (RNN) are able to handle the shorter sentences very well but the are not able to handle the lengthy sentences and perform the translation well and therefore to overcome this drawback attention mechanism is proposed which while translation process focuses only sub part of the sentence rather that concentrating on the whole sentence [13]. The diagram below Fig. 3 illustrate the attention mechanism.

Looking from a high level of abstraction during decoding phase the encoder allows the decoder to look at all the hidden states $h_t$ at each time step. Architecture without a attention employed a context vector $c_t$ to be passed an input to decoder but now a concatenation layer is employed which combines the hidden state and context vector to generate attention hidden state [14]. The attention hidden state is given to softmax layer to generate output at particular time step.

## 4 Experiments

To perform the Machine translation task we applied the above described methods to two Indian language pairs [15]. The process involves the splitting of the data into the respective languages and then data is tokenized and all the irrelevant words also called the stop words are removed from the data. Then after that the data is split into validation and testing datasets. The preprocessor objective is to build the dictionary data structure. The sentence starts from the <start> token and ends with the <EOS> token indicating that these token are the start and the end of the sentence.

### *4.1 Dataset*

The data for both language pairs is obtained from Tatoeba Corpus the which contains the collection of sentence and its corresponding translation for specific language pair [16]. A tab-delimited bilingual sentence pairs for many language pairs. The number of unique input- output tokens are for bilingual pair is given in Table 1.

**Table 1** Number of input–output tokens

| Parameters | Marathi-English | Hindi-English |
|---|---|---|
| Unique input tokens | 71 | 70 |
| Unique output tokens | 85 | 88 |
| Length for inputs | 19 | 107 |
| Length for outputs | 42 | 123 |

**Marathi – English.** Total number of sentences present in the Marathi-English document are 41,028. The data is encoded in utf-8 format. The vocabulary size of English is 347 and vocabulary size of Telugu is 408. The data is applied to sequence to sequence and Transformer where transformer performed slightly better than previous model. [http://www.manythings.org/anki/].

**Hindi – English.** Total number of sentences present in the Hindi-English document are 2916. The data is encoded in utf-8 format. The vocabulary size of English is 2419 and vocabulary size of Hindi is 3062. A significant performance difference can be seen where transformer gave better accuracy than sequence to sequence. [http://www.manythings.org/anki/].

## 5 Results

The rapid developments have introduced many performance efficient frameworks and the once which we have used for analysis in our research work are Keras and Tensorflow. The Results are calculated and the visual representation of the analysis has been created using Seaborn and Matplotlib library.

### *5.1 Sample Output*

The below is the sample summary of the output produced after the implementation of the attention layer to the architecture for the Hindi and Marathi Language. The below is the graph for the Marathi-English language see Fig. 4.

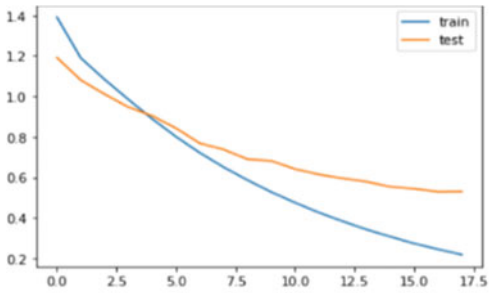



**Fig. 4** Train-Test Plot for Marathi-English language using attention mechanism

English – Marathi:

```
Review: how many rackets do you have
Original summary: तुझ्याकडे किती रॅकेट आहेत
Predicted summary:  तुमच्याकडे किती बहिणी आहेत
```

English-Hindi:

```
Review: dont speak french in the class
Original summary: वर्गात फ्रेंचमध्ये बोलू नका
Predicted summary:  वर्गात जास्त कठीण राहू नकोस
```

```
['democracy is the worst form of government except all the others that have been tried',
 'लोकतंत्र सरकार का सबसे घिनौना रूप है अगर बाकी सारी तरह की सरकारों को अंदेखा किया जाए तो।'],
['if my boy had not been killed in the traffic accident he would be a college student now'
 'अगर मेरा बेटा ट्रेफ़िक हादसे में नहीं मारा गया होता तो वह अभी कॉलेज जा रहा होता।'],
```

**Discussion.** The results obtained after implementation of the Indian bilingual datasets clearly signifies the incapability of Long short term memory to handle more complex sentences which the attention based model is capable of handling i.e. the attention based models when implied to both datasets show 91 and 92% of validation accuracy which is quite higher than the conventional model i.e. 77 and 73% of validation accuracy see (Table 2)**.** A comparative study is performed with similar methods such where it was found that in most of the cases proposed attention method gives better accuracy than LSTM, as shown in Table 3.

**Table 2** Shows results of seq2seq_LSTM and attention based models

| Language pair | Accuracy | Validation_loss | Validation_accuracy |
|---|---|---|---|
| Marathi-English (seq2seq) | 0.9641 | 1.147 | 0.777 |
| Telugu-English (seq2seq) | 0.9787 | 1.333 | 0.735 |
| Marathi-English (attention) | 0.9789 | 0.5303 | 0.9127 |
| Telugu-English (attention) | 0.9812 | 0.456 | 0.9203 |

**Table 3** Comparative results of Indian language pair [17]

| Language pair | Accuracy | Validation_accuracy |
|---|---|---|
| [10] | 0.9641 | 0.8774 |
| [11] | 0.9637 | 0.8756 |
| Proposed attention | 0.9882 | 0.9405 |

## 6 Conclusions

The paper implements the long short term memory (LSTM) transformer and attention mechanism approach on Indian language pairs. To our surprise the models performed quite well even with the limited amount of data publically available. We conclude that the attention and transformer based methods perform better. The conventional methods secure low accuracy scores. The validation accuracy and accuracy for both languages is much lower that what is secured by the attention-transformer based methods**.**

For the future the machine translation task can be more fine-tuned by using pre-trained models such as Bidirectional encoder representation transformer (BERT), Albert. Using the above mentioned approach accuracy levels for low resourced and zero resourced language can be improved further.

## References

1. Stahlberg F (2020) Neural machine translation: a review. J Artif Intell Res 69:343–418
2. Vaswani A, Bengio S, Brevdo E, Chollet F, Gomez AN, Gouws S, Jones L, et al (2018) Tensor2tensor for neural machine translation. arXiv preprint arXiv:1803.07416
3. Rikters M, Pinnis M, Krišlauks R (2018) Training and adapting multilingual nmt for less-resourced and morphologically rich languages. In: Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018)
4. Zhang J, Zong C (2020) Neural machine translation: challenges, progress and future. Sci China Technol Sci 63:1–23
5. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473
6. Cho K, Merriënboer BV, Bahdanau D, Bengio Y (2014) On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259
7. Wu Y, Schuster M, Chen Z, Le QV, Norouzi M, Macherey W, Krikun M, et al (2016) Google's neural machine translation system: bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144
8. Ortega JE, Mamani RC, Cho K (2020) Neural machine translation with a polysynthetic low resource language. Mach Transl 34(4):325–346
9. Datta D, David PE, Mittal D, Jain A (2020) Neural machine translation using recurrent neural network. Int J Eng Adv Technol 9(4):1395–1400
10. Revanuru K, Turlapaty K, Rao S (2017) Neural machine translation of Indian languages. In: Proceedings of the 10th Annual ACM India Compute Conference, pp 11–20
11. Pathak A, Pakray P (2019) Neural machine translation for Indian languages. J Intell Syst 28(3):465–477

12. Luong MT, Pham H, Manning CD (2015) Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025.
13. Pal S, Zampieri M (2020) Neural machine translation for similar languages: the case of Indo-Aryan languages. In: Proceedings of the fifth conference on machine translation, pp 424–429
14. Singh S, Anand Kumar M, Soman KP (2018) Attention based English to Punjabi neural machine translation. J Intell Fuzzy Syst 34(3):1551–1559
15. Jain M, Punia R, Hooda I (2020) Neural machine translation for Tamil to English. J Stat Manag Syst 23(7):1251–1264
16. Srivastava S, Tiwari R (2019) Self-attention based end-to-end Hindi-English Neural Machine Translation. arXiv preprint arXiv:1909.09779
17. Laskar SR, Pakray P, Bandyopadhyay S (2019) Neural machine translation: Hindi-Nepali. In: Proceedings of the fourth conference on machine translation, vol 3: Shared Task Papers, Day 2, August 2019, pp 202–207

# Malaria Detection from Blood Cell Images Using Convolutional Neural Network Model


**Harsha Tiwari and Avinash Dhole**

**Abstract** Malaria disease is the major and worldwide reason for death in the human community. To identify the malaria, sample of red blood cell is examined under the microscopic instrument by the technician. But, due to the varying structural nature of malaria parasite in the blood cell; relying on the single method may endanger the diagnosis of this disease. With the advancement of advanced technique including machine and deep learning mechanism, this solution can be enhanced to verify the severity of this diseases and make them to properly diagnose the infected case. In this context, an artificial intelligence based convolutional neural network model has been developed to categorize the blood cell into malaria infected and uninfected type of classes. Performance of the model has been evaluated using standard evaluation measures on the public dataset of malaria blood cell images. Obtained quantitative scores confirms its performance which outperforms the available state-of-the-art techniques.

**Keywords** Computer-added diagnosis · Convolutional neural network · Deep learning · Medical imaging · Malaria parasite · Red blood cell examination

## 1 Introduction

Malaria is one of the severe and deadly disease produced by the harmful parasites which are transferred to individuals from the infected bites of female anopheles' mosquitoes. Despite of its features, it is preventable and can be operated to cure. From the report of world health organization (WHO) 2019, there are 4,09,000 number of estimated deaths worldwide from this malaria disease [1]. However, early treatment and diagnosis of malaria assist the doctor to save the human life. The report reveals that the children who are aged below 5 years are the most susceptible elements affected by malaria and it has been observed that 67% (2,74,000) of all children deaths

H. Tiwari (✉) · A. Dhole
Department of Computer Science and Engineering, Raipur Institute of Technology, Mandir Hasaud, Chhatouna, Chhattisgarh 492001, India
e-mail: harshatiwari.com@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_56

in 2019 worldwide are marked. The primary symptom of malaria is high fever, excessive sweating which generally occurs during the suffering. Observing the statistics of malaria, it is one of the lethal disease in the world, and responsible for the child death on the planet. This ailment keeps the children from basic schooling, taking a colossal toll on livelihoods and overall progress of country. Medical practitioner generally identifies the source of malaria through the blood cell examination in which presence of malaria parasite affects the red blood cell in an unusual manner. Malaria parasites usually recognized by investigative approach under a microscopic instrument where the sample of the patient's blood is used and called as a "blood smear". WHO mentions that the cases of so-called malaria be definite using parasite-based analytic testing before directing the medical treatment. Outcomes of such investigation verification can be presented in $\pm 30$ min. Further they identify a blood cell into two categories: normal and malaria infected blood cell. But, the structure of parasite in blood cell are not uniform and have varied representation. An affected but unidentified malaria parasite in blood cell examination may lead to the death. Nevertheless, chances of manual mistake may lead to unexpected treatment due to this short duration of investigation. Therefore, using intelligent technique a precise detection of such parasite is required to avoid any chances of human death. Nowadays medical field is also being facilitated with the advancement of several intelligent techniques, which assists the precise decision making of doctors and physician. In this domain, several image processing and optimization approaches have been applied to understand the different data patterns in several applications [2–7], but their performance has been experimented against the standard conditions. In this direction, recent development in machine and deep learning techniques have leveraged the performance of the artificial intelligence (AI) based systems [8–13] and [14–18]. AI techniques has the prospective to benefit some of the most challenging issues when united with standard for the betterment of our society. Among the recent AI based techniques Deep learning assists to build strong, accessible and operative solutions which can also be implemented to Malaria detection problem in the blood cell. In this proposed approach, a convolutional neural network (CNN) based model has been employed which not only accurately classify between parasitized and uninfected images but also directs the implementation of other disease's classification or detection in the similar domain. The schematic of the proposed approach is illustrated in the Fig. 1.

The structure of rest of this paper is prepared as follows: Sect. 2 deliberates the associated mechanisms in the existing field while Sect. 3 demonstrations the implementation and investigational effects and Sect. 4 comprehends the final observations and upcoming directions.

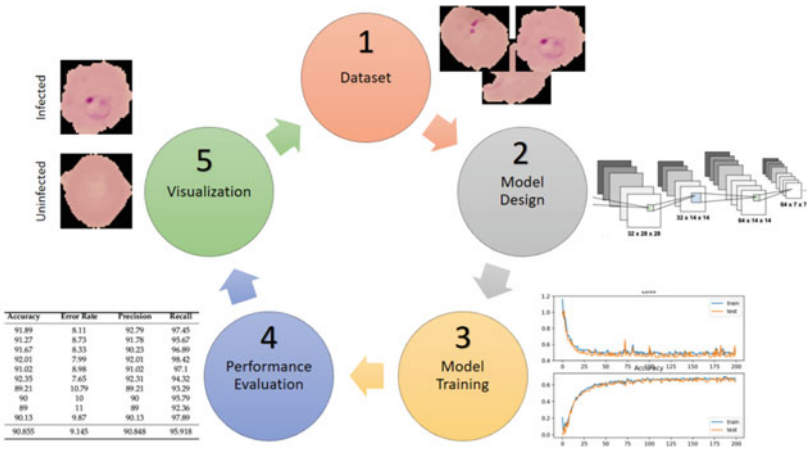



**Fig. 1** A process flow indicating the classification of blood cell against the malaria disease into infected and uninfected category

## 2 Related Works

Several morphological processing has been used to recognize the nature of red blood cell suffering from malaria parasite. Due to the presence of this parasite, area and perimeter of blood cell were affected and this change was determined with the classification accuracy of 95% against infected and non-infected cell [19]. A decision support system for the cells infected with parasites was carried out by iterative thresholding and associated element tagging [20]. The idea in the segmentation process is to implement Laplacian of Gaussian and coupled edge profile active contours to solve the issues associated with textual variation, and achieved 90% F1 score to identify infected cell using ANN classifier. The technique employed in [21] used adaptive image segmentation to perform segmentation in the blood cell images under varying lighting conditions. However, the performance for the selection of appropriate and significant features were restricted. In order to solve the global and challenging issues related with malaria detection in the blood cell, Roy et al. had applied pixel based discrimination method followed by various segmentation techniques to recognize the parasites from thin smear red blood cell images [22]. However, validation of the model had not been tested under different test images and thus restricted against the accurate classification of the malaria parasite due to varying quality of test images taken from the internet. Identification of significant features in this domain is challenging. Kazarine et al. image scanning cytometry and demonstrated that third harmonic generation imaging combined with image processing and malaria can be detected at single red blood cell [23]. To deal with this issue, features of color histograms and texture features both were implemented under a deep belief network (DBN) [24]. It was employed to recognize the existence of parasites in the blood smear images. The approach attained F-score of 89.66%, sensitivity of 97.60

and 95.92% of specificity. Malaria detection on microscopic images have also been experimented under machine learning strategy. In this direction, Saiprasath et al. had employed seven different machine learning approaches [25]. Out of which random forest attained the accuracy of 96.5% among Adaboost, decision tree, K- nearest neighbor, linear regression, naïve bayes and extra trees algorithms. Haryanto et al. discussed malaria detection using the image pigment color space which were further compared to find out the optimum color channels which describes the significant features of parasite plasmodium [26].

Likewise, with histograms of oriented gradients (HOG) and local binary patterns (LBP) features various machine learning techniques like k-nearest neighbor, Naïve Bayes and support vector machine (SVM) were applied [27]. The approach was based on the inductive method which segmented the parasites region through the adaptive machine learning and obtained the highest accuracy of 97.31%. Using the same methods, Devi et al. had extracted the parasite features in terms of prediction error and R-G color channel difference histogram and achieved 98.5% accuracy [28]. To extract suitable parasite information from the blood cell in order to identify the correct class of malaria, Vijayalakshmi et al. had implemented neural network based model using transfer learning approach [29]. With the model, authors achieved the 93.1% of classification accuracy for the identification of infected falciparum malaria. With the development of deep learning approaches, a 16-layer based CNN model was applied on the blood cell images to extract and learn the significant features [30]. Based on the learned features model classified the cell into infected and non-infected blood cell categories with 97.37% accuracy. With the deep learning approach, Delgado-Ortet et al. presented malaria detection in three stages which includes segmentation of erythrocytes, masking, and classification of malaria into infected and non-infected categories [31]. The presented pipeline achieved 93.72% global accuracy and 87.04% of specificity against malaria detection. Similarly, Dong et al. had implemented various deep learning based models including LeNet-5, AlexNet, GoogLeNet, and SVM, out of which GoogLeNet attained the accuracy of 98.13% using hand crafted features [32].

The mentioned approached have been utilized with the advancement of image processing and deep learning techniques. However, the performance in classifying the infected and un-infected blood cell can be improved using an optimal model design. Considering the study in [30] in which the classification accuracy is promising but having high computation cost as it used CNN with 16-layers. Obtaining feature with minimum number of model is also a potential idea. In the presented paper, a CNN based model has been proposed which has least design cost and capable to achieve high performance in terms of various standard evaluation metrics.

## 3 Materials and Methods

In order to determine the malaria parasite affected blood cell from the two classes namely: infected and unaffected (normal), following sequential phases have been considered.

### *3.1 Dataset*

For the offered method, a public dataset "Malaria Cell Images Dataset" has been considered from the Kaggle dataset [33]. In this dataset, two categories are available namely parasitized and uninfected. Total size of the dataset is 388 MB and contains 27,560 images in the mentioned categories. Image resolution in the dataset (width*height) is not uniform and thus making the classification task a bit challenging. The dataset has been arbitrarily distributed into the 70%, 15%, and 15% sections for training, validation, and testing, correspondingly, and has undergone through 100 epochs in the model training.

### *3.2 Model Design*

The feature which describes the parasite in the red blood cell are not uniform. They have varying structure and cannot be identified through the particular feature representation. It necessitates the extraction of a more number of important features which refer to the parasite affected blood cell. To generate and identify those significant features, CNN model is employed which further ensures and classifies the affected cell. It consists of an input layer, a set hidden layer, and an output layer which at last classifies the image. In the main components of the hidden layers, there are activation function, some pooling layers strategy, fully connected and normalization layers. All these components are the core building blocks to extract the significant feature of the target data through the set of operating kernel(s). It helps a model to identify the available features and generate various feature mappings to detect or classify the image in multiple categories. The developed model for the proposed work whose detailed configuration of the architecture has been given in Table 1. In the model design, hyper parameters including learning rate, selection of optimizer, batch size, number of epochs and number of layers with kernel have been thoroughly examined and applied to ensure the stability of learning while extracting the significant features is the prime objective for the presented domain.

**Table 1** Detailed configuration details of the proposed CNN model

| Layers | Output shape | Parameter |
|---|---|---|
| Convolutional | (None, 178, 178, 32) | 896 |
| Batch normalization | (None, 178, 178, 32) | 128 |
| Convolutional | (None, 176, 176, 32) | 9248 |
| Batch normalization | (None, 176, 176, 32) | 128 |
| Max pooling | (None, 88, 88, 32) | 0 |
| Convolutional | (None, 86, 86, 64) | 18,496 |
| Batch normalization | (None, 86, 86, 64) | 256 |
| Convolutional | (None, 84, 84, 64) | 36,928 |
| Batch normalization | (None, 84, 84, 64) | 256 |
| Max pooling | (None, 42, 42, 64) | 0 |
| Convolution | (None, 40, 40, 128) | 73,856 |
| Batch normalization | (None, 40, 40, 128) | 512 |
| Convolution | (None, 38, 38, 128) | 147,584 |
| Batch normalization | (None, 38, 38, 128) | 512 |
| Max poolin | (None, 19, 19, 128) | 0 |
| Flatten | (None, 46,208) | 0 |
| Dense | (None, 128) | 5,914,752 |
| Dropout | (None, 128) | 0 |
| Dense | (None, 1) | 129 |

### *3.3 Model Training*

The concept discussed in the Sect. 3.2 have been implemented and the behavior of trained model is analyzed. Model training is influenced from various hyper-parameter, design and optimizers used and the overall training performance is represented through two important factors which are the accuracy and loss curve over 6,203,681 parameters. Considering the Fig. 2(a) which describes the model accuracy gradually increases with the number of epochs and reaches up to its maximum level at approximately $100^{th}$ epoch. The training gap between the training and validation accuracy indicates the absence of overfitting problem and ensures the optimal impact on the classification result. Further, the network model remains less-fluctuated throughout its training period and does represent the stability in learning the significant features. Similarly, Fig. 2(b) point out the amount of loss through the loss curve during training and it started from 0.8 and the gradually decreased as the training proceeds. The network stability, gap between training and validation loss and the achieved minimum loss confirms the model training is up to the mark and considered its robustness in context of significant feature extraction.

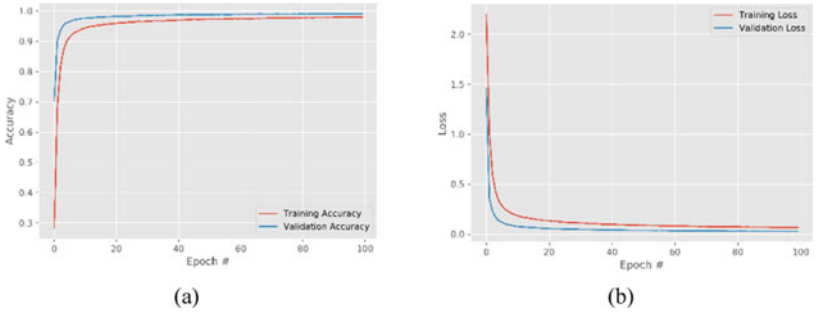



**Fig. 2** Training graph. **a** Model performance in terms of training and validation accuracies. **b** Behavior analysis of model training in terms of training and validation loss

## 4 Discussion and Performance Evaluation

Apart from the qualitative assessment of the model discussed in Sect. 3.3, quantitative evaluation has also been carried out to confirm the model performance represented through Fig. 2. For this, standard evaluation metrics is considered which are universally known parameters in machine/deep learning method including accuracy, precision, recall and F1 score. The mentioned metrics with the model's quantitative scores is represented in the Table 2. The obtained scores given in Table 2 matches with the training graph and confirms its effectivity in both approaches (qualitative and quantitative). With these average attained scores, performance of the other state-of-the-art techniques is also compared. It indicates that the proposed approach outperforms other related methods mentioned. Further, the model performance on classifying the infected and uninfected blood cell using the mentioned dataset has been illustrated in Fig. 3 and marked with probability scores.

**Table 2** Quantitative classification performance of proposed CNN model and its comparison with the relevant state-of-the-art technique

| Method | Accuracy | F1 | Precision | Recall |
|---|---|---|---|---|
| Poostchi et al. [19] | 95.00 | 94.50 | 95.40 | 94.40 |
| Saiprasath et al.[25] | 96.50 | 64.50 | 77.50 | 55.30 |
| Devi et al. [28] | 98.50 | 93.82 | – | 95.86 |
| Vijayalakshmi et al. [29] | 93.13 | 91.66 | 89.95 | 93.44 |
| Proposed model | 98.70 | 98.71 | 98.22 | 98.91 |

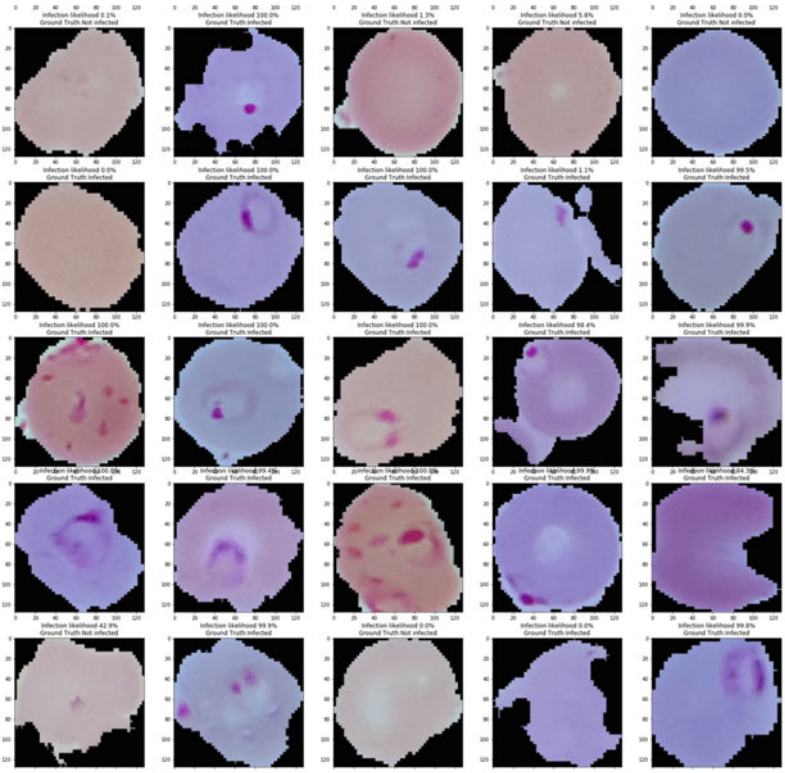


**Fig. 3** Run time behavior of the proposed CNN model using the public dataset. Sample output indicates the class prediction of each class with their probability scores

## 5 Conclusion

The proposed CNN based model which classifies the blood cell into malaria parasite infected and uninfected categories obtained with an accuracy measure of 98.70%. As the dataset used was balanced, the accuracy parameter has been utilized to assess the parameter. However, it has achieved decent score of 98.71% using F1 parameter (mostly used for unbalanced dataset). Hence, this model can be tested with unseen data too. However, the model can be optimized with different optimization techniques. In future, the model can be extended to classify other categories of malaria parasites.

## References


1. W.-H.-O. Malaria death report (2021) Key facts. https://www.who.int/news-room/fact-sheets/detail/malaria
2. Dewangan DK, Rathore Y (2011) Image quality costing of compressed image using full reference method. Int J Technol 1(2):68–71
3. Ali U, Dewangan KK, Dewangan DK (2018) Distributed denial of service attack detection using ant bee colony and artificial neural network in cloud computing. Adv Intell Syst Comput 652:165–175. https://doi.org/10.1007/978-981-10-6747-1_19
4. Chaudhuri A, Sahu TP (2021) A hybrid feature selection method based on Binary Jaya algorithm for micro-array data classification. Comput Electr Eng 90:106963. https://doi.org/10.1016/j.compeleceng.2020.106963
5. Pandey P, Dewangan KK, Dewangan DK (2018) Enhancing the quality of satellite images by preprocessing and contrast enhancement. In: Proceedings 2017 IEEE International Conference on Communication Signal Processing, ICCSP 2017, vol 2018-Janua, pp 56–60. https://doi.org/10.1109/ICCSP.2017.8286525.
6. Chaudhuri A, Sahu TP (2021) Feature selection using Binary Crow Search Algorithm with time varying flight length. Expert Syst Appl 168:114288. https://doi.org/10.1016/j.eswa.2020.114288
7. Bhattacharya N, Dewangan DK (2015) Fusion technique for finger knuckle print recognition. In: International conference on electrical, electronics, signals, communication and optimization, EESCO 2015. https://doi.org/10.1109/EESCO.2015.7253990
8. Dewangan DK, Sahu SP (2021) Deep learning-based speed bump detection model for intelligent vehicle system using raspberry pi. IEEE Sens J 21(3):3570–3578. https://doi.org/10.1109/JSEN.2020.3027097
9. Dewangan DK, Sahu SP (2021) RCNet: road classification convolutional neural networks for intelligent vehicle system. Intell Serv Robot 14(2):199–214. https://doi.org/10.1007/s11370-020-00343-6
10. Chaudhuri A, Sahu TP (2020) PROMETHEE-based hybrid feature selection technique for high-dimensional biomedical data: application to parkinson's disease classification. Electron Lett 56(25):1403–1406. https://doi.org/10.1049/el.2020.2517
11. Dewangan DK, Sahu SP (2021) Driving behaviour analysis of intelligent vehicle system for lane detection using vision-sensor. IEEE Sens J 21(5):6367–6375. https://doi.org/10.1109/JSEN.2020.3037340
12. Dewangan DK, Sahu SP (2021) PotNet: pothole detection for autonomous vehicle system using convolutional neural network. Electron Lett 57(2):53–56. https://doi.org/10.1049/ell2.12062
13. Dewangan DK, Sahu SP (2021) Road detection using semantic segmentation-based convolutional neural network for intelligent vehicle system. In: Data engineering and communication technology. lecture notes on data engineering and communications technologies. Springer Singapore, pp 629–637
14. Dewangan DK, Sahu SP (2021) Predictive control strategy for driving of intelligent vehicle system against the parking slots. In: 2021 5th international conference on intelligent computing and control systems (ICICCS), pp 10–. https://doi.org/10.1109/ICICCS51141.2021.9432362
15. Ojha A, Sahu SP, Dewangan DK (2021) Vehicle detection through instance segmentation using mask R-CNN for intelligent vehicle system. In: 2021 5th international conference on intelligent computing and control systems (ICICCS), pp 954–959. https://doi.org/10.1109/ICICCS51141.2021.9432374
16. Banjarey K, Sahu SP, Dewangan DK (2021) A survey on human activity recognition using sensors and deep learning methods. In: 2021 5th international conference on computing methodologies and communication (ICCMC), pp 1610–1617. https://doi.org/10.1109/ICCMC51019.2021.9418255

17. Pardhi P, Yadav K, Shrivastav S, Sahu SP, Dewangan DK (2021) Vehicle motion prediction for autonomous navigation system using 3 dimensional convolutional neural network. In: 2021 5th international conference on computing methodologies and communication (ICCMC), pp 1322–1329.https://doi.org/10.1109/ICCMC51019.2021.9418449
18. Sahu SP, Dewangan DK, Agrawal A, Priyanka TS (2021) Traffic light cycle control using deep reinforcement technique. In: 2021 international conference on artificial intelligence and smart systems (ICAIS), pp 697–702. https://doi.org/10.1109/ICAIS50930.2021.9395880
19. Poostchi M et al (2018) Malaria parasite detection and cell counting for human and mouse using thin blood smear microscopy. J Med Imaging 5(04):1. https://doi.org/10.1117/1.jmi.5.4.044506
20. Somasekar J, Sharma A, Madhusudhana Reddy N, Padmanabha Reddy YCA (2020) Image analysis for automatic enumeration of rbc infected with plasmodium parasites-implications for malaria diagnosis. Adv Math Sci J 9(3):1221–1230. https://doi.org/10.37418/amsj.9.3.48
21. Dave IR, Upla KP (2017) Computer aided diagnosis of Malaria disease for thin and thick blood smear microscopic images. In: 2017 4th international conference on signal processing integration networks, SPIN 2017, pp 561–565. https://doi.org/10.1109/SPIN.2017.8050013
22. Roy K, Sharmin S, Mukta RBM, Sen A (2018) Detection of malaria parasite in giemsa blood sample using image processing. Int J Comput Sci Inf Technol 10(1):55–65. https://doi.org/10.5121/ijcsit.2018.10105
23. Kazarine A, Baakdah F, Gopal AA, Oyibo W, Georges E, Wiseman PW (2019) Malaria detection by third-harmonic generation image scanning cytometry. Anal Chem 91(3):2216–2223. https://doi.org/10.1021/acs.analchem.8b04791
24. Bibin D, Nair MS, Punitha P (2017) Malaria parasite detection from peripheral blood smear images using deep belief networks. IEEE Access 5:9099–9108. https://doi.org/10.1109/ACCESS.2017.2705642
25. Saiprasath G, Babu RN, Priyan JA, Vinayakumar R, Sowmya V, Soman K (2019) Performance comparison of machine learning algorithms for malaria detection using microscopic images. Int J Res Anal Rev 6(1):86–90. www.ijar.org.
26. Haryanto SEV, Mashor MY, Nasir ASA, Jaafar H (2017) Malaria parasite detection with histogram color space method in Giemsa-stained blood cell images. In: 2017 5th international conference cyber it service and management, CITSM 2017, pp 4–7. https://doi.org/10.1109/CITSM.2017.8089291
27. Denton D, Mcllroy P (2018) Plasmodium life cycle stage classification based quantification of malaria parasitaemia in thin blood smears, pp 2008–2018
28. Devi SS, Roy A, Singha J, Sheikh SA, Laskar RH (2018) Malaria infected erythrocyte classification based on a hybrid classifier using microscopic images of thin blood smear. Multimedia Tools Appl. 77(1):631–660. https://doi.org/10.1007/s11042-016-4264-7
29. Vijayalakshmi A, Rajesh Kanna B (2020) Deep learning approach to detect malaria from microscopic images. Multimedia Tools Appl. 79(21–22):15297–15317. https://doi.org/10.1007/s11042-019-7162-y
30. Liang Z, et al (2017) CNN-based image analysis for malaria diagnosis. In: Proceedings - 2016 IEEE international conference on bioinformatics and biomedical, BIBM 2016, pp 493–496. https://doi.org/10.1109/BIBM.2016.7822567
31. Delgado-Ortet M, Molina A, Alférez S, Rodellar J, Merino A (2020) A deep learning approach for segmentation of red blood cell images and malaria detection. Entropy 22(6):1–16. https://doi.org/10.3390/e22060657
32. Dong Y, et al (2017) Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. In: 2017 IEEE EMBS international conference on biomedical healths and informatics, BHI 2017, pp 101–104. https://doi.org/10.1109/BHI.2017.7897215
33. Arunava (2019) Malaria cell images dataset cell-images for detecting malaria. https://www.kaggle.com/iarunava/cell-images-for-detecting-malaria.

# Miniaturized Defected Ground Structure Microstrip Patch Antenna Design for X and Ku Band Applications

**P. Anitha, S. Latha, and Kalyan Reddy**

**Abstract** A compact Micro-strip patch antenna is designed for wideband applications in the frequency range of 6-22 GHz covering standard microwave frequency bands X (8–12 GHz) and Ku (12–18 GHz). The proposed structure is fabricated on FR-4 ($\varepsilon r = 4.4$) substrate with a thickness of 0.8 mm has an overall volume of 180 $mm^3$ on a partial ground plane. The proposed structure is resonating at 11 GHz with a minimum return loss of −21 dB shows a better impedance matching ($S_{11} \leq 10$ dBi). The results show that the antenna covers the frequency spectrum of 6.5–22.5 GHz with a −10 dB return loss, maximum gain of 2.7 dBi and has Efficiency of 93%. The defected ground structure (DGS) along with microstrip line feed results in a compact wide band patch antenna. The proposed structure is simulated using HFSS software and the simulation results are in good agreement with the measured results. Impedance bandwidth has been enhanced by utilizing a partial ground plane of 6 × 15 $mm^2$ compared to conventional MPA, due to which a single antenna may be useful for multi functions.

**Keywords** DGS · Ku band · Microstrip Line Feed · X band

## 1 Introduction

Rapid development in satellite and radar communication system requires a new type of antennas such as compact broadband high gain and multi-frequency antennas. Attractive merits of Micro-strip patch antenna such as low cost, light weight, low profile planar configuration and conformal to host surface would satisfy the above requirements. Especially, X band and Ku band microstrip antennas are widely used

P. Anitha (✉) · S. Latha
Department of ECE, SJBIT, Bengaluru, Karnataka, India
e-mail: anitha.peram@gmail.com

K. Reddy
Department of ECE, AITT, Tirupathi, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_57

for point-to-point wireless communications, microwave imaging, satellite communication, industrial scientific medical, spacecraft, and military, missiles, marine radar, tracking, and mapping [1, 2].

Despite of the attractive features, patch antennas suffer from narrowband, low gain problems: However, reduction in size, enhancement in bandwidth, improvement in gain is the major design constraints for commercial applications.

## 2 Related Work

To improve the bandwidth and to have a compact structure, techniques such as creating slots on ground plane and on patch, stacking of patches, using a thick substrate, lowering the dielectric constant, using gap coupled Multi-resonator and loading of shorting pins are reported [3–5]. A U slot on the circular ground plane with U-shaped patch resulted in wide bandwidth [6]. Compactness is achieved by Shorting Wall technique [7]. Gain enhancement is achieved by etching arc-shaped slot into the radiating patch [8]. A notch loaded with multiple shorting techniques is designed for X and Ku band applications [9]. A RMSA with rounded corners and defected, partial and slotted ground plane structure is designed for multiband applications [10]. A wide multi band antennas [11–13] are proposed on partial ground plane with rectangular and circular patch variants but occupies a large volume. A 30 $\times$ 30 $\times$ 1.6 $mm^3$ elliptical triangular patch on defected ground structure is proposed for wireless applications [14]. A 20 $\times$ 20 $\times$ 1.6 $mm^3$ circular patch with polygon slits on partial ground plane [15] is designed for multiband applications. A CPW fed patch with DGS [16] is proposed with an impedance bandwidth of 76.8% to cover multiple bands. The various techniques deployed in the survey to overcome the drawbacks are well appreciated, the simplest one out of which is using a partial ground plane, microstrip line feed to work in the specified two bands is deployed here to meet the requirements.

In the proposed work, we have designed an antenna with reduced size of 15 $\times$ 15 $\times$ 0.8 $mm^3$ using defective ground plane technique to meet the requirements with respect to resonating frequency, impedance bandwidth, gain and efficiency. The fabricated antenna is tested for $S_{11}$ & bandwidth, Gain & radiation pattern results using VNA and anechoic chamber.

## 3 Methodology

### *3.1 Antenna Design*

The antenna is designed to operate at a frequency of 11 GHz. The substrate used is a (flame retardant) FR-4 Epoxy material which has a dielectric constant of 4.4 and a

loss tangent of 0.01. The dimensions of the patch such as patch length (L), width (W), ground plane length and width (Lg & Wg), feed line length and width ($L_T$ & $W_T$) are calculated according to the transmission line model [17] at a given $f_0$ (11 GHz) with a substrate thickness of 'h'(1.6 mm) and a substrate dielectric constant of $\varepsilon_r$ (4.4). The width of the patch (W) is depends upon the substrate dielectric constant, resonant frequency and velocity of light and is evaluated by using Eq. (1) as

$$W = \frac{C}{2f}\sqrt{\frac{2}{\varepsilon r + 1}} \tag{1}$$

The length of the patch depends upon guided wavelength and fringing length and is calculated using Eq. (2) as

$$L = \frac{\lambda g}{2} - 2\Delta L \tag{2}$$

where λg is the guided wavelength depends upon operating frequency and effective dielectric constant and is calculated using Eq. (3) as.

$$\lambda g = \frac{c}{f\sqrt{\varepsilon reff}} \tag{3}$$

where f is the desired resonance frequency, C is speed of the light in a vacuum and εreff is the effective dielectric constant and is determined using Eq. (4) as

$$\varepsilon reff = \left(\frac{\varepsilon r + 1}{2}\right) + \left(\frac{\varepsilon r - 1}{2}\right)\left(1 + 12\frac{h}{w}\right)^{-1/2} \tag{4}$$

Due to the fringing fields around the periphery of the patch, the antenna looks larger than its physical dimensions. ΔL accounts for this, which is given by Eq. (5) as.

$$\Delta L = 0.412h\left[\frac{(\varepsilon reff + 0.3)}{(\varepsilon reff - 0.258)}\frac{\left(\frac{w}{h} + 0.264\right)}{\left(\frac{w}{h} + 0.8\right)}\right] \tag{5}$$

Ground plane dimensions such as Length (Lg) and width (Wg) of the ground plane are calculated according to Eq. (6) as

$$Lg = 6h + L \;\&\; Wg = 6h + W \tag{6}$$

The patch impedance is calculated using Eq. (7), according to the transmission line model [17].

$$Z_a = 90\frac{\varepsilon_r^2}{\varepsilon_r - 1}\left(\frac{L}{W}\right)^2 \tag{7}$$

Feed line impedance is the geometric mean of patch impedance (Za) and characteristic impedance (Zo = 50Ω) which is calculated using Eq. (8) as

$$Z_T = \sqrt{Z_a Z_0} \tag{8}$$

$Z_T$ also depends upon height and width of the feed line and is represented as

$$Z_T = \frac{60}{\sqrt{\varepsilon_r}} \ln\left(\frac{8h}{W_T} + \frac{W_T}{4h}\right) \tag{9}$$

From Eq. (9), we can evaluate the feed line width $W_T$. The feed line length can be evaluated using Eq. (10) as

$$L_T = \frac{\lambda_0}{4\sqrt{\varepsilon_{re}}} \tag{10}$$

The various parameters such as Reflection co-efficient, return loss (RL) and VSWR are evaluated using the Eqs. (11), (12) and (13) respectively

$$\text{Reflection Co-efficient } \Gamma = \frac{Z - Z_{in}}{Z + Z_{in}} \tag{11}$$

where Z is the impedance of the microstrip line feed and is 50 Ω. $Z_{in}$ is the total input impedance of the proposed antenna is evaluated using TEM analysis [17]

$$\text{Return Loss } = -20 \log|\Gamma| \tag{12}$$

$$\text{VSWR } = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{13}$$

Defected ground structure is implemented by etching off a simple shape (rectangle) on the ground plane. In the proposed antenna, the radiating element is a rectangular patch. The geometry and the fabricated structure of the proposed antenna is shown in Figs. 1 & 2 and the specifications of the proposed antenna is given in Table 1.

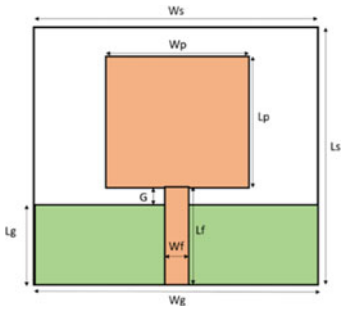



**Fig. 1** Geometry of the proposed antenna

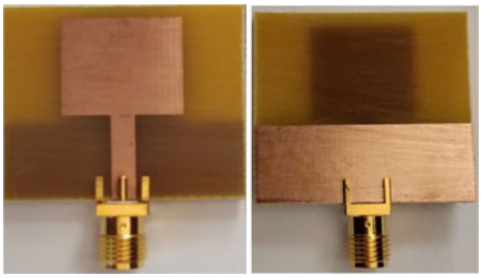


**Fig. 2** Front & back view of the fabricated structure

**Table 1** Antenna specifications

| Parameter | $W_S$ | $L_S$ | $W_G$ | $L_G$ | $W_P$ | $L_P$ | $W_F$ | $L_F$ | G | H |
|---|---|---|---|---|---|---|---|---|---|---|
| Values (mm) | 15 | 15 | 15 | 6 | 7.5 | 7.5 | 1.5 | 6.5 | 0.5 | 0.8 |

## *3.2 Parametric Analysis*

From various simulations that we have conducted using HFSS, analysis is presented with respect to the most important parameters which have affected the antenna return loss performance are (a) Patch dimensions variation (b) Ground plane dimensions variation (C) Gap between ground planes and patch variations (D) substrate dimensions variation.

**Patch Dimensions' Variation.** The Fig. 3 represents $S_{11}$ for different patch dimensions ($W_p$ & $L_p$) variation. It is observed that with patch size of 7.5 × 7.5 $mm^2$, attained a maximum bandwidth with good impedance matching. When we try increase the patch dimensions, impedance bandwidth has drastically reduced. When the patch size is reduced to 5.5 × 5.5 impedance matching is improved but there is degradation in bandwidth.

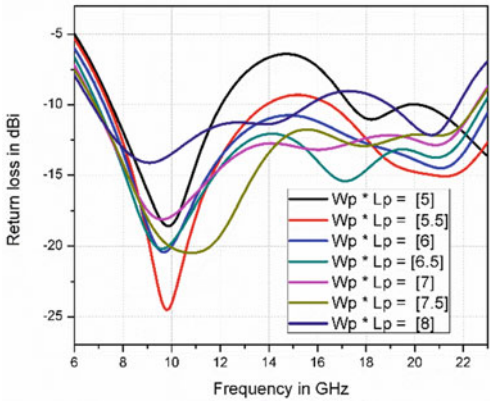



**Fig. 3** $S_{11}$ for patch dimension variations

**Ground Plane Dimensions' Variation.** The Fig. 4 represents $S_{11}$ for different ground plane ($L_g$ & $W_g$) dimensions. It is observed that when the ground plane size is $15 \times 6\,mm^2$, maximum bandwidth and good impedance matching has attained. When the ground plane dimensions are varied in either directions, there is degradation in bandwidth/impedance matching.

**Gap Between Ground and Patch Variation.** The Fig. 5 represents $S_{11}$ for different gap (between the ground plane and patch of an antenna) variations. It is observed that when gap size is 0.5 mm, maximum bandwidth and good impedance matching has attained. When the gap is varied in either directions, a significant degradation in bandwidth and impedance matching is found.

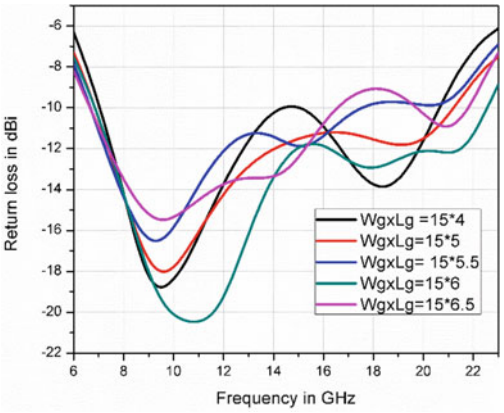



**Fig. 4** $S_{11}$ for Ground plane variation

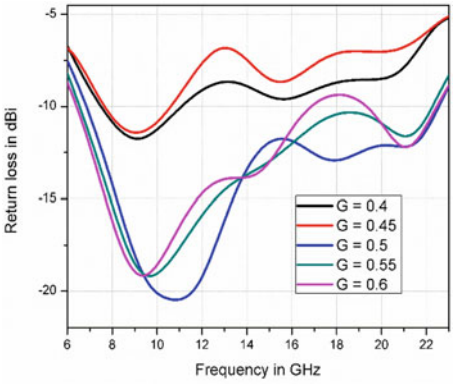



**Fig. 5** Return loss for gap variation

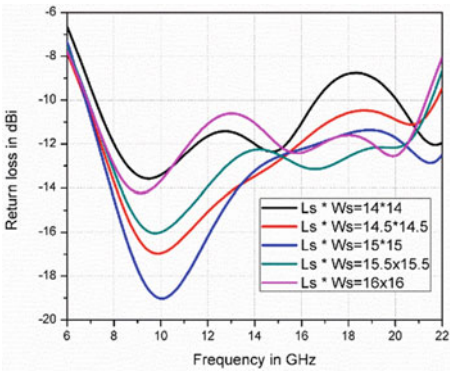



**Fig. 6** Return loss for substrate variation

**Substrate Dimensions' Variation.** The Fig. 6 represents $S_{11}$ for different substrate ($L_s$ & $W_s$) variations. It is observed that when substrate size is $15 \times 15$ mm$^2$, attained a maximum bandwidth and good impedance matching. When the substrate dimension is either increased or decreased, there is drastic degradation in bandwidth and impedance matching is found.

## 4 Results and Discussions

The proposed micro-strip patch antenna is designed in the X and Ku band frequency range. The same is fabricated and tested for return loss, radiation pattern, gain etc., It is noticed that the bandwidth is increased to a considerable extent. Simulated results

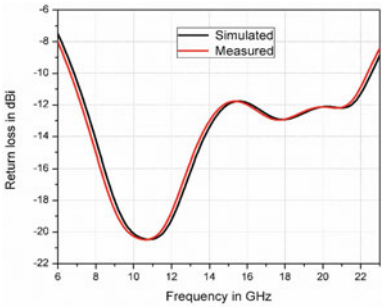



**Fig. 7** Return loss plot

obtained through HFSS which are in good agreement with the measured results obtained after testing through VNA and Anechoic Chamber.

## *4.1 Return Loss*

The Fig. 7 represents the simulated and measured results of return loss of the designed antenna. From the result it is observed that the antenna is resonating at 11 GHz and has a minimum return loss of −21 dB.This antenna has an impedance bandwidth of 16 GHz (6.5–22.5 GHz).

## *4.2 VSWR*

The Fig. 8 represents the simulated and measured results of VSWR. The VSWR of the structure is 1.2 and it is observed that the designed antenna is operating in the frequency range 6.5–22.5 GHz resulting in a bandwidth of 145%.

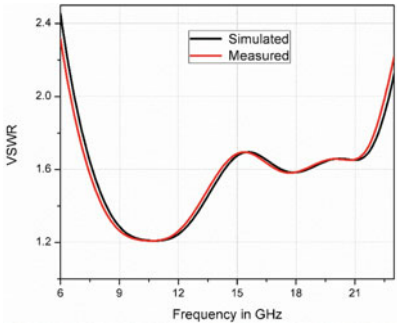



**Fig. 8** VSWR plot

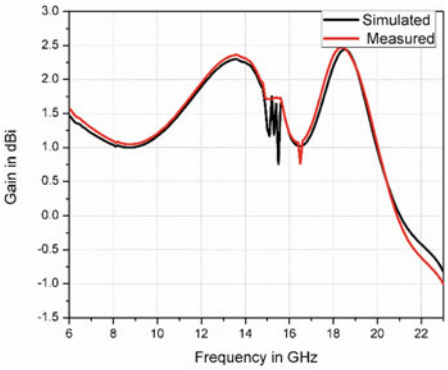



**Fig. 9** Gain Plot

## *4.3 Gain*

From Fig. 9 we can observe that gain is varying from 1–2.7 dB in the operating frequency range of 6–21 GHz and has a maximum gain of 2.7 dBi at 19 GHz. It is maintained above 1.5 dBi throughout the X and Ku band.

## *4.4 Radiation Pattern*

From Fig. 10 one can observe that the simulated and the measured results are almost same. In the E plane the antenna is radiating in all directions, where as in H plane it is omnidirectional.

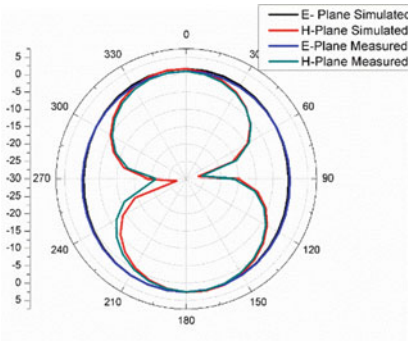



**Fig. 10** Radiation Pattern of E and H Plane at 11 GHz

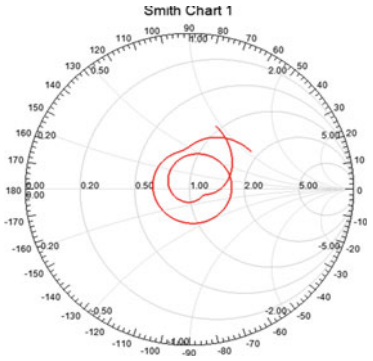



**Fig. 11** Smith Chart

## *4.5 Smith Chart*

Smith chart representation is shown in the Fig. 11. From smith chart with VSWR circle of radius 1 is giving a bandwidth of 16 GHz (6.5–22.5 GHz) for the designed antenna.

## *4.6 Directivity*

The directivity of the proposed structure is varying from −0.5 to 3.25 dBi in the frequency range of 6–23 GHz which is shown in Fig. 12. By using directivity and gain value at any frequency one can calculate the efficiency of the patch antenna. At 10 GHz the directivity is 1.25 dBi and gain is 1.167 dBi. Therefore, efficiency at 10 GHz is 93.05%. Table 2 gives the comparison of various parameters such as

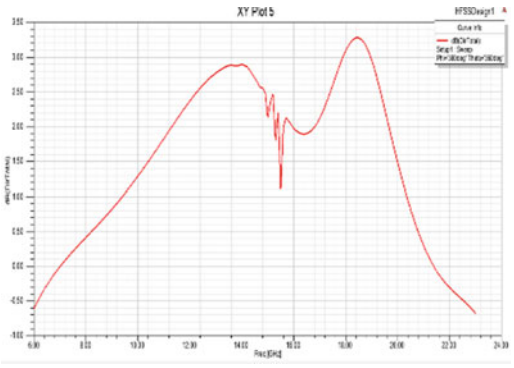



**Fig. 12** Directivity Plot

**Table 2** Comparison of the proposed antenna with respect to literature

| Reference | Overall size ($mm^3$) | Bandwidth (in GHz) | Gain (in dBi) | Efficiency (in %) | Return loss ($S_{11}$ in dB) |
|---|---|---|---|---|---|
| [9] | $20 \times 25 \times 1.57$ (785 $mm^3$) | 1 GHz (in both X & Ku band) | 7.42 | 70 | −21 |
| [10] | $30 \times 36.7 \times 1.6$ (1761.6 $mm^3$) | 2.5 and 2.36 (in both X & Ku band) | 2.31 | 80 | −37 to –48 |
| **Proposed structure** | **$15 \times 15 \times 0.8$** (180 $mm^3$) | **16 GHz(6.5–22.5 GHz)** | **2.7** | **93** | **-21** |

Antenna size, Bandwidth, gain efficiency and return loss of the proposed structure with respect to the literature.

## 5 Conclusion

A wideband micro-strip patch antenna is designed with an overall volume of 180 $mm^3$. The structure is operating in the frequency range of 6.5–22.5 GHz which covers the X and Ku band applications. The defective ground plane along line feed has increased the bandwidth. The proposed structure has achieved a good impedance matching of −21 dB with good antenna efficiency of 93%. The antenna gain varies from 1–2.7 dBi in the band. The simulated and tested results are in good agreement with each other. The proposed structure is fabricated using Flame retardant 4 (FR-4 with $\varepsilon r = 4.4$) type of substrate and simulated using HFSS software V13.0 with 32-bit RAM PC. The simulated and tested results are in good agreement with each other. The gain can be improved further by using an array of proposed structure.

## References

1. Balanis CA (2005) Antenna theory analysis and design, 3rd edn. Wiley, Hoboken, p 816
2. Kumar G, Ray KP (2003) Broadband microstrip antennas. Artech: Artech House
3. Deshmukh AA, Ray KP, Kadam A (2014) Analysis of slot cut broadband and dual band rectangular microstrip antennas. IETE J Res 59:193–200
4. Kula J, Psychoudakis D, Liao W-J, Chen C-C, Volakis J, Halloran J (2006) Patch antenna miniaturization using recently available ceramic substrates. IEEE Ant Propag Mag 48(6):13–20
5. Wong KL, Chen WS (1997) Compact microstrip antenna with dual frequency operation. Electron Lett 33(8):646–647
6. Mandal K, Sarkar PP (2016) A compact low profile wideband U-shape antenna with slotted circular ground plane. Int J Electron Commun (AEU) 70(3):336–340
7. Shakib MN, Moghavvemi M, Mahadi WNL (2015) A low-profile patch antenna for ultrawideband application. IEEE Ant Wirel Propag Lett 14:1790–1793

8. Tang M-C, Shi T, Ziolkowski RW (2016) Planar ultra-wideband antennas with improved realized gain performance. IEEE Trans Ant Propag 64(1):61–69
9. Yadav S, Jaiswal K, Patel AK, Singh S, Pandey AK, Singh R (2019) Notch-loaded patch antenna with multiple shorting for x and ku band applications. In: Khare A, Tiwary U, Sethi I, Singh N (eds) Recent trends in communication, computing, and electronics. LNEE, vol 524. Springer, Singapore
10. Sharma IB, Lohar FL, Maddila RK, Deshpande A, Sharma MM (2018) Design of five-band microstrip antenna for radar and satellite communications. In: Janyani V, Tiwari M, Singh G, Minzioni P (eds) Optical and Wireless Technologies. LNEE, vol 472. Springer, Singapore
11. Anitha Peram A, Reddy Subba Rami, Giri Prasad MN (2019) Miniaturized single layer ultra wide band (UWB) patch antenna using a partial ground plane. Wirel Pers Commun 106(3):1275–1291
12. Bag B, Biswas P, De S et al (2020) A wide multi-band monopole antenna for GSM/WiMAX/WLAN/X-Band/Ku-band applications. Wirel Pers Commun 111:411–427
13. Sharma PK, Shrama D, Rani CJ (2020) A flag-shaped microstrip patch antenna for multiband operation. In: Khanna A, Gupta D, Bhattacharyya S, Snasel V, Platos J, Hassanien A (eds) International conference on innovative computing and communications. AISC, vol 1059. Springer, Singapore
14. Ghouz HHM, Sree MFA, Mohamed HA, Ibrahim MA (2020) Novel compact microstrip monopole antenna for UWB wireless applications. In: International applied computational electromagnetics society symposium (ACES), pp1–2
15. Abirami R, Gnanamanoharan E, Sivagnanam S, Priya KP (2020) Design of compact microstrip patch antenna for K Band and KU band applications. In: International conference on system, computation, automation and networking (ICSCAN), pp 1–7
16. Anjaneyulu G, Varma JS (2021) Design of a CPW fed microstrip patch antenna with defective ground structure for wireless applications. In: 2021 6th international conference on inventive computation technologies (ICICT), pp 25–30
17. Garg R (2001) Microstrip design handbook. Artech House. Inc., Norwood

# Multilayer Perceptron Neural Network Supervised Learning Based Solar Radiation Prediction


**M. Shyamala Devi, A. Peter Soosai Anandaraj, K. Venkata Thanooj, P. V. Sandeep Guptha, and A. Jayanth Reddy**

**Abstract** Worldwide solar radiation is a basic parameter for the plan and operation of solar radiation frameworks. Long-standing records of worldwide solar radiation data are not accessible in many places because of the cost and maintenance of the measuring instruments. Sun based radiation expectation contains an incredible significance in power generation from sun based energy and makes a difference to measure photovoltaic control frameworks. With this overview, this paper enabled to predict the solar radiation using machine learning algorithms. With this context, we have utilized solar radiation dataset extracted from UCI Machine Learning. The proposed model Multilayer perceptron (MLP) based Neural Network for forecasting solar radiation are attained in four ways. Firstly, the data set is analyzed and preprocessed with Feature Scaling and missing values. Secondly, the correlation of the features is done and the relation of each features are visualized. Thirdly, the raw solar radiation data set is fitted to all the regressors and the implementation is furnished before and after scaling. Fourth, the raw data set is subjected to multilayer perceptron with various activation layers like identity, logistic, tanh and relu layers. The performance is analyzed with EVS, MAE, MSE, RScore and run time of the neural network layer. The execution is done using python language under Spyder platform with Anaconda Navigator. Experimental results show that the Gradient boost regressor have the RScore of 0.98 before and after feature scaling. The MLP regressor with TANH activation layer is tends to retain 0.99 Rscore before and after scaling.

**Keywords** Activation layer · Correlation · Multilayer perceptron · Regressor

## 1 Introduction

The accessibility of sensibly exact worldwide sun powered radiation information is imperative for the victory of any sun powered venture [1]. In any case, as it were

M. Shyamala Devi (✉) · A. Peter Soosai Anandaraj · K. Venkata Thanooj · P. V. Sandeep Guptha · A. Jayanth Reddy
Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India
e-mail: shyamaladevim@veltech.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_58

a couple of meteorological stations around the world capture these information as a result of the tall taken a toll of measuring hardware and the need of specialized capability in calibrating them. In an endeavor to resolve this challenge, engineers and analysts have created different options to create the information [2]. Due to the shortage of hardware and the tall costs of upkeep, distant less perceptions of sun based radiation are made than perceptions of temperature, precipitation and other climate variables [3].

## 2 Related Works

This paper builds the solar power predictor using the ensemble methods. They have used four approaches for getting the best forecast out of any learning like Profound Learning, Machine Learning, Ensemble Learning and Conventional Strategy [1] (ARIMA, LSTM). This paper design sun oriented radiation predictor using ANN models with distinctive back engendering calculations and construct a sun powered radiation predictor. After collecting the information, the complete dataset is separated into two categories such as preparing set and testing set. It is prepared with the assistance of the ANN model. Among the collected information, 80% of information was utilized for training [2]. This paper constructs a clustering show for solar based prediction. It uses different ML approaches to create model that diminishing the fluctuation, inclination as well as progressing the forecasts. The execution of the proposed approach will be evaluated by Mean Absolute Error (MAE) Normalized Root Mean Square Error (nRMSE) Prediction Accuracy [3]. The solar radiation model is built with the ML algorithms The input information was decided in five diverse bunches at the conclusion of the determination handle, and the improvement forms of the most excellent ML models were clarified for each group [4].

This paper built the solar radiation model using the root regularization based outfit methods and construct a sun powered radiation determining model. A few nonlinear models counting SVMs and ANNs have been considered to figure worldwide sun based radiation. To overcome the disadvantages of a single, demonstrate, which yields low estimating precision, a novel worldwide sun oriented radiation estimating strategy called RS-SRSCAD-FA has been proposed [5]. This paper utilizes nearby climate figures and foresee hourly sun oriented irradiance. The proposed model employments climate parameters comparable to those utilized by the reference model. Beneath the presumption that information communication is performed on an everyday premise, the input information of the model utilized for expectation was overhauled once each 24 h, and the model anticipated with the overhaul to set up the 24-h sun powered irradiance of the following day. The proposed model has three points of interest over the existing sun powered irradiance expectation model [6]. This paper coordinates ML strategies and make a day ahead solar expectation method. It estimates are determined from each gathering part by distinguishing estimate hailstorms, coordinating the estimate storms with watched hailstorms, extricating information inside the storm ranges, and after that fitting a machine learning model.

Making strides accost forecast with more exact data around anticipated salute areas and concentrated will permit individuals to moderate a few of the potential effect of hail [7]. This paper builds the precipitation prediction method using profound learning approach. The method is composed of two systems: an auto encoder organizes and a multilayer perceptron network. This paper has displayed a profound learning approach based on using the auto encoders and neural systems to anticipate the amassed precipitation for another day [8]. This paper construct a wind speed predictor using machine learning strategies and construct a wind speed predictor. The essential concept behind the ANN is to create an instrument that ought to perform computation for illustrating the brain function. Mutual data highlight selection finds the critical highlights to decrease the complexity of the wind speed determining model [9]. This paper built a heat transfer rate predictor using AI approach using SVM model that determines the rate of heat transfer based on wire on tube type of heat exchanger [10]. Sun oriented radiation expectation incorporates an awesome significance in power era from sun powered vitality and makes a difference to estimate photovoltaic control frameworks [11]. The major objective of this work is to create an ANN demonstrate for precisely foreseeing sun powered radiation [12]. The prediction of solar radiation is done by applying all the machine learning model and classifiers [13, 14]. The most objective of this paper is to display calculation to predict hourly sun powered radiation within the short/medium term, combining data around cloud scope level and chronicled sun powered radiation registers, which expanded the execution and the precision [15].

## 3 Proposed Architecture

The overall workflow of MLP based Neural Network model is shown in Fig. 1. The paper contributions are given below.

(i) Firstly, the data set is preprocessed with Feature Scaling and Missing Values.
(ii) Secondly, the correlation of the features is done and the relation of each features are visualized.
(iii) Thirdly, raw data set is fitted to all the regressor like Linear Regression, Ridge Regression, ElasticNet Regression, Lars Regression, LarsCV Regression, Lasso Regression, LassoLarsCV Regression, BayesianRidge, ARDRegression, Decision Tree Regression, Extra Tree Regression, AdaBoost Regression, GradientBoosting Regression and RandomForest Regression with and without the presence of feature scaling.
(iv) Fourth, the raw data set is subjected to multilayer perceptron with various activation layers like identity, logistic, tanh and relu layers.

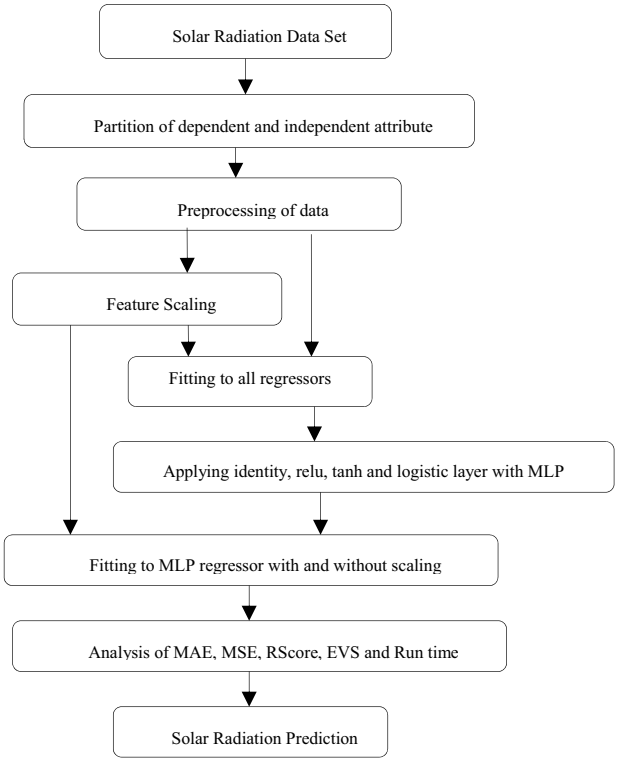



**Fig. 1** Overall architecture flow

## 4 Results and Discussion

### *4.1 Qualitative Analysis with Dataset*

The solar radiation dataset extracted from the UCI machine learning repository is used for implementation. The dataset consists of the features namely (UNIXTime, Data, Time, Radiation, Temperature, Pressure, Humidity, Wind Direction in Degrees, Speed, Sun Rise time, SunSet time). The code is drafted with python under Anaconda Navigator with Spyder IDE. The data set is split with 80:20 for training and testing dataset. The exploratory data analysis of the solar radiation dataset is shown in Fig. 2.

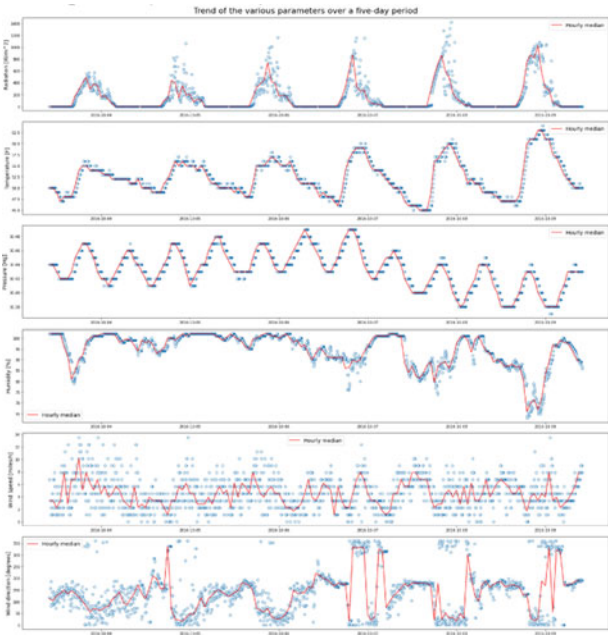



**Fig. 2** Distribution of parameters of the solar radiation dataset

## *4.2 Quantitative Analysis with Different Classifier Algorithms*

The raw data set is fitted to all the regressors like Linear Regression, Ridge Regression, ElasticNet Regression, Lars Regression, LarsCV Regression, Lasso Regression, LassoLarsCV Regression, BayesianRidge, ARDRegression, Decision Tree Regression, Extra Tree Regression, AdaBoost Regression, GradientBoosting Regression and RandomForest Regression with and without the presence of feature scaling and performance is shown in Tables 1 and 2, the RScore, running time, regressor performance comparison is shown in Figs. 3, 4 and 5.

**Table 1** Regressor performance of the raw dataset before scaling

| Regressor | EVS | MSE | MAE | RScore | Running time (ms) |
|---|---|---|---|---|---|
| Linear | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.14 |
| Ridge | 0.57 | 0.000439535 | 1.5779 | 0.57 | 0.01 |
| ElasticNet | 0.55 | 0.000453448 | 1.6249 | 0.55 | 0.01 |
| Lars | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.01 |
| LarsCV | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.07 |
| Lasso | 0.56 | 0.00044339 | 1.5937 | 0.56 | 0.01 |
| LAssoLarsCV | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.05 |
| Bayesian | 0.57 | 0.000439536 | 1.5778 | 0.57 | 0.02 |
| ARD | 0.57 | 0.000439555 | 1.5780 | 0.57 | 0.03 |
| DecisionTree | 0.52 | 0.000488893 | 1.0998 | 0.52 | 0.16 |
| ExtraTree | 0.49 | 0.000512972 | 1.1300 | 0.79 | 0.05 |
| AdaBoost | 0.59 | 0.000451648 | 1.6225 | 0.55 | 0.49 |
| Gradient boost | 0.69 | 0.000316029 | 1.0973 | 0.96 | 1.84 |
| RandomForest | 0.58 | 0.000427887 | 1.3801 | 0.88 | 0.93 |

**Table 2** Regressor performance of the raw dataset after scaling

| Regressor | EVS | MSE | MAE | RScore | Running time (ms) |
|---|---|---|---|---|---|
| Linear | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.01 |
| Ridge | 0.57 | 0.000439537 | 1.5777 | 0.57 | 0.01 |
| ElasticNet | 0.50 | 0.000508116 | 1.7213 | 0.50 | 0.00 |
| Lars | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.00 |
| LarsCV | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.04 |
| Lasso | 0.57 | 0.000439636 | 1.5786 | 0.57 | 0.00 |
| LAssoLarsCV | 0.57 | 0.000439537 | 1.5778 | 0.57 | 0.05 |
| Bayesian | 0.57 | 0.000439537 | 1.5777 | 0.57 | 0.01 |
| ARD | 0.57 | 0.000439555 | 1.5780 | 0.57 | 0.01 |
| DecisionTree | 0.51 | 0.000492491 | 1.1065 | 0.51 | 0.16 |
| ExtraTree | 0.49 | 0.000512972 | 1.1300 | 0.76 | 0.05 |
| AdaBoost | 0.58 | 0.000470879 | 1.6892 | 0.53 | 0.44 |
| Gradient boost | 0.69 | 0.000316029 | 1.0973 | 0.97 | 1.85 |
| RandomForest | 0.58 | 0.000427887 | 1.3801 | 0.88 | 0.91 |

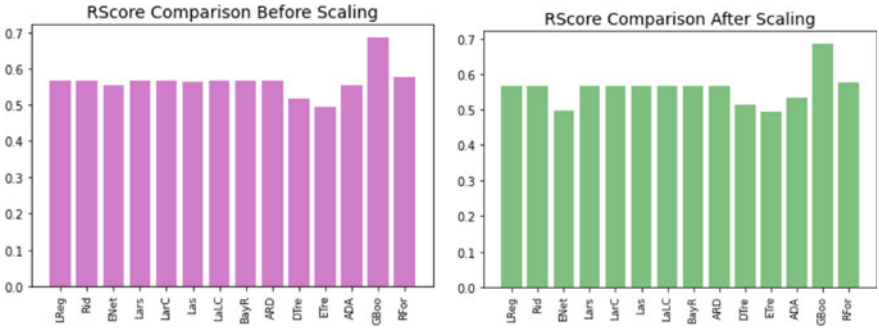



**Fig. 3** RScore Analysis of raw dataset before and after feature scaling

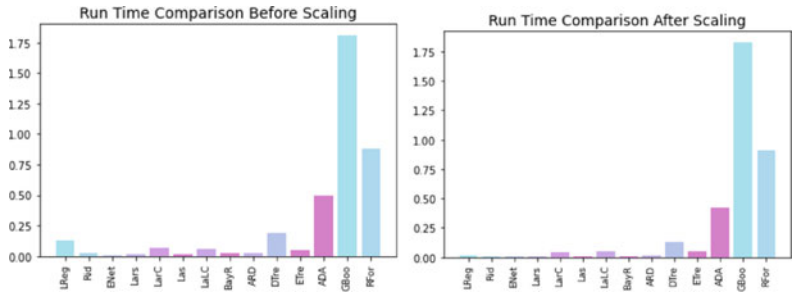



**Fig. 4** Response time analysis of raw dataset before and after feature scaling

## *4.3 Quantitative Analysis of MLP Based Neural Network Model*

The dataset is fitted with MLP neural network based regressor with different activation layers like Relu, identity, Logistic and tanh and the evaluation metric for the MLP neural network before and after feature scaling and is shown in Tables 3 and 4.

## *4.4 Comparative Analysis of MLP Based NN Model with Existing Models*

This proposed model Multilayer perceptron (MLP) based Neural Network is compared with existing models using MAE, MSE and RScore evaluation metric and is shown in Table 5.

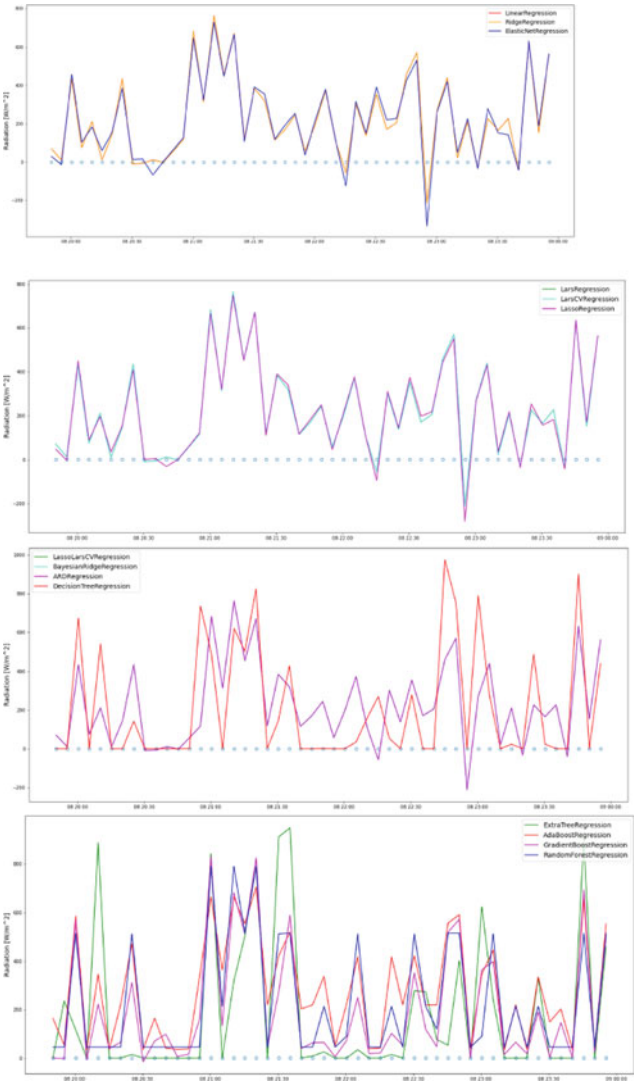



**Fig. 5** Testing versus Prediction dataset of the various regressors

**Table 3** Evaluation metric for MLP neural network performance dataset before scaling

| MLP regressor | EVS | MSE | MAE | RScore | Running time (ms) |
|---|---|---|---|---|---|
| Relu | 0.6666 | 0.000337967 | 1.236305 | 0.6662 | 50.0926 |
| Identity | 0.5533 | 0.000452422 | 1.621438 | 0.5532 | 9.3057 |
| Logistic | 0.6779 | 0.000326444 | 1.157363 | 0.6776 | 122.9012 |
| Tanh | 0.6807 | 0.000323257 | 1.153343 | 0.9907 | 97.4085 |

**Table 4** Evaluation metric for MLP neural network performance of dataset after scaling

| MLP regressor | EVS | MSE | MAE | RScore | Running time (ms) |
|---|---|---|---|---|---|
| Relu | 0.6856 | 0.000318462 | 1.110951 | 0.6855 | 98.2089 |
| Identity | 0.5659 | 0.000439755 | 1.576327 | 0.5657 | 6.9842 |
| Logistic | 0.6618 | 0.000342664 | 1.199573 | 0.6616 | 120.9924 |
| Tanh | 0.6993 | 0.000304608 | 1.079012 | 0.9992 | 163.4098 |

**Table 5** Evaluation metric comparative analysis for MLP neural network performance

| Regressor model | MSE | MAE | RScore |
|---|---|---|---|
| ANN model I – GD [2] | 11.2127 | 8.6860 | 0.5202 |
| ANN model I – LM [2] | 4.5497 | 3.4979 | 0.9223 |
| ANN model I – SCG [2] | 5.1477 | 5.1477 | 0.7990 |
| ANN model I – RP [2] | 5.0187 | 5.0187 | 0.7616 |
| ANN model II – GD [2] | 10.7656 | 9.0903 | 0.4844 |
| ANN model II – LM [2] | 3.6461 | 3.0281 | 0.9272 |
| ANN model II – SCG [2] | 4.0794 | 3.3172 | 0.8332 |
| ANN model II – RP [2] | 4.9601 | 3.9474 | 0.5852 |
| Multilayer feed forward NN [4] | 0.0508 | 0.0352 | 0.9488 |
| Square root smoothly clipped [5] | 0.0666 | 20.210 | 0.9800 |
| Wavelet-coupled SVM [14] | 0.05942 | 0.04696 | 0.9650 |
| Proposed MLP regressor RELU | 0.00031 | 1.110951 | 0.6855 |
| Proposed MLP identity | 0.00043 | 1.576327 | 0.5657 |
| Proposed MLP logistic | 0.000342 | 1.199573 | 0.6616 |
| Proposed MLP regressor tanh | 0.000304 | 1.079012 | 0.9992 |

## 5 Conclusion

This paper performs the exploratory data analysis of the solar radiation dataset and also explores the correlation between the features. Dataset is fitted with all regressors to analyze the performance in terms of MAE, MSE, EVS and RScore. Experimental results show that the Gradient boost regressor have the RScore of 0.98 before and after feature scaling. The MLP regressor with TANH activation layer tends to retain 0.99 Rscore before and after feature scaling.

## References

1. Innani A, Dhatwalia S, Tripathi Y, Jambhulkar V (2020) Solar power predictor using ensemble learning. Int Res J Eng Technol 7(4)

2. Neelamegama P, Amirthamb VA (2016) Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. J Appl Res Technol 14:206–214
3. Sharma P, Kamble M (2019) Clustering model for solar irradiation prediction using machine learning algorithm. Int J Sci Technol Res 8(12)
4. Guher AB, Tasdemir S, Yaniktepe B (2020) Effective estimation of hourly global solar radiation using machine learning algorithms. Int J Photoenergy 2020:1–26
5. Dong Y, Jiang H (2019) Global solar radiation forecasting using square root regularization-based ensemble. Math Prob Eng 2019:1–20
6. Litta AJ, Idicula SM, Mohanty UC (2013) Artificial neural network model in prediction of meteorological parameters during premonsoon thunderstorms. Int J Atmos Sci 2013:1–14. https://doi.org/10.1155/2013/525383
7. Jeon B-K, Kim E-J (2020) Next-day prediction of hourly solar irradiance using local weather forecasts and LSTM trained with non-local data. Energies 13(20):1–16
8. Gagne D, Brotzge J, Mcgovern A (2015) Day-ahead hail prediction integrating machine learning with storm-scale numerical weather models. ResearchGate
9. Senthil Kumar P (2019) Improved prediction of wind speed using machine learning. EAI Endorsed Trans Energy Web 6(23):157033
10. Kashiwaoa T, Nakayamaa K, Andoc S, Ikedad K, Leee M, Bahadori A (2015) A neural network-based local rainfall prediction system using meteorological data on the internet: a case study using data from the Japan Meteorological Agency. ScienceDirect
11. Demirtas M, Yesilbudak M, Sagiroglu S, Colak I (2012) Prediction of solar radiation using meteorological data. In: The proceedings of international conference on renewable energy research and applications, pp 1–4
12. Premalatha N, Arasu AV (2016) Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. J Appl Res Technol 14(3):206–214
13. Voyant C, Notton G, Kalogirou S et al (2017) Machine learning methods for solar radiation forecasting: a review. J Renew Energy 105:569–582
14. Deo RC, Wen X, Qi F (2016) A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. Appl Energy 168:568–593
15. Faceira J, Afonso P, Salgado P (2015) Prediction of solar radiation using artificial neural networks. In Proceedings of the 11th Portuguese conference on automatic control. Lecture notes in electrical engineering, vol 321

# Myocardial Infarction Analysis Using Deep Learning Neural Network Based on Image Processing Approach


**G. Rajakumar, V. Nagaraju, B. R. Tapas Bapu, P. Stella Rose Malar, R. Santhana Krishnan, and K. Lakshmi Narayanan**

**Abstract** In recent decades, the health monitoring system has been an attractive research topic. In order to diagnose Cardiovascular Disease (CVD), Electrocardiogram (ECG) is one of the popular instruments. In recent times, ECG monitoring systems are exponentially increasing. As a result, it is very difficult for the scientists and health experts to choose, compare and evaluate the systems, and identify which will meet their needs and necessary monitoring. In this research article, we offer a global taxonomy and review from an expert of ECG surveillance systems. It provides support, which is an evidence to monitor components, contexts, functions and challenges. For monitoring the ECG systems, an architectural model is proposed in this paper and provides in-depth review. In this research paper, an adaptive medium filter is used for signal preprocessing and the segmentation is accomplished using watershed transform. Finally, feature extraction is performed using several texture features and signal classification is accomplished using probabilistic neural networks (PNN).

**Keywords** Cardiovascular disease · Electrocardiogram · Health monitoring system · Machine learning · Probabilistic neural network

G. Rajakumar
Department of ECE, Francis Xavier Engineering College, Tirunelveli, India

V. Nagaraju
ECE, Saveetha School of Engineering, Chennai, Tamilnadu, India
e-mail: nagarajuv.sse@saveetha.com

B. R. T. Bapu
ECE, S.A. Engineering College, Chennai, Tamilnadu, India

P. S. R. Malar
ECE Department, J.P. College of Engineering, Tenkasi, Tamilnadu, India

R. S. Krishnan
SCAD College of Engineering and Technology, Tirunelveli, Tamilnadu, India

K. L. Narayanan (✉)
Francis Xavier Engineering College, Tirunelveli, Tamilnadu, India
e-mail: klnarayanan@francisxavier.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_59

## 1 Introduction

Myocardial Infarction (MI) is also called heart attack which occurs due to insufficient supply of blood flow to the heart muscles [1–4]. The symptoms of Myocardial Infarction consist of pain in the chest, jaw, arm. That Pain is first originating at the center or from left of the chest and radiates to the arm, jaw, lower back, or shoulder. Some of the patients may have symptoms such as vomiting, nausea and chronic breathing issues. Sometimes patients do not have any above symptoms [6, 7]. In such cases, MI can be detected mostly by means of cardiac imaging or other such techniques. The failure of detecting myocardial infarction is a major problem because when the coronary heart cells die, they cannot be replaced with new cells. Machine Learning (ML) helps to predict MI at the earlier stage, thereby we can prevent damage occurring in the heart [8–10].

## 2 Related Works

Petmezas et al. [11] developed a new hybrid model using focal loss in order to deal with data imbalance problems. Initially, Convolutional Neural Network (CNN) was used for feature extraction and then long short term memory network was used for classifying four ECG rhythm types such as junctional rhythm, normal, atrial flutter and atrial fibrillation. Malik et al. [12] used different types of features and kernels of Support Vector Machine (SVM) for ECG signal classification. Shaker et al. [13] developed a data augmentation method named generative adversarial network for restoring the dataset balance. Dias et al. [14] has used three combinations of features such as higher order statistics, RR intervals, and signal morphology for data classification on the MIT-BIH dataset. Fatimah et al. [15] initially divides ECG signal into one minute segments and then a Fourier decomposition method was used for separating the frequency bands. Then, feature extraction was accomplished by entropy, and mean absolute deviation for classifying the ECG segments.

## 3 Proposed Method

The proposed method consists of Adaptive - Median filtering which is used to preprocess the input image. The Watershed segmentation algorithm is used in the segmentation process. The GLCM algorithm is used in the feature extraction process. Probabilistic Neural Network classifier is used to classify ECG signals.

### 3.1 Medical Techniques for DETECTING MI

Pathological methods and imaging methods are the two medical methods to detect MI. Some imaging techniques to detect MI include Echocardiography, Angiography and ECG.

### 3.2 ECG

Electrocardiography is a commonly used technique of image processing [5]. This technique helps to identify the causes of chest pain which is caused other than MI, such as unrestricted flow of blood and valvular heart disease, thereby we can avoid misdiagnosis. In this process, the electrodes are placed on the patient's chest and limbs. From which the electrical impulse of the heart is calculated and then on the screen or a paper, a wavy line is represented which indicates the heart's rhythm and the weaknesses on the different parts. An ECG wave has 3 important components. Those include P wave, QRS complex and T wave. The atria contraction is represented by P wave, the depolarization of ventricles is measured by QRS complex and the repolarization of ventricles is measured by T wave. The period between polarization and depolarization of ventricles is represented by the ST segment. The total duration of ventrial recovery is specified by the combination of T wave and U wave. The time period between the QRS complex is the R-R interval. By using the time period of two QRS complex, instantaneous heart rate can be calculated. Block diagram of the proposed system is given in Fig. 1.

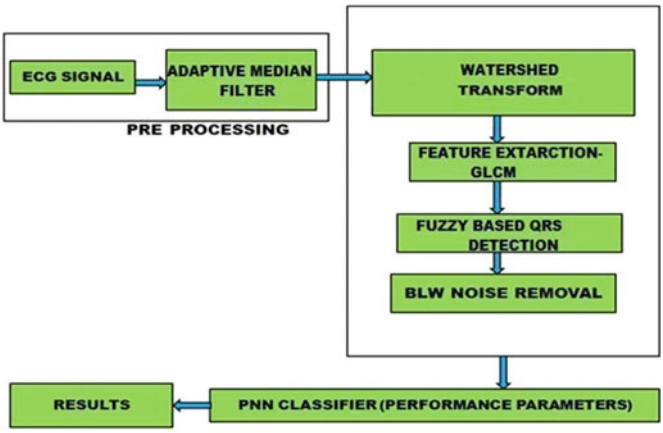



**Fig. 1** Block diagram of proposed system

### 3.3 Pre-processing

The preprocessing prepares the data for data analysis. To eliminate the effects of the detector, the data has to be pre-processed before starting the actual process. The most important aspect of data processing is pre-processing. As the result of an experiment, the particular data is acquired from which we will extract the useful information. Data processing classifies the data into three types and processes it accordingly. Therefore, data filtering, data ordering, data editing, and noise modeling plays an important role in this process. Additionally, it helps to convert data format from real to integer representation and retain the sensitivity and then the output is processed. In the compression system of system development, few important compression techniques for ECG classification rely on regions of interest. These signals are generally of reduced contrast and noisy nature. So denoising of signal is needed to uphold the quality of signal by noise suppression. Quality of signal and feature extraction algorithm becomes unreliable due to noise. The denoising and feature extraction method implemented in this research will enhance the signal processing reliability.

### 3.4 Segmentation

When processing digital signals and the view of the computer, the segmentation of the signal is the method of subdividing a digital signal in several segments (pixel arrangements, also known as signal objects). The character segmentation is normally used to locate objects and limits (lines, curves, etc.) The signal will therefore analyze, becoming easier. We use different signal segmentation algorithms to divide and group a particular set of pixels together from the signal. Actually, we point to pixels to pixels, and pixels with the same label fall under a category where it has any or the other thing in them. The concept of partition, separation, collected and then the labeling and then the use of this information to train several ML models, in fact, numerous business problems have been undergone. From the topographic analogy, the idea of watershed is derived. For calculating the watersheds, the first algorithm is in the field of surveying. In grayscale segmentation problems, watershed segmentation is used. Feature outlines cannot be produced by the basin transform applied signal. Rather, the signal is divided into the associated areas by using intensity gradient and treats that signal, where the altitude is denoted by intensity of pixel. After completion, the signal basin is defined by the resulting dam network.

### 3.5 GLCM Feature Extraction

There are many methods for feature extraction. GLCM is one of those methods. Here for feature extraction, GLCM technique is used. Based on the signal histogram, the

texture filter function provides the statistical view of texture. The useful information about the texture of a signal is provided by the functions but it cannot provide any information about its shape. Some of the features; mean, standard deviation, variance, contrast, and correlation are mathematically depicted in the Eqs. (1–5).

$$Mean(\mu_i) = \sum_{i,j=0}^{N-1} i\left(P_{i,j}\right), \mu_j = \sum_{i,j=0}^{N-1} j(P_{i,j}) \tag{1}$$

$$Standard\ deviation(\sigma_i) = \sqrt{\sigma_i^2}, \sigma_j = \sqrt{\sigma_j^2} \tag{2}$$

$$Varaince(\sigma_i^2) = \sum_{i,j=0}^{N-1} P_{i,j}(i-\mu_i)^2, \sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j}(j-\mu_j)^2 \tag{3}$$

$$Contrast = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \tag{4}$$

$$Correlation = \sum_{i,j=0}^{N-1} P_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right] \tag{5}$$

### 3.6 Classification

The intention of the classification process is to categorize all pixels in a digital signal in one of the different types of terrestrial coverage or "topics". These categorized data can be used to make thematic cards of the ground cover available in a signal. Character Classification is a complex method based on different components. Here are some of the strategies presented, additional questions and possibilities of signal orders. The main focus is in avant-garde classification methods used to improve the accuracy of characterization. The Probabilistic Neural Network (PNN) is a neuronal neural network. For the classification and pattern recognition problems, PNN is most commonly used. The higher level probability distribution function (PDF) of each class approaches a packet window and a non-parametric function in the PNN algorithm. The assistance of new input data is estimated using a PDF of each class and the rule is used to assign the class with the highest backup probability for the new input data. This method minimizes the probability of errors in the classification.

## 4 Result and Discussion

The discussed ECG classification method in this paper is implemented in MATLAB to analyze ECG signals. The experiments for the proposed methodology are carried out and validated using the MIT-BIH dataset. The expected performance for the

sample ECG signals at each subsequent stages of the proposed methodology is exhibited for detailed analysis. To eliminate all these noises, orientation- specific encoding schemes like adaptive median filter is used for analyzing the texture features of ECG signal. Analogous to input signal, the output of filter is more precise and accurate. The input and pre-processed signals are represented in Figs. 2 and 3.

For further processing with minimum data redundancy and to constraint the dataset integration, the filtered output is normalized. The signal ECG is considered as projection of the heart's electrical vector on the corresponding lead vector as a time function (amplified by the absolute magnitude of the lead vectors). It is depicted in Fig. 4.

Generally, the coefficients are dispersed based on the bandwidth. The energies in the ECG signal is gathered together using watershed so as to represent the most

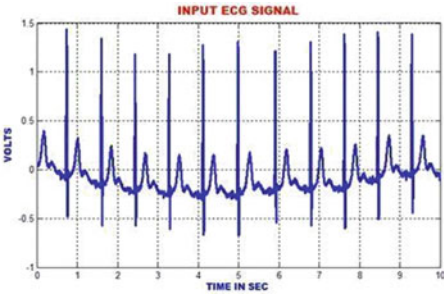



**Fig. 2** Input ECG signals

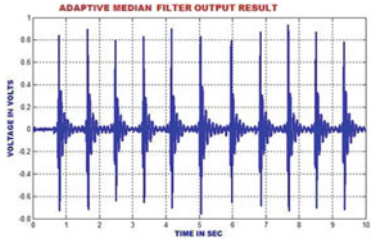



**Fig. 3** Adaptive median filter output

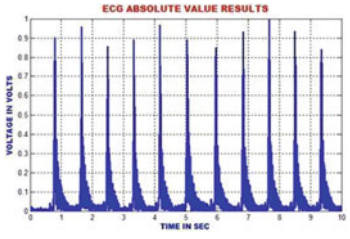



**Fig. 4** ECG absolute value results

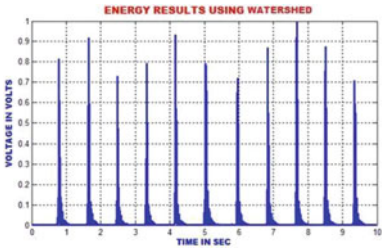



**Fig. 5** Output of watershed approach

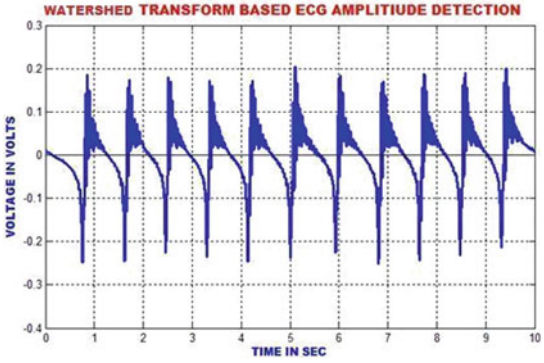



**Fig. 6** Segmented Output

important coefficient at the low frequency. The features that are extracted using the watershed approach indicates the time-recurrence attributes of the ECG signal and are unsymmetrical in nature. Also the peak values in QRS polarity and the unexpected variations in QRS amplitude are detected. It is depicted in Figs. 5 and 6.

The advantage of fixing the length of each heart occasion is to find the R-top precisely compared with the P and T waves since they have low magnitude and are sensitive to the commotion. These unbalanced time–recurrence coefficients are needed to be processed for the ECG signal so as to represent their morphological qualities, which are used for further examination. The moving average filter is dedicated to removing high frequency noises from the ECG signal by computing the running mean on the predetermined window length. This is a moderately straightforward estimation which will smoothen both the signal and its anomalies. The R-top in the ECG sign is smoothed to around 33% of its unique height. The low frequency contents of the ECG signal are represented in the Fig. 7.

The R top in the QRS interim is the most significant component for examining the ECG signal. R top discovery in ECG is a strategy that is generally used to analyze heart anomalies and gauge pulse fluctuation. It is natural that the magnitudes of genuine R tops are more than those for bogus pinnacles. The QRS detection ensures the efficient extraction of beat intervals and the abnormalities in the heart function.

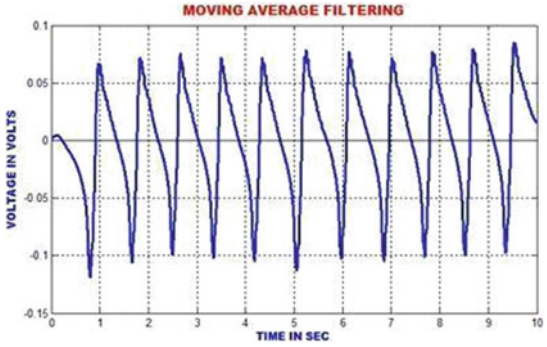



**Fig. 7** Feature extraction output

The improvements in the QRS sections are executed by the proposed technique to eliminate the pattern meandering, which is depicted in Figs. 8 and 9.

The RR-interim is resolved to extract the dynamic qualities of the ECG signal. The 4 RR attributes that are discussed in this paper are pre-RR, post-RR, neighborhood RR, and mean RR interim. The interim between a past R-top and the present R-top is processed to find the pre-RR attribute, while the interim between a specific R-top and the successive R-top is estimated to find the post-RR highlight. By using the

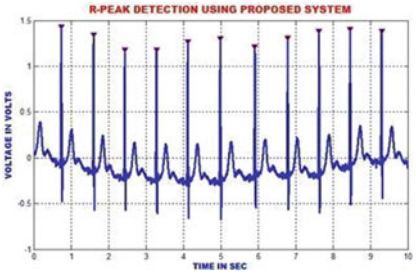



**Fig. 8** R peak detection using the proposed system

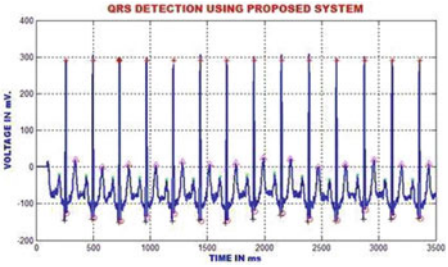



**Fig. 9** Detection of QRS using the proposed system

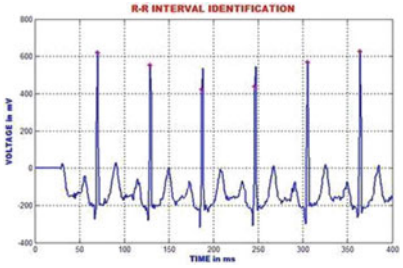



**Fig. 10** Identification of R-R interval

**Table 1** Comparative analysis

| Methods | Accuracy (%) | Sensitivity (%) |
|---|---|---|
| SVM [12] | 92 | 86 |
| CNN [13] | 90 | 92 |
| **PNN** | **98.5** | **98.3** |

results of all the processes and the final image obtained from the PNN classifier, the output indicating the results of ECG is obtained. The RR interval is stated in Fig. 10.

The performance of the PNN is compared with the existing classifiers such as CNN [13] and SVM [12]. After completing a number of test samples, this method shows greater performance with a highest accuracy of 98.5%. The reliability of this method is consistently high. The true positive value of the ECG classification is indicated by the sensitivity of this method. The sensitivity performance of this method is more than the existing method with a highest value of 98.3% whereas the maximum sensitivity of CNN and SVM lies at 92% and 86%. The comparative analysis is given in Table 1.

## 5 Conclusion

The diagnosis of myocardial infarction at the earlier stage is important to reduce mortality rate in our country. For diagnosing the several types of heart attacks, many researchers have focused on 12 lead ECG. As always, multiple lead ECG devices are available, it will be incorporated only in hospitals and clinics. In recent years, for prevention and monitoring of myocardial infarction the prevalence of portable ECG test equipment plays an important role. This document suggested a model for the classification of myocardial infarction ECG, based on PNN. The structure of the network had deep structural features that could acquire the spatial and temporal properties of ECG signals. Therefore, it is an effective method for the automatic classification of myocardial infarction ECG.

## References


1. Sun Z, Wang C, Zhao Y, Yan C (2020) Multi-label ecg sign, p 117996
2. Zhai X, Tin C (2018) Automated ECG classification using dual heartbeat coupling based on convolutional neural network. IEEE Access 6:27465–27472
3. Kiranyaz S, Ince T, Gabbouj M (2016) Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Trans Biomed Eng 63(3):664–675
4. Rajendra AU (2017) A deep convolutional neural network model to classify heartbeats. Comput Biol Med 89:389–396
5. Martis RJ, Rajendra Acharya U, Min LC (2013) ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. Biomed Signal Process Control 8(5):437–448
6. Ansari S, Farzaneh N, Duda M, Horan K, Andersson HB, Goldberger ZD, Nallamothu BK, Najarian K (2017) A review of automated methods for detection of myocardial ischemia and infarction using electrocardiogram and electronic health records. IEEE Rev Biomed Eng 10:264–298. https://doi.org/10.1109/RBME.2017.2757953
7. Lee T-S, Tsui BMW (2019) Task-based evaluation of image reconstruction methods for defect detection and radiation dose reduction in myocardial perfusion SPECT. IEEE Trans Radiat Plasma Med Sci 3(1):89–95
8. Pontre B, Cowan BR, DiBella E, Kulaseharan S, Likhite D, Noorman N, Tautz L, Tustison N, Wollny G, Young AA, Suinesiaputra A (2017) An open benchmark challenge for motion correction of myocardial perfusion MRI. IEEE J Biomed Health Inf 21(5):1315–1326. https://doi.org/10.1109/JBHI.2016.2597145
9. Liu W, Zhang M, Zhang Y, Liao Y, Huang Q, Chang S, Wang H, He J (2018) Real-time multilead convolutional neural network for myocardial infarction detection. IEEE J Biomed Health Inf 22(5):1434–1444
10. Scannell CM, Villa ADM, Lee J, Breeuwer M, Chiribiri A (2019) Robust non-rigid motion compensation of free-breathing myocardial perfusion MRI data. IEEE Trans Med Imaging 38(8):1812–1820
11. Petmezas G, Haris K, Stefanopoulos L, Kilintzis V, Tzavelis A, Rogers JA, Katsaggelos AK, Maglaveras N (2021) Automated atrial fibrillation detection using a hybrid CNN-LSTM network on imbalanced ECG datasets. Biomed Signal Process Control 63:102194
12. Malik GK, Kumar Y, Panda M (2021) Multi-kernel SVM approach for arrhythmias classification. In: Proceedings of integrated intelligence enable networks and computing. Springer, Singapore, pp 733–739
13. Shaker AM, Tantawi M, Shedeed HA, Tolba MF (2020) Generalization of convolutional neural networks for ECG classification using generative adversarial networks. IEEE Access 8:35592–35605
14. Dias FM, Monteiro HL, Cabral TW, Naji R, Kuehni M, Luz EJDS (2021) Arrhythmia classification from single-lead ECG signals using the inter-patient paradigm. Comput Methods Programs Biomed 202:105948
15. Fatimah B, Singh P, Singhal A, Pachori RB (2020) Detection of apnea events from ECG segments using Fourier decomposition method. Biomed Signal Process Control 61:102005

# Novel Single CDDITA Based Resistively Tunable All-Pass Filter Configuration with Grounded Passive Elements


**Priyanka Joshi, Kapil Bhardwaj, and Mayank Srivastava**

**Abstract** The design and analysis of a new voltage-mode all pass filter developed using CDDITA (Current Differencing Differential Input Trans conductance Amplifier) along with two grounded resistances and one grounded capacitor has been depicted in this paper. As the detailed literature survey has revealed, proposed all pass filter (APF) configuration is the most compact filter design as compared to any CDDITA based APF presented so far. The proposed configuration has the advantages as; employment of all passive grounded elements, availability of gain control through grounded resistance, low input impedance and high output impedance. For the evaluation of simulation results of the developed configuration the passive component is taken as: $R_1 = R_2 = 1\,\mathrm{K}\Omega$, $C1 = 0.1\,\mathrm{nF}$ with supply voltages $\pm 3$ V DC and bias current values $I_{bias} = 20\,\mu\mathrm{A}$. The operability of the circuit has been validated through simulation performed using PSPICE OrCAD 9.1 Version using TSMC technology parameters.

**Keywords** All pass filter · CDDITA · Resistance control · Voltage-mode all-pass filter

## 1 Introduction

The APFs are the prime circuits in analog signal processing field, and generally find its application in phase equalization as well as for introducing a frequency dependent delay with input signal amplitude kept constant over desired range of frequencies. In many other filters the applied input signal's magnitude changes for some frequency value but in APF the magnitude remains the same throughout the range of frequencies. The conventional passive APF has the disadvantages such as source loading occurs; circuit is large in size due to the large value of resistor and capacitor used in the filtering circuit; due to the use of inductor the circuit becomes bulky; the circuit

P. Joshi (✉) · K. Bhardwaj · M. Srivastava
NIT Jamshedpur, Jamshedpur, India
e-mail: priajoshi31@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_60

is even tuned for a fixed frequency, etc. To overcome all these drawbacks active realization of the filtering circuit is performed.

First order Voltage Mode-APFs (VM-APFs) using Active building block has been proposed earlier in many works. However, most of the circuits suffer from several disadvantages among which use of floating capacitance is found to be inevitable. The encouragement behind this paper is to configure CDDITA based first-order VM -APF having minimum active and passive components with the feature of tunability. The CDDITA has been used in several analog signal generation/processing circuits as a building block for circuit realization [1–8].

## 2 Related Work

Previously different ABBs have been used in the design of Voltage mode and current mode APFs such as given in [9–37]. In these filters, the configurations reported in [9, 26–37] realize CM-APFs and others are voltage mode APFs [10–25]. On closely investigating these previously proposed VM-APF circuit configurations [10–25], it is found the circuits suffered with several disadvantages as follows: (1). In many work there was more than one active element used in the circuit realization. (2). More than one passive elements used in the circuit. (3). Floating capacitors are used in the circuits. (4). Many circuits are not resistor tunable.

The proposed configuration uses minimal number of components: one ABB, one grounded capacitor and two grounded resistors. The modified CDDITA called as the Z-copy CDDITA (ZC-CDDITA) is used in the proposed work. The current obtained from zc (Z-copy) terminal is of the same magnitude as z terminal with current output in reverse direction. Following advantages are obtained from the new proposed VM-APF configuration: (1). Grounded capacitor used in the circuit makes it apt for monolithic integration because the use of grounded capacitor configuration can recompense the stray capacitance at the nodes it's connected. (2). Cascading of the circuit with low input and high output impedance becomes convenient for higher order filter synthesis. (3). At pole frequency good active and passive sensitivity can be found.

## 3 Proposed Work

### *3.1 CDDITA Circuit Idea*

The CDDITA first reported in [38], is a popular active building block used in numerous circuits. In Fig. 1 CDDITA block diagram depicted and Fig. 2 depicts the behavioral model of CDDITA depicts it consist of current differencing unit (CDU) following this is a trans conductance amplifier. Ideally, the CDDITA offers single

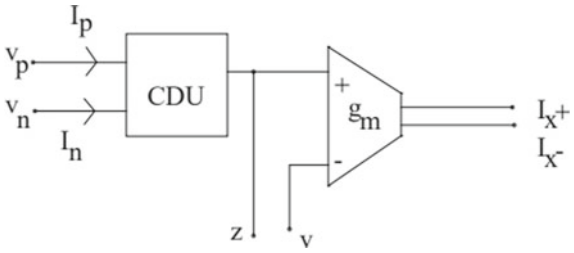
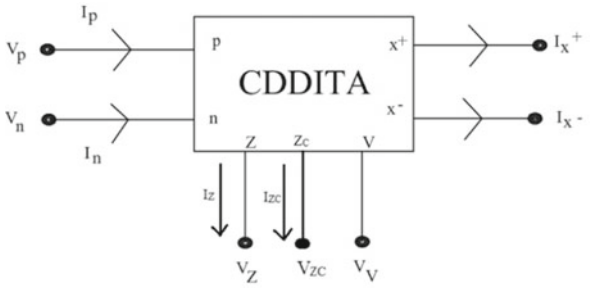
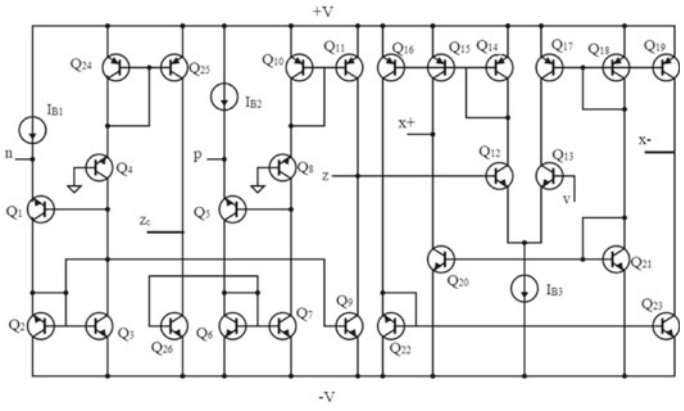



**Fig. 1** Behavioral model of an ideal CDDITA




(a)




(b)

**Fig. 2** **a** Symbolic depiction of ZC-CDDITA **b** Transistor implementation of ZC-CDDITA

positive $z$ output obtained from the first trans conductance stage. In this work, we have also extracted the $Z$—port of the CDDITA to achieve the negative $Z$ current output. Figure 3 depicts the CMOS implementation of the ZC-CDDITA. It has input ports $P$ and $N$, $X+$ and $X$—are output ports, $Z$ and $Z_c$ are the auxiliary ports and $V$ is buffered port. In CDDITA except $P$ and $N$ all the other ports are high impedance

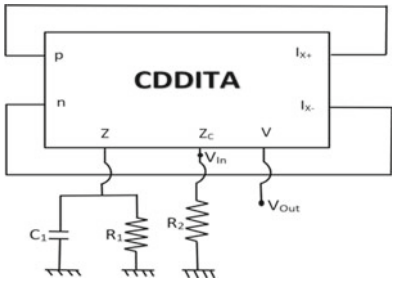



**Fig. 3** All pass filter configuration

port. The voltage-current relationship between all the ports of CDDITA is given Eq. (1) to (4):

$$V_p = V_n = 0 \tag{1}$$

$$I_z = \left(I_p - I_n\right) \tag{2}$$

$$I_{x+} = g_m(V_z - V_v) \tag{3}$$

$$I_{x-} = -gm(V_z - V_v) \tag{4}$$

## *3.2 Proposed CDDITA All-Pass Filter*

The designed first order VM-APF configuration is depicted in Fig. 3.

$$\frac{\boldsymbol{V_{Out}(s)}}{\boldsymbol{V_{in}(S)}} = \frac{1}{2}\left(\frac{\boldsymbol{sC_1 - g_m}}{\boldsymbol{sC_1 + g_m}}\right) \tag{5}$$

Where,

$$\frac{1}{R_1} = \frac{1}{R_2} = g_m \tag{6}$$

It can be clearly depicted from Eqs. (5) and (6) that the voltage gain of circuit implemented in Fig. 3 is unity.

The phase response and the pole frequency obtained from Eqs. (7) and (8) are as follows,

$$\angle \frac{V_{Out}(s)}{V_{in}(S)} = \Phi(\omega_p) = \pi - 2tan^{-1}\left(\frac{\omega C_1}{g_m}\right) \tag{7}$$

$$\omega_o = \left(\frac{g_m}{C_1}\right) \tag{8}$$

Here, $\omega_o$ is the angular frequency. The angular pole frequency can be electronically tuned through trans conductance $g_m$.

### 3.3 Non-Ideal Analysis of Proposed Configuration

In non-ideal condition the voltage-current relationship of CDDITA is defined by the following Eqs. (9) to (13):

$$I_z = \left(\alpha_p I_p - \alpha_n I_n\right) \tag{9}$$

$$I_{x+} = \beta_{+gm}(V_z - V_v) \tag{10}$$

$$I_{x-} = \beta_{-gm}(V_z - V_v) \tag{11}$$

Here $\alpha_p$, $\alpha_n$ are the current transfer error to $z$ terminal from $p$ and $n$ terminals respectively and $\beta+$, $\beta-$ are the trans conductance error gain in $z$ and $zc$ terminals respectively. On considering the above non-ideal condition, the transfer function of VM-APF depicted in Fig. 4 is shown below;

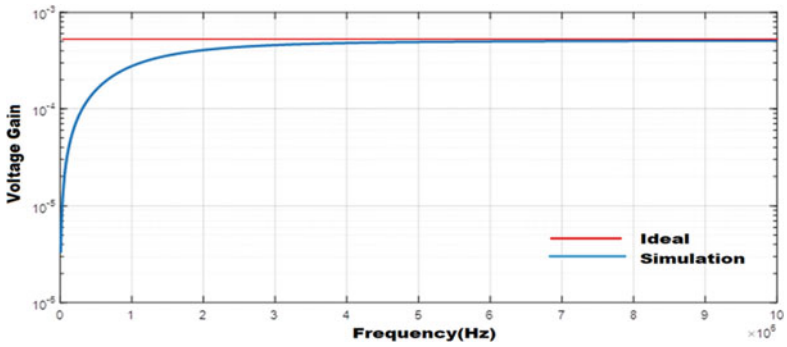



**Fig. 4** Magnitude (Gain) vs. frequency plot of proposed APF configuration along with the comparison shown with conventional RLC APF filter response

$$\frac{V_{Out}(s)}{V_{in}(S)} = \left( \frac{sR_1 + g_m - g_m\left(\alpha_p\beta_+ + \alpha_n\beta_-\right)}{(sR_1C_1 + g_m)\left(\alpha_p\beta_+ + \alpha_n\beta_-\right)} \right) \tag{12}$$

$$\angle\frac{V_{Out}(s)}{V_{in}(S)} = \pi - tan^{-1}\left( \frac{\omega C_1}{g_m\left(\alpha_p\beta_+ + \alpha_n\beta_-\right)} \right) - tan^{-1}\left( \frac{\omega C_1}{g_m} \right) \tag{13}$$

It's obvious from Eqs. (12) and (13), the gain and even the phase of the VM-APF is affected by the current transfer error as well as by the trans conductance error gain and thus a good CDDITA design should be contemplated to reduce the effect of non-ideal parameters. The sensitivity of angular zero frequency (obtained using Eq. 13) to the non-ideality factors and externally connected elements are shown in Eq. (14):

$$S^{\omega_z}_{\alpha_n,\alpha_p,\beta_+,\beta_-} = 0,\ S^{\omega_z}_{g_m} = 0,\ S^{\omega_z}_{C_1} = 0 \tag{14}$$

Similarly, the sensitivity of angular pole frequency to the non-ideality factors along with externally connected elements are shown in Eq. (15):

$$S^{\omega_p}_{g_m} = 1,\ S^{\omega_p}_{C_1} = -1,\ S^{\omega_p}_{\alpha_{p,\beta_+}} = \frac{\alpha_p\beta_+}{2(\alpha_p\beta_+ + \alpha_n\beta_- - 1)},$$
$$S^{\omega_p}_{\alpha_{p,\beta_+}} = \frac{\alpha_n\beta_-}{2\left(\alpha_p\beta_+ + \alpha_n\beta_- - 1\right)} \tag{15}$$

## 4 Results and Discussion

The reported circuit has been verified by simulating CMOS implementation of CDDITA using PSPICE OrCAD 9.1 Version (180 μm Technology) (Fig. 3). And the transistor model parameters have been taken as PR100 N(PNP (IS = 73.5E−18, BF = 110, VAF = 51.8, IKF = 2.359E−3, ISE = 25.1E−16, NE = 1.650, BR = 0.4745, VAR = 9.96, IKR = 6.478E−3, RE = 3, RB = 327, RBM = 24.55, RC = 50, CJE = 0.18E−12, VJE = 0.5, MJE = 0.28, CJC = 0.164E−12, VJC = 0.8, MJC = 0.4, XCJC = 0.037, CJS = 1.03E−12, VJS = 0.55, MJS = 0.35, FC = 0.5, TF = 0.610E−9, TR = 0.610E−8, EG = 1.206, XTB = 1.866, XT1 = 1.7)) and NR100 N(NPN (IS = 121E−18, BF = 137.5, VAF = 159.4, IKF = 6.974E−3, ISE = 36E−16, NE = 1.713, BR = 0.7258, VAR = 10.73, IKR = 2.198E−3, RE = 1, RB = 524.6, RBM = 25, RC = 50, CJE = 0.214E−12, VJE = 0.5, MJE = 0.28, CJC = 0.983E13, VJC = 0.5, MJC = 0.3, XCJC = 0.034, CJS = 0.913E−12, VJS

= 0.64, MJS = 0.4, FC = 0.5, TF = 0.425E−8, TR = 0.5E−8, EG = 1.206, XTB = 1.538, XT1 = 2)).

***Comparison of Configured VM-APF with Previously Reported Works***

The Table 1 depicts a detailed comparison of some popular all pass filters reported previously considering different design aspects.

For the evaluation of simulation results of the developed configuration reported in Fig. 3 the passive component is taken as: $R_1 = R_2 = 1\,\text{K}\Omega$, $C_1 = 0.1\,\text{nF}$ with supply voltages ±3 V DC and bias current values $I_{bias} = 20\,\mu\text{A}$.

The simulation result of magnitude of the gain has been depicted in Fig. 4, showing the comparative analysis of the conventional passive APF with the proposed one. The parameter $R = 2.5\,\text{k}\Omega$ and $C = 0.85\,\text{nF}$ has been chosen to get the desired result of the passive APF configuration. Also, the Bandwidth of the proposed configuration is approx. $10^7$ Hz which is greater than the circuit reported in previous literature [10, 21–23] and the configuration is linear throughout the range. The proposed configuration is Resistor tunable which is not available in the configuration reported in [10, 21–24]. And coming to performance comparison, it is important to discuss that for different

**Table 1** Comparative analysis of VM-APFs given in [10–25] using different active building blocks

| Ref. no | Active elements used | Passive elements used | | Demonstration of electronic tunability |
|---|---|---|---|---|
| | | Capacitor | Resistor | |
| [10] | 1-LTI228 | 2-F* | 1-F | No |
| [11] | 1-DOCCTA | 2-G** | 1C-F | No |
| [12] | 1-DDCC | 3(2-G,1-F) | 1-G | Yes |
| [13] | CDBA | 2-G | 1-F | Yes |
| [14] | CDBA | 2-F | 2(1-F,1-G) | Yes |
| [15] | 1-UVC | 2-G | 1-F | Yes |
| [16] | 1-CCII | 1-G | 1-F | Yes |
| [17] | 1-CCII | 2-G | 1-F | Yes |
| [18] | 1-CCII | 4(3-F, 1-G) | 1-G | Yes |
| [19] | 2-CCII | 2-G | 2-G | Yes |
| [20] | 1-CCII | 2-F | 1-F | No |
| [21] | 1-FTFN | 4(1-G, 3-F) | 4(1-G,3-F) | No |
| [22] | 1-FOC | – | 1-G | No |
| [23] | 2-CDBA | 4(3-G, 1-F) | 2(1-G,1-F) | No |
| [24] | 1-MO-ICCII<br>1 MO-CCII | 1-G | 1-G | No |
| [10] | 2-LT1228 | 2-F | 2(1-G, 1-F) | No |
| [25] | 1-CDDITA | 2-F | 1-F | Yes |
| Proposed | 1-CDDITA | 2-G | 1-G | Yes |

* F—Floating, ** G—Grounded

parameter values, the gain, bandwidth and other quantities of the active filter vary and therefore, two active filter performance cannot be compared directly.

The phase response of gain vs. frequency is depicted in Fig. 5, confirming that the proposed configuration is a VM-APF.

The variation in the realized value of gain on selecting different values of capacitance $C_1$ (keeping $R_1 = R_2 = 1\,\mathrm{K}\Omega$ and bias current $I_{bias} = 20\,\mu\mathrm{A}$) are shown in Fig. 6. On increasing the value of capacitor $C_1$ the all pass filter magnitude vs. frequency curve shifts downwards.

Additionally, the resistor tuning of realized all pass filter through grounded resistance $R_1$ (keeping $R_2 = 1\,\mathrm{K}\Omega$, $C_1 = 0.1\,\mathrm{nF}$ and $I_{\mathrm{bias}} = 20\,\mu\mathrm{A}$) is depicted in Fig. 7. On increasing the resistance $R_1$ value, the magnitude of the VM-APF flatten for the relatively lower frequency range as well showing APF characteristics for a wider range of frequency. Thus the plot depicts the resistor tuning of the circuit can be obtained through grounded resistor $R_1$.

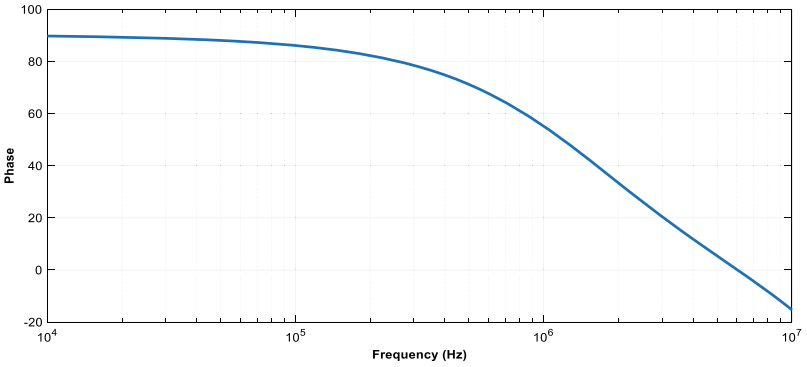



**Fig. 5** Phase vs. Frequency plot of proposed APF configuration

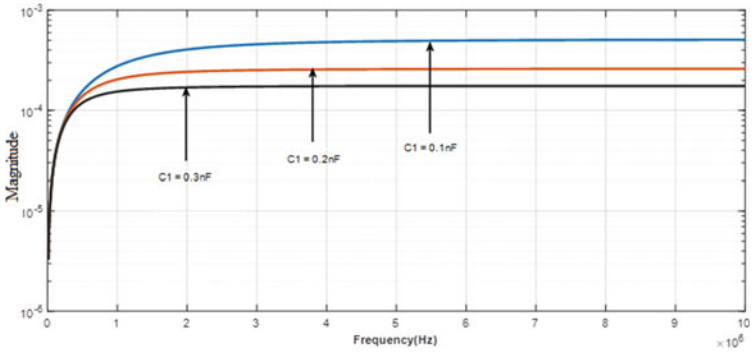



**Fig. 6** Effect of capacitance ($C_1$) variation on the Magnitude versus frequency plot

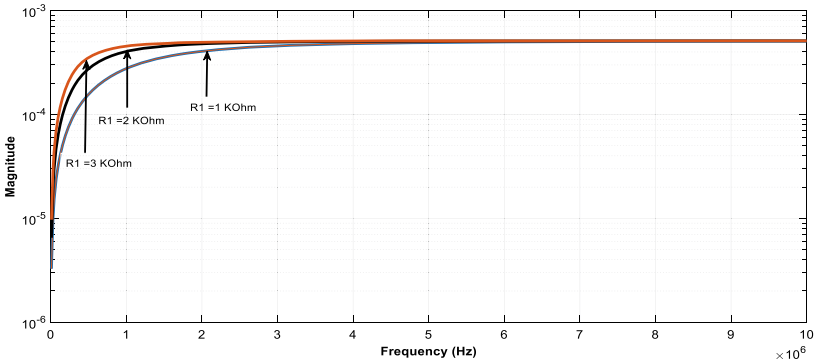



**Fig. 7** Tunable Magnitude response through resistance $R_1$

## 5 Conclusion

The first order active VM-APF configuration is reported, having only single CDDITA accompanying three grounded passive components (two-resistor and one-capacitor). The advantages of this proposed circuit are; all the passive elements employed in the circuit are grounded, low input impedance and high output impedance and also the configuration is electronically/resistor tunable. For the verification task, PSPICE simulations have been used and generated results validate the working of the proposed APF.

## References

1. Bhardwaj K, Srivastava M (2021) Novel CDDITA-based-grounded inductance simulation circuits. In: Advances in VLSI, communication, and signal processing 2021. Springer, Singapore, pp 571–582
2. Khanam PA, Srivastava M (2019) Minimum component grounded inductor simulator employing CDDITA. In: 2019 3rd international conference on recent developments in control, automation & power engineering (RDCAPE), 10 October 2019. IEEE, pp 580–584
3. Panwar K, Prasad D, Srivastava M, Haseeb Z (2018) New current mode lossy integrator employing CDDITA. Circ Syst 9(08):117
4. Bhardwaj K, Srivastava M (2021) New resistorless FDNR simulation configuration employing CDDITAs. In: Advances in VLSI, communication, and signal processing 2021. Springer, Singapore, pp 561–570
5. Panwar K, Srivastava M, Bhardwaj K, Prasad D. Grounded parallel RL impedance simulator using CDDITA
6. Prasad D (2019) Realization of analog signal processing/generation circuits. CSI Trans ICT 7(3):257–259
7. Kaur H. Implementation of analog circuits using CD-DITA (Doctoral dissertation)
8. Bhardwaj K, Srivastava M, Panwar K, Prasad D, Roy A (2019) Grounded series R-L impedance simulator using CDDITA, pp 254–257. https://doi.org/10.1109/ICSC45622.2019.8938336

9. Cheta W, Siripruchyanun M, Trachu K, Suwanjan P, Sotner R, Jaikla W (2018) Single VDCC based voltage-mode first-order all pass filter with electronic controllability. In: 2018 18th international symposium on communications and information technologies (ISCIT), 26 September 2018. IEEE, pp 255–260
10. Chaichana A, Siripongdee S, Jaikla W (2019) Electronically adjustable voltage-mode first-order allpass filter using single commercially available IC. In: IOP conference series: materials science and engineering, 1 June 2019, vol 559, no 1. IOP Publishing, p 012009
11. Kumngern M, Mongkol V, Junnapiya S (2013) Voltage-mode allpass section employing only one DDCCTA and one capacitor. In: 2013 eleventh international conference on ICT and knowledge engineering, 20 November 2013. IEEE, pp 1–4
12. Horng JW, Wu CM, Herencsar N (2014) Fully differential first-order all pass filters using a DDCC. Indian J Eng Mater Sci 21:345–350
13. Maheshwari S (2006) Voltage-mode all-pass filters including minimum component count circuits. Active Passive Electron Comp 2007. Article ID: 79159
14. Kaçar F, Ozcelep Y (2011) CDBA based voltage-mode first-order all-pass filter topologies. IU-JEEE 11:1327–1332
15. Herencsar N, Koton J, Jerabek J, Vrba K, Cicekoglu O (2011) Voltage-mode all-pass filters using universal voltage conveyor and MOSFET-based electronic resistors. Radioengineering 20:10–18
16. Cicekoglu O, Kuntman H, Berk S (1999) All-pass filters using a single current conveyor. Int J Electron 86:947–959. https://dx.doi.org/10.1080/002072199132941
17. Pal K, Rana S (2004) Some new first-order all-pass realizations using CCII. Act Passive Electron Compon 27:91–94. https://dx.doi.org/10.1080/0882751031000116188
18. Soliman M (1997) Generation of current conveyor-based all-pass filters from op amp-based circuits. IEEE Trans Circ Syst-II 44:324–330. https://dx.doi.org/10.1109/82.566650
19. Horng JW (2005) Current conveyors based all pass filters and quadrature oscillators employing grounded capacitors and resistors. Comput Electr Eng 31:81–92. https://dx.doi.org/10.1016/j.compeleceng.2004.11.006
20. Ibrahim MA, Kuntman H, Ozcan S, Suvak O, Cicekoglu O (2004) New first-order inverting-type second generation current conveyor-based all-pass sections including canonical forms. Electr Eng 86:299–301. https://dx.doi.org/10.1007/s00202-003-0205-3
21. Tarunkumar H, Ranjan A, Pheiroijam NM (2018) Fourth order band pass and all pass filter using single FTFN. In: 2018 international conference on computer communication and informatics (ICCCI), 4 January 2018. IEEE, pp 1–4
22. Bhat MV, Bhat SS, Kamath DV (2012) G m-C current mode fractional all pass filter of order α (0< α< 1). In: 2019 3rd international conference on electronics, communication and aerospace technology (ICECA), 12 June 2019. IEEE, pp 240–245
23. Bhagat R, Bhaskar DR, Kumar P (2019) Inverse band reject and all pass filter structure employing CMOS CDBAs. Int J Eng Res Technol 08(09)
24. Herencsar N (2020) General view on fractional-order all-pass filters using generalized current conveyors. In: 2020 43rd international conference on telecommunications and signal processing (TSP), 7 July 2020. IEEE, pp 689–693
25. Prasad D, Panwar K, Bhaskar DR, Srivastava M (2015) CDDITA-based voltage-mode first order all pass filter configuration. Circ Syst 6(11):252
26. Keskin AU, Pal K, Hancioglu E (2008) Resistor less first order all-pass filter with electronic tuning. Int J Electron Commun (AEU) 62:304–306. https://dx.doi.org/10.1016/j.aeue.2007.04.001
27. Kumar P, Keskin AU, Pal K (2007) Wide-band resistor less all-pass sections with single element tuning. Int J Electron 94:597–604. https://dx.doi.org/10.1080/00207210701289676
28. Toker A, Gune EO, Ozoguz S (2001) New high-Q band-pass filter configuration using current controlled current conveyor based all-pass filters. Proc ICECS 1:165–168. https://dx.doi.org/10.1109/icecs.2001.957706
29. Minaei S, Cicekoglu O (2006) A resistor less realization of the first-order all-pass filter. Int J Electron 93:177–183. https://dx.doi.org/10.1080/00207210600562173

30. Kilinc S, Cam U (2004) Realization of allpass filters using operational trans resistance amplifier (OTRA). In: Proceeding of the IEEE 12th signal processing and communications applications conference, vol 1, pp 133–136. http://dx.doi.org/10.1109/SIU.2004.1338276
31. Kilinc S, Cam U (2004) Operational trans resistance amplifier based first-order all-pass filter with an application example. In: Proceeding of the MWSCAS 2004, vol 1, pp 65–68. http://dx.doi.org/10.1109/mwscas.2004.1353898
32. Cakir C, Cam U, Cicekoglu O (2005) Novel all pass filter configuration employing single OTRA. IEEE Trans Circ Syst-II 52:122–125. https://dx.doi.org/10.1109/TCSII.2004.842055
33. Tanaphatsiri C, Jaikla W, Siripruchyanun M (2008) An electronically controllable voltage-mode first-order all-pass filter using only single CCCDTA. In: International symposium on communications and information technologies (ISCIT 2008), Vientiane, 21–23 October 2008, pp 305–309. http://dx.doi.org/10.1109/iscit.2008.4700203.
34. Maheshwari S, Khan IA (2004) Novel first-order current-mode all pass sections using CCIII. Active Passive Elec. Comp. 27:111–117
35. Higashimura M (1991) Current-mode all pass filter using FTFN with grounded capacitor. Electron Lett 27:1182–1183
36. Maheshwari S (2004) New voltage and current-mode APS using current controlled conveyor. Int J Electron 91:735–743
37. Frey DR (1993) Log-domain filtering: an approach to current-mode filtering. IEE Proc G Circ Devices Syst 140:406–416
38. Biolek D, Senani R, Biolkova V, Kolka Z (2008) Active elements for analog signal processing; classification, review and new proposals. Radioengineering 17:15–32

# Optimization Control Techniques for the Aircraft Yaw Control Lateral Dynamics


**A. C. Pavithra and N. V. Archana**

**Abstract** Presently, advanced control theory plays a major role in the aircraft system for the control of Yaw, Roll and Pitch angles, which was very much necessary for the stabilization of aircraft system. Following the works of control system community for aircraft applications this paper concentrates on the control of Yaw angle by designing various optimal Linear Quadratic Regulator (LQR) controllers using standard existing tools. The optimal controllers are designed for the aircraft lateral Yaw dynamics and the performance of each optimal LQR feedback controllers are tested and validated using MATLAB/SIMULINK environment and the results are compared with different system conditions to select the best optimal LQR feedback controllers in future aircraft control system.

**Keywords** LQR · Optimal · Yaw control · Multi stage · Genetic algorithm

## 1 Introduction

The modern aircraft system requires advanced automatic control to monitor the aircraft subsystems, particularly for the application of military and civil aviation. The architecture, operations and capacity etc. of aircraft systems are rapidly changing day to day in the present scenario. Advanced aircraft has various control structures, among them primary and mandatory flight controls are pitch, roll and yaw control which are basically exist in deflection of elevators, ailerons and rudder or combinations of them. The present paper concentrates on the control system design and stability

A. C. Pavithra (✉)
NIEIT, Mysore, India
e-mail: pavithraac26@gmail.com

Electronics and Communication Department, ATME College of Engineering, Mysore 570 028, Karnataka, India

N. V. Archana
Electrical and Electronics Department, NIE Institute of Technology, Mysore 570 028, Karnataka, India
e-mail: archana@nie.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_61

analysis of yaw control (rotation around the vertical axis) with rudder control input by implementing the four different optimization techniques for the lateral dynamics of aircraft system and the results are compared with respect to peak overshoots and settling time.

## *1.1 Related Works*

A brief literature survey on the control of aircraft system, which will give strong foundation on the existing control techniques for the control and stability analysis is as follows:

The authors in [1] proposed an optimal control algorithm which reduces the error compared to reference value for the yaw angle control in subway systems considering curved road driving. The novel control allocation method was developed in [2] with two optimization objectives, where the results obtained are negligible differences with respect to aerodynamic efficiency. Generalized Dynamic Inversion (GDI) control technique [3] for the linear state dynamic equations of yaw and roll axes control and its effectiveness is verified through numerical simulations. The authors in [4] scrutinized the aerodynamic change of adding a yaw-wise rotational degree of freedom to a single slotted flap of airplane via computational fluid dynamic analyses and the outcome reveals that a suitable gap has to be matched for ameliorate the lift further. The concurrent approach for the state variables associated in lateral dynamics such as yaw rate and slide slip angle were proposed by the authors in [5] and the simulation results show that the proposed controller yields tire-road friction adaptation with all the considered feedback controllers.

The authors in [6] designed an optimal control model for the reduction in noise and successfully applied for the two aircraft system and the results show the reduction of noise at reception points. In [7], analyzed the control navigation strategy of small Unmanned Aerial Vehicle (UAV) flight control system using time domain time constant specifications ζ & Wn. The simulation results of the proposed method relatively low cost and easy operation which was suitable for static stability. The authors in [8] proposed pole placement technique control based on LQR for the linearized aircraft model. The simulation results of the proposed control techniques show that it will be suitable for small aircraft system.

Addresses the Elevators and Ailerons of two control surfaces in [9] namely Elevators and Ailerons for controlling longitudinal and roll control movement. The control surfaces were modeled and implemented with different intelligent controllers such as optimization techniques Genetic Algorithm & Particle Swarm Optimization. The simulation results performance of the proposed control techniques was evaluated based on time response specification of controllers. In [10] developed the model of an aircraft roll control system which will be useful for control strategy to an actual aircraft system was designed for Matlab/Simulink environment. The state and output equations with time domain methods for the automatic flight control system

was proposed by authors in [11] with flight control system considering reference aircraft CHARLIE under different flight conditions.

## 2 Modelling of Yaw Control System

Presently, aircraft system has two types of dynamical equations; one is lateral and other is longitudinal where both represents the dynamics of aircraft with respect to lateral and longitudinal axis respectively. The state variables yaw, roll and slide slip motions comes under the first category of lateral dynamics [12]. Where, the longitudinal dynamics includes the pitch motions. The current section explains the modeling of Yaw control system. Figure 1, represents the control surfaces of aircraft and the forces, moments and velocity components respectively.

The Lateral equations of the aircraft system in state space form is as follows:

$$\dot{\mathrm{x}} = \mathrm{Ax} + \mathrm{Bu} \quad x = \begin{bmatrix} \Delta\beta \\ \Delta P \\ \Delta r \\ \Delta\theta \end{bmatrix} \quad u = \begin{bmatrix} \Delta\delta_a \\ \Delta\delta_a \end{bmatrix}$$

where, δa, δr: aileron & rudder deflection; β, Θ: sideslip & roll angle; P, r: roll & yaw rate.

For, the current research, rudder deflection δr is considered as control input. The numerical values [12] for the state space matrices is as follows:

$$\mathrm{A} = \begin{bmatrix} -0.254 & 0 & -1 & 0.183 \\ -15.969 & -8.395 & 2.19 & 0 \\ 4.549 & -0.349 & -0.76 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} [\mathrm{B}\delta\mathrm{r}] = \begin{bmatrix} 0 \\ 23.09 \\ -4.613 \\ 0 \end{bmatrix}$$

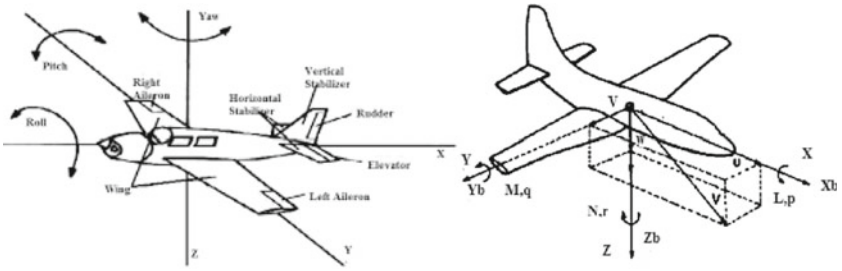



**Fig. 1** Aircraft motions: yaw, roll, pitch and definitions of force, moments and velocity components in a body fixed frame

## 3 Optimization Control Techniques

In this work the optimal Linear Quadratic Regulator (LQR) controller is implemented for the aircraft Yaw control. Since, LQR will guarantee stability and also compromise between output performance and control cost. The proposed LQR is tuned for the weighting matrices Q & R by applying four optimization algorithms such as (i) Multistage (ii) Genetic Algorithm (GA) (iii) Particle Swarm Optimization (PSO) and (iv) Artificial Bee Colony (ABC).

The proposed four optimal LQR tuning algorithms are applied for the aircraft Yaw control system with three flight conditions as shown in Figs. 2 and 3. Figure 2, indicates the flight Yaw control system with initial conditions and the step input as reference input and Fig. 3, shows the flight condition with considering both initial conditions as well as reference input. For the sake of completeness, the proposed LQR tuning algorithms are explained in detail:

### *3.1 Multistage LQR Algorithm (MS-LQR)*

This algorithm was proposed by (R.K. Pandey in 2010) and successfully implemented it for the UPFC based FACTS controllers. The design procedure of this algorithm (currently four stages are considered) is as follows [14]:

1. 1st stage: The weighting matrices Q & R for the initial stage to be considered as Bryson rule

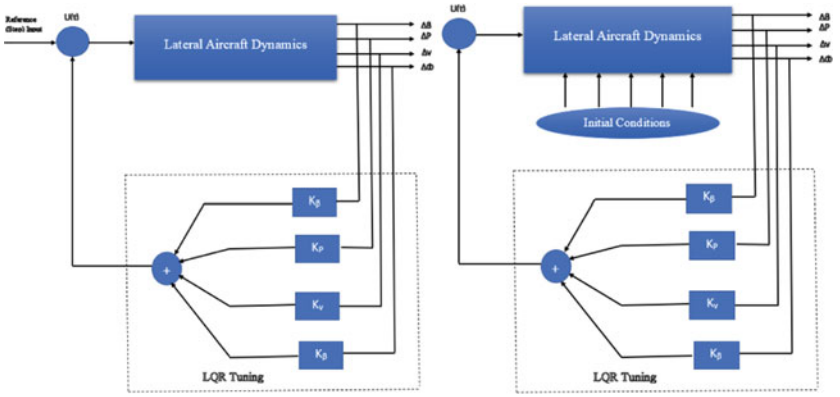



**Fig. 2** Aircraft system with initial conditions and reference inputs

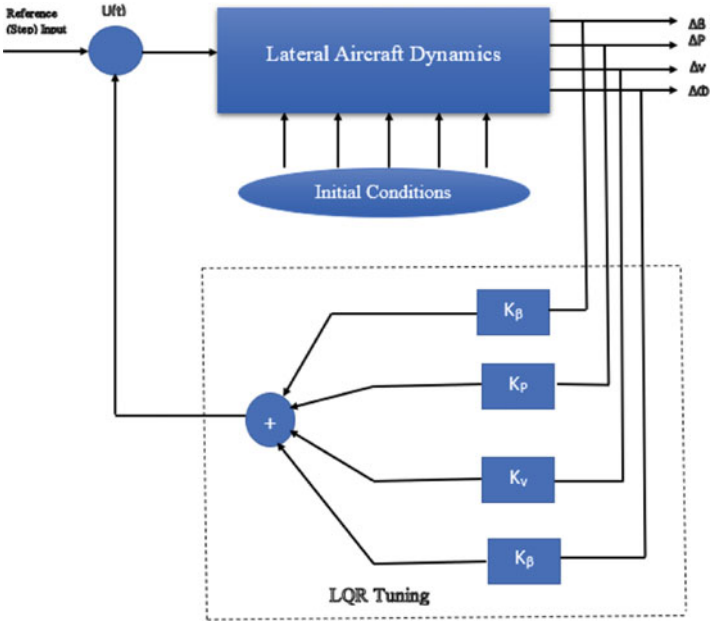



**Fig. 3** Aircraft system with initial conditions and reference input

$$Q = \begin{bmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{bmatrix} \mathrm{R} = [1][\mathrm{K1},\ \mathrm{S},\ \mathrm{E}] = \mathrm{lqr}(\mathrm{A},\ \mathrm{B},\ \mathrm{Q},\ \mathrm{R})$$

$\mathrm{K1} = [0.2096\ 0.6711\ -0.6348\ 1.1047]$.

2. $2^{\mathrm{nd}}$ stage: Choose Q1 & R matrices as

$$\mathrm{Q1} = \begin{bmatrix} 10\,0\,0\,0 \\ 0\ \ 1\,0\,0 \\ 0\ \ 0\,1\,0 \\ 0\ \ 0\,0\,1 \end{bmatrix} \mathrm{R} = [1] \quad \mathrm{Select},\ \mathrm{A1} = \mathrm{A} - (\mathrm{B} * \mathrm{K1})$$

$[\mathrm{K2}, \mathrm{S}, \mathrm{E}] = \mathrm{lqr}\ (\mathrm{A1}, \mathrm{B}, \mathrm{Q1}, \mathrm{R})\ \mathrm{K2} = [1.6287\ 0.3023 - 0.6765\ 0.4759]$.

3. $3^{\mathrm{rd}}$ stage: Choose Q2 & R matrices as

$$\mathrm{Q2} = \begin{bmatrix} 100\,0\,0\,0 \\ 0\ \ 1\,0\,0 \\ 0\ \ 0\,1\,0 \\ 0\ \ 0\,0\,1 \end{bmatrix} \mathrm{R} = [1] \quad \mathrm{Select},\ \mathrm{A2} = \mathrm{A1} - (\mathrm{B} * \mathrm{K2})$$

$$[K3, S, E] = \text{lqr}(A2, B, Q1, R)\ K3 = [6.4593\ 0.1077 - 1.4258\ 0.3498].$$

4. 4th stage: Choose Q3 & R matrices as

$$Q3 = \begin{bmatrix} 1000 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} R = [1] \quad \text{Select, } A3 = A2 - (B * K3)$$

$$[K4, S, E] = \text{lqr}(A3, B, Q1, R1)\ K4 = \begin{bmatrix} 22.0275 & -0.1000 & -2.6207 & 0.1862 \end{bmatrix}.$$

## 3.2 Genetic Algorithm (GA-LQR)

Genetic Algorithm (GA) is one of the optimization techniques which will be useful to find optimal or near-optimal solutions for the optimization problems using the principles of Genetics and Natural Selection. Five phases are considered in a Genetic Algorithm.

(i) Initial Population; (ii) Fitness Function; (iii) Selection; (iv) Crossover and (v) Mutation.

The Fitness function for the current paper is defined as follows [15]:

$$F = S.t_t.\ t_{rmax} + S.ts.t_{smax} + S.O.Mo$$

where,

F = Fitness Function; tr = Rise Time; $t_{rmax}$ = Maximum Rise Time; ts = Settling Time; $t_{smax}$ = Maximum Settling Time; O = Overshoot; Mo = Maximum Overshoot.

Table 1, shows the control criterions of the GA optimization algorithm chosen for the current research [15].

The weighting matrices Q and R obtained for the GA-LQR controller are:

$$Q = \begin{bmatrix} 2.766661 & 0 & 0 & 0 \\ 0 & 0.010324 & 0 & 0 \\ 0 & 0 & 2.5977 & 0 \\ 0 & 0 & 0 & 157.767 \end{bmatrix} R = [0.00195375]$$

The optimal feedback gain matrix $K_{GA}$ is

**Table 1** Criterion numerical values of the GA method

| GA criterion | Numerical value/approach |
|---|---|
| Population size | 20 |
| Max no. of generations | 100 |
| Selection method | Normalized geometric |

$$\text{K4} = \begin{bmatrix} 18.9878 & 8.0624 & -3.5734 & 282.2613 \end{bmatrix}$$

### 3.3 Particle Swarm Optimization (PSO-LQR)

This optimization rule was proposed by Eberhart and Kennedy in 1995. The fitness function proposed for the current paper is as follows [15]:

$$F = \omega max(1 - e^{-1(M0+ess)} + \omega mine^{-(ts-tr)})$$

where, $\omega$ = weighting function and ess = steady state error.

The tuning criterion values for the PSO algorithm in MATLAB environment is as in [15]. The weighting matrices Q and R obtained for the PSO-LQR controller are:

$$Q = \begin{bmatrix} 4.31164 & 0 & 0 & 0 \\ 0 & 0.0117533 & 0 & 0 \\ 0 & 0 & 1.62006 & 0 \\ 0 & 0 & 0 & 187.453 \end{bmatrix} \text{R} = [\,0.0011896]$$

The optimal feedback gain matrix $\text{K}_{\text{PSO}}$ is

$$\text{K}_{\text{PSO}} = \begin{bmatrix} 23.6153 & 9.1774 & -2.0721 & 396.3961 \end{bmatrix}$$

### 3.4 Artificial Bee Colony (ABC-LQR)

The Artificial Bee Colony (ABC) algorithm was introduced by Karaboga in 2005 for optimizing numerical problems. The weighting matrices vector Q and R obtained for the ABC-LQR controller are [15]:

$$Q = \begin{bmatrix} 3.64177 & 0 & 0 & 0 \\ 0 & 0.000800173 & 0 & 0 \\ 0 & 0 & 1.63733 & 0 \\ 0 & 0 & 0 & 135.344 \end{bmatrix} \text{R} = [\,0.000149047]$$

## 4 Simulation Results

The aircraft yaw control system is simulated using the proposed optimal LQR controllers under three cases in MATLAB/SIMULINK environment as shown in Table 3 and the results are compared. The deviation in Yaw rate ($\Delta r$) responses for the case (i), cas (ii) & case (iii) with the legends MS-LQR, GA-LQR, PSO-LQR & ABC-LQR are shown in Figs. 4 and 5 followed by comparison of settling time and peak overshoots from Tables 3 and 4 (Table 2).

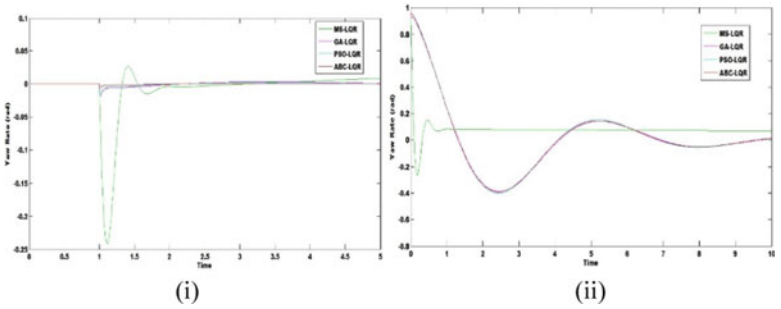



**Fig. 4** Yaw rate deviation for case (i) and case (ii)

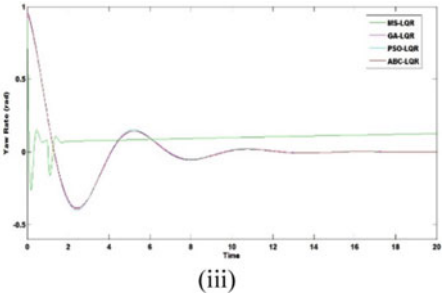



**Fig. 5** Yaw rate deviation for case (iii)

**Table 2** Aircraft yaw control MATLAB simulations

| Case | System operating | LQR controllers applied |
|---|---|---|
| (i) | Initial conditions | Multi-stage, GA, PSO & ABC |
| (ii) | Step input | Multi-stage, GA, PSO & ABC |
| (iii) | Initial conditions and step input | Multi-stage, GA, PSO & ABC |

**Table 3** Comparison of peak overshoots (MP) and settling time (TS) for proposed control methods

| Case | (i) | (ii) | (iii) |
|---|---|---|---|
| MS-LQR | 1s | Above 5s | Above 20s |
| GA-LQR | Above 10s | 3.5s | 15s |
| PSO-LQR | Above 10s | 3.55s | 15s |
| ABC-LQR | Above 10s | 3s | 15s |

| Case | (i) | (ii) | (iii) |
|---|---|---|---|
| MS-LQR | -0.25 | -0.24 | -0.25 |
| GA-LQR | -0.38 | -0.018 | -038 |
| PSO-LQR | -0.4 | -0.017 | -0.4 |
| ABC-LQR | -0.35 | -0.008 | -0.35 |

**Table 4** Comparison of peak overshoots (MP) and settling time (TS) for best proposed control method with other control techniques

| Case | (i) | (ii) |
|---|---|---|
| Proposed MS-LQR | 1s | - |
| Proposed GA-LQR | - | 3.5s |
| GDI by [3] | Above 5s | - |
| Bryson-LQR by [12] | - | 4s |

| Case | (i) | (ii) |
|---|---|---|
| Proposed MS-LQR | -0.25 | -0.25 |
| GDI by [3] | -0.3 | - |
| Bryson-LQR by [12] | - | 1.1 |

## 5 Discussion

The depicted Figs. 4 and 5 and Tables 3 and 4 indicates the comparison of all the four proposed optimal LQR Controllers (MS-LQR, GA-LQR, PSO-LQR & ABC-LQR) for three operating cases of the system. Case (i) results, reveals that the Multistage-LQR provides better performance with respect to peak overshoots & settling time compared to other LQR tuning controllers and Case (ii) represents for the system operating with reference input the ABC-LQR controller has good effectiveness both in peak overshoots as well as settling time in damping compared to other optimal LQR controllers. Finally, Case (iii) the system operating with initial conditions as well as reference input represents again the proposed MS-LQR optimal controller provides better peak overshoots but settling time is quite higher compared to other optimal LQR Controllers. Finally, the overall results reveal that for different operating cases of aircraft, if the multi-stage LQR is tuned with further stages by selecting appropriate weighting matrices Q & R results in improved performance compared to other optimization techniques shown in the current research.

## 6 Conclusion

In the present paper, an aircraft Yaw control system was developed using optimal LQR control technique. The designed controller is validated by simulating using MATLAB/SIMULINK platform under three cases (i) Initial Conditions, (ii) Reference (Step) Input and (iii) Both Initial Conditions and Reference Input to validate and compare the behavior of the proposed control system. In the current paper, the LQR controller was tuned using MS-LQR, GA-LQR, PSO-LQR & ABC-LQR Algorithms. The simulation results conclude that for the case (i) & case (iii) MA-LQR provides better performance with respect to peak overshoots and settling time and for the case (ii) ABC-LQR has more effectiveness in damping compared to other optimization controls.

In the future, further the current research is carried out by implementing Linear Switched Control System to switch between two optimized feedback controllers to utilize the good properties of both sub systems and to obtain a new property which is not present in any of the sub system.

## References

1. Won JH et al (2018) A study on the control method of lateral displacement and yaw angle in curved driving of IRWs system. IEEE Trans Appl Supercond 28(3):1–5
2. Shearwood TR et al (2020) A novel control allocation method for yaw control of tailless aircraft. Aerosp MDPI J 150(7):1–21
3. Bajodah AH et al (2020) Aircraft motion decoupling of roll and yaw dynamics using generalized dynamic inversion control. In: 26th mediterranean conference on control and automation (MED) Zadar, Croatia, 19–22 June 2020
4. Chiba K et al (2020) Aerodynamic efficacy of adding yaw-wise rotational degree of freedom to an airplane flap. Aerosp Syst 3(1):207–217
5. Lenzo B et al (2021) Yaw rate and sideslip angle control through single input single output direct yaw moment control. IEEE Trans Control Syst Technol 29(1):124–139. http://www.springer.com/lncs. Accessed 21 Nov 2016
6. Nahayo F, Khardi S et al (2011) Optimal control of two-commercial aircraft dynamic system during approach. The noise levels minimization. Gen Math Notess 3(2):27–4. ISSN 2219-718
7. Hlaing MS et al (2016) Analysis of autopilot system based on bank angle of small UAV. In: Proceedings of 13th research world international conference, Singapore, 4th March 2016
8. Jiewang Z et al (2008) Modeling, simulation and optimal control for an aircraft of aileron-less folding wing. WSEAS Trans Syst Control 3 (10):869–878. ISSN 1991-8763
9. Singh AK et al (2017) Dynamic modeling and control of aircraft surfaces using hybrid intelligent controllers. IOSR J Electr Electron Eng (IOSR-JEEE) 12(6):21–40. e-ISSN 2278-1676, p-ISSN 2320-3331, Ver. I
10. Akyazi1 O et al. (2012) A self-tuning fuzzy logic controller for aircraft roll control system. Int J Control Sci Eng 2(6):181–188
11. Raj KDS et al (2016) Design and simulation of flight path control systems for CHARLIE aircraft. IOSR J Electron Commun Eng (IOSR-JECE) 11(3):35–44. e-ISSN 2278-2834, p-ISSN 2278-8735, Ver. IV
12. Nair VG et al (2012) Aircraft yaw control system using LQR and fuzzy logic controller. Int J Comput Appl 45(9). ISSN 0975–8887

13. Pandey RK (2010) Analysis and design of multi-stage LQR UPFC. IEEE. 978-1-4244-8542-0/10
14. Shankar S et al (2019) Multi-stage switching control of multi-LQR's for STATCOM operating over wide range of operating conditions in power system. Int J Recent Technol Eng 7(6s):371–379. ISSN 2277-3878
15. Abdulla AI et al (2017) Roll control system design using auto tuning LQR technique. Int J Eng Innov Technol (IJEIT) 7(1), 10–17. ISSN 2277–3754

# Optimization of 2D-Wavelet Filters Based on Taylor Hybrid BAT Algorithm


**T. Rajesh Kumar, K. Kalaiselvi, C. M. Velu, B. Sripathy, C. Karthikeyan, and Soubraylu Sivakumar**

**Abstract** The aim of image compression is to minimize the redundancy of the image data in order to process or transfer the images quickly. It also decreases the file size and allows small space for additional files to be stored. In lossy wavelet based image compression, it is not possible to fully recover the input image, since the quantization error effect is greater. But it can be minimized by using an optimized filter bank. The regular wavelet filter coefficients and its inverse filter were optimized in this proposed novel Taylor hybrid bat Algorithm. The optimized wavelet filters are utilized by SPIHT encoder/decoder for picture compression. Performance metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measures (SSIM) are utilized in image restoration during the decompression process for performance of optimized filters. The obtained results show that by restricting the errors between the input and the decompressed image, the proposed filters beat the traditional wavelet filters.

**Keywords** Adaptive discrete wavelet transform · Discrete wavelet transform · Optimization · Taylor hybrid bat algorithm · Wavelet filter

T. Rajesh Kumar (✉) · C. M. Velu
Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Intitute of Medical and Technical Sciences, Chennai, Tamil Nadu, India
e-mail: t.rajesh61074@gmail.com

K. Kalaiselvi
Department of CSE, SRM Institute of Science and Technology, Chennai, India

B. Sripathy
Department of Mathematics, School of Arts, Humanities and Education, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India

C. Karthikeyan
Department of CSE, Koneru Laksmaiah Education Foundation, Guntur, Andhra Pradesh, India

S. Sivakumar
Computer Science Engineering, Sri Ramachandra Faculty of Engineering and Technology, Porur, Chennai, Tamil Nadu, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_62

## 1 Introduction

Several applications that need large data collection, upload and retrieval, such as record creations, multi-media, video conferencing and image processing in medical field, image compression is very important. For storage and transmission uncompressed images need substantial storage space and bandwidth. Lossless or lossy may be the compression of image. The compressed image can be used for lossless compression algorithm to restore to the accurate duplicate of the original i/p image and nothing of the detail is missing while the process of compression [1]. The real i/p image from the squeezed data cannot be perfectly reproduced by lossy compression. The truth is that much of the detail in the photo has been deleted without altering the images look significantly. Although lossless compression is efficient for precise recovery, adequate compression rates are typically not given to allow image compression advantageous.

Due to its potential to reduce redundant data, the research has been focused on discrete wavelet transform, a conventional technique for lossy image compression in real time implementations and applications [2–4]. A transform utilizing any bulk of filter using wavelet, quantization and based on encoding with normal encoders requires DWT based image compression. As with decompression, based on decoder, de-quantization and converse transform, the compressed data is decoded. For some fair kind of wavelet filter bank, the transformation/inverse transform should be feasible. The method for quantization is nothing more than a procedure for thresholding [5, 6]. Inaccuracies will arise in picture data due to this quantization process and it is impossible to restore the explicit picture after decompression. In this article, a novel and heavily trained image compression approach based on the Taylor hybrid Bat algorithm is proposed for optimized discrete two dimensional wavelet transformation, leads to improve the quality than traditional image compression techniques based wavelet. In order to conclude how correctly the image will be reproduced with respect to the mentioned input image, formal performance parameters such as PSNR and SSIM, have been calculated.

## 2 Related Works

Moore et al. [7] and Peterson [8] concluded that it would reduce this calculation error by optimizing the filter bank with Meta heuristic algorithms Ravi et al. [9]. In order to mitigate the consequences of the quantization inaccuracy, further investigation work will be carried out in this field in some articles. On the other hand, the effects of a vast range of research input and calculations of image worth in each of these works have not been completely tested. Problems with moving salespeople, planning and preparation problems, benefit optimization, etc. are Optimization algorithms [11] which are extremely effective algorithms aimed at seeking solutions to very complex optimization problems. Environment-stimulated heuristic algorithms are a series of

novel methods and techniques derived from natural phenomena to solve problems. Wavelets [10], Genetic Algorithm [13], PSO [14], Ant colony [15], bat algorithm [12], etc. are frequent examples of Environment-stimulated optimization algorithms.

## 3 Proposed Methodology

### *3.1 Transform Domain Flow in Image Compression*

First, an input image is applied with discrete Fourier transformation, then source file is quantized and decoded, then apply with entropy compressing technique. Subsequently, encoding processes is performed and inverse quantization technique is applied and also inverse discrete Fourier transformed will be performed in order to decompress the image. Here also entropy encoding is applied for obtaining qualitative coefficients. Several linear processes of transformation method were developed during last decade of years, which consists of discrete Fourier transform (DFT), discrete cosine transform (DCT) and discrete wavelet transform (DWT) etc., each with its own advantages and hassles. Generally, many to one mapping in quantization leads to errors. The key source of failure in compression is quantization error. Although quantization is a system of loss of pixels, in the process of compression of images, it is highly essential. And the meaningless data of is deleted from the picture only because of this degradation function. An encoder based on entropy that may be quantized using encoding with minimum loss is the next step. It usually uses a technique to measure the probability of each value and depending on these probabilities, which symbolizes the information in coded type.

### *3.2 Image Compression Based on Wavelet Transform*

At advanced compression, coding dependent on wavelets allows significant changes to the images accuracy. A variety of powerful and sophisticated wavelet systems for image compression have been developed and introduced in recent years. The new JPEG 2000 format is having wavelet based compression algorithm, which one is the best advantages available. A wavelet-based coder operates by transforming input, quantizing the transformation coefficient values and next coding of entropy is applied to the quantizer output to remove redundancy. Lack of data is triggered by the quantization process; whereby less important parts of the data are deliberately rejected.

**Errors Obtained During Quantization.** A main stage in image compression is the quantization process. It is a process of approximating a continuous set of values with a set of values with finite levels in the image data. The input of a quantizer is the transformed coefficients in a wavelet-based lossy image compression,

also a bunch of discrete finite-level values are output. This method trims down the required amount of bits for an image to be processed and reverse process certainly not fully possible, i.e. Information missing during the period of quantization will not be perfectly restored. Or else during the quantization process, any unnecessary information would be discarded.

**BAT Algorithm**

a) To spot distance of an item, Bat used his "echolocation ability". It also utilizes this capacity to distinguish even in the darkness between food, predators, and background obstacles.
b) Bats, spontaneously flies at specific characteristics like height, fixed frequency, prey position loudness.
c) Bats also display level of loudness from wide to low loudness. Using differing wavelengths and loudness, bats identify the prey while preserving height, location and speed. They will change the wave, frequency and pulse rate emitted.

**Taylor Series.** Taylor series in standard Taylor equation is shown as per given below,

$$\begin{aligned} H\,(t+1) = &-9.92\,e^{-5}\,H(t-8) + 1.38\,e^{-3}\,H(t-7) - 0.010\,H(t-6) \\ &+ 0.055\,H(t-5) - 0.2259\,H(t-4) + 0.6795\,H(t-3) \\ &- 1.3590\,H(t-2) + 1.3591\,H(t-1) + 0.5\,H(t) \end{aligned} \tag{1}$$

The benefits of Taylor series are in the way in which, the technique is humble and effortless to calculate, under the occurrence of the complex functions. Also it guarantees the exact inference of the filter coefficients and it reaches union easily. The modified Eq. (1) is given in Eq. (2).

$$\begin{aligned} H\,(t) = &-9.92\,e^{-5}\,H(t-9) + 1.38\,e^{-3}\,H(t-8) - 0.010\,H(t-7) \\ &+ 0.055\,H(t-6) - 0.2259\,H(t-5) + 0.6795\,H(t-4) \\ &- 1.3590\,H(t-3) + 1.3591\,H(t-2) + 0.5\,H(t-1) \end{aligned} \tag{2}$$

Reorganize Eq. (2), which provides Eq. (3).

$$H\,(t-1) = 2\begin{bmatrix} H(t) - 1.3591\,H(t-2) + 1.3590\,H(t-3) - 0.6795\,H(t-4) + 0.2259\,H(t-5) \\ -0.055\,H(t-6) + 0.010\,H(t-7) - 1.38\,e^{-3}\,H(t-8) + 9.92\,e^{-5}\,H(t-9) \end{bmatrix} \tag{3}$$

$$H\,(t-1) = \begin{bmatrix} 2H(t) - 2.7182\,H(t-2) + 2.718\,H(t-3) - 1.359\,H(t-4) + 0.4518\,H(t-5) \\ -0.111\,H(t-6) + 0.0208\,H(t-7) - 0.00276\,e^{-3}\,H(t-8) + 0.00019\,H(t-9) \end{bmatrix} \tag{4}$$

The narrative Taylor coefficients are dig out by substituting the Eq. (4) in projected algorithm that is given in Eq. (5).

$$H\,(t,\ f) = \Delta H_T\,(t,\ f) + \begin{bmatrix} 2H(t) - 2.7182\,H(t-2) + 2.718\,H(t-3) - 1.359\,H(t-4) + 0.4518\,H(t-5) \\ -0.111\,H(t-6) + 0.0208\,H(t-7) - 0.00276\,e^{-3}\,H(t-8) + 0.00019\,H(t-9) \end{bmatrix} \tag{5}$$

The algorithm commences with the initialization of the Taylor hybrid bats population. For each Taylor hybrid bat, a starting position with the first solution is given. The pulse rate and loudness are randomly perceived. In all iteration, the Taylor hybrid bat moves from restricted to global solutions. When, after leaping, a bat finds a safer alternative, the loudness and pulse values are modified. The process continues until the final conditions are satisfied. At the end, solution meets the right solution.

### *3.3 Training Process*

Taylor hybrid bat algorithm training method estimates the filter coefficients as follows. The Bior 5.5 coefficients are known to be Taylor hybrid bats initial solution with position. The input picture will be converted using these coefficients. To achieve lossy image compression, the coefficients obtained from discrete wavelet transform will be stated with quantization process and encoded using traditional method. In order to execute the decompression process, using decoding process, steam of bits is decompressed and reverse discrete wavelet transformed process. An error is measured among decompressed and input original image utilizing the PSNR function for exact fitness. In addition, during each Taylor-bat algorithm iteration process, the coefficients will be updated and performance of assessment of exact fitness based on compression, decompression happens. The method proceeds until the full number of iteration processes is reached. The filter coefficients will be chosen and considered as the optimized filter coefficient, which gives the exact fitness of PSNR value. The parameters used in Taylor hybrid bat algorithm for control are size of population $(\psi_i) = 10$, loudness $(Ł_j) = 0.9$, Pulse Rate $(p) = 0.1$, Minimum Frequency $(\gamma_{min}) = 0$ Hz, Maximum Frequency $(\gamma_{max}) = 2$ Hz, Maximum Iterations are 1000 iterations.

## 4 Discussion on Results

### *4.1 Discussion on Training Method*

The investigators used the test image of "camera man" of scale $256 \times 256$ in this work as the image to train the algorithm of the Taylor hybrid bat. The fitness value (PSNR) has also been altered after each iteration during the training phase. The iterate curve representing the number of iterations of PSNR Vs is seen below. There is no shift in PSNR up to the 20th iteration. The PSNR value is retained at 38.40. But, after the 20th iterations, the Taylor hybrid bat algorithm has given slight increases towards an abrupt change and maintain a constant value. During the $50^{th}$ iteration the PSNR value enters into 38.43 and during $100^{th}$ iterations the PSNR value is 38.442 and it maintains the same for $1000^{th}$ iteration. So, we practically examined that Taylor

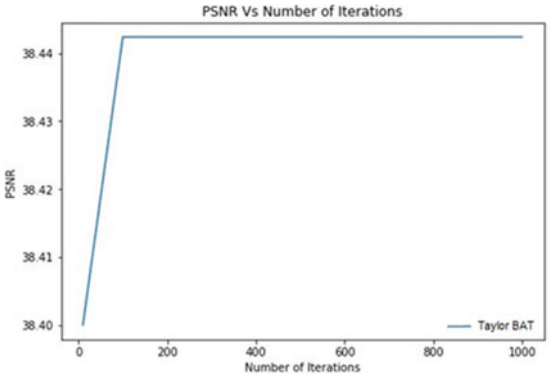



**Fig. 1** PSNR vs number of iterations

hybrid bat algorithm, certainly not providing adequate changes in PSNR after 100$^{th}$ iteration, which is depicted in Fig. 1.

## *4.2 Image Compression and Decompression Results*

The basic images for the experimentally examination are shown in Fig. 2. All displayed, experimental images were compressed and decompressed after compression using the Taylor hybrid BAT optimized wavelet filter. The validation metrics, PSNR, and SSIM were taken among the original input image and the processed

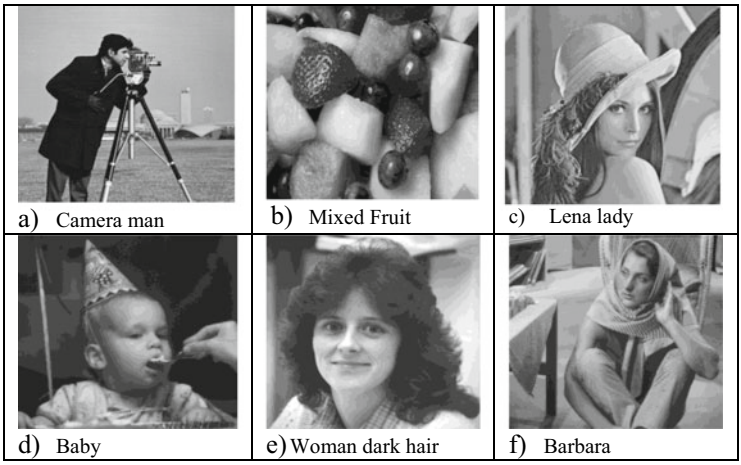



**Fig. 2** 256 × 256 size examination images

decompressed image to check reconstruction ability of the proposed optimized wavelet filter.

The decompression findings of PSNR are depicted in Table 1, when bpp = 0.5, the proposed Taylor hybrid bat optimized results as 35.4983, 34.1035, 32.6802, 39.3870, 39.50566 and 32.5523 respectively for above test images. In same manor, validation based on SSIM while bpp = 0.5, we obtained 0.7372, 0.7683, 0.7231, 0.6524, 0.7655 and 0.7294 in proposed method respectively is depicted in Table 2. The other wavelet filters of db4 [12], rbio 4.4 [13], bior 5.5 [14] shown in Tables 1 and 2, when bpp = 0.5.

Similarly, when bpp = 1, the PSNR provides the following output for the test images in proposed Taylor hybrid bat optimization. They are 36.5606, 36.6479, 39.4422, 45.7051, 41.8007, and 38.1801. When we perform the validation using SSIM, we obtained 0.7271, 0.8842, 0.8574, 0.7143, 0.8223 and 0.8526 in proposed method using test images, which are depicted in the Tables 3 and 4.

For the given 6 test images, Taylor-Bat optimized Wavelet Filter is compared with db4 [12], rbio4.4 [13], and bior5.5 [14] wavelet filter for PSNR, SSIM at the rate of bpp = 0.5 and bpp = 1 are stipulated in the above tabular column. The coefficient of the proposed Taylor-Bat optimized Filter show a better result, when compared with other three wavelet filters of type db4, ribo4.4, bior5.5. The Peak Signal to Noise Ratio (PSNR) coefficient are obtained in float values which are greater than 30, but the structural similarity index measurement (SSIM) values are always less than one.

**Table 1** PSNR @ bpp = 0.5

| Images | Wavelet filter | | | |
|---|---|---|---|---|
| | Taylor-BAT optimized | db4 [12] | rbio 4.4 [13] | bior 5.5 [14] |
| Cameraman | 35.4983 | 32.0224 | 31.1519 | 32.0074 |
| Mixed fruit | 34.1035 | 32.7265 | 32.0117 | 32.3676 |
| Lena lady | 32.6802 | 32.4519 | 30.5803 | 34.2983 |
| Baby | 39.3870 | 38.4817 | 37.2601 | 38.5805 |
| Woman dark hair | 39.5066 | 38.4908 | 37.5557 | 38.4508 |
| Barbara | 32.5523 | 31.2302 | 30.628 | 30.9749 |

**Table 2** SSIM @ bpp = 0.5

| Images | Wavelet filter | | | |
|---|---|---|---|---|
| | Taylor-BAT optimized | db4 [12] | rbio 4.4 [13] | bior 5.5 [14] |
| Cameraman | 0.7372 | 0.4261 | 0.4119 | 0.4011 |
| Mixed fruit | 0.7683 | 0.7123 | 0.7325 | 0.7369 |
| Lena lady | 0.7231 | 0.6865 | 0.6510 | 0.6791 |
| Baby | 0.6524 | 0.628 | 0.5868 | 0.6275 |
| Woman—dark hair | 0.7655 | 0.7235 | 0.7226 | 0.7684 |
| Barbara | 0.7294 | 0.7179 | 0.6875 | 0.7024 |

**Table 3** PSNR@ bpp = 1

| Images | Wavelet filter | | | |
|---|---|---|---|---|
| | Taylor-BAT optimized | db4 [12] | rbio 4.4 [13] | bior 5.5 [14] |
| Cameraman | 36.5606 | 36.1369 | 35.6056 | 35.1889 |
| Mixed fruit | 36.6479 | 36.1292 | 35.4609 | 36.8015 |
| Lena lady | 39.4422 | 38.7125 | 36.4418 | 36.9258 |
| Baby | 45.7051 | 42.6677 | 41.5913 | 43.1339 |
| Woman—dark hair | 41.8007 | 41.2169 | 42.1259 | 42.1039 |
| Barbara | 38.1801 | 37.6133 | 35.0011 | 36.5184 |

**Table 4** SSIM@ bpp = 1

| Images | Wavelet filter | | | |
|---|---|---|---|---|
| | Taylor-BAT optimized | db4 [12] | rbio 4.4 [13] | bior 5.5 [14] |
| Camera man | 0.7271 | 0.7828 | 0.7332 | 0.7198 |
| Mixed fruit | 0.8842 | 0.8723 | 0.8578 | 0.8559 |
| Lena lady | 0.8574 | 0.7431 | 0.7911 | 0.8997 |
| Baby | 0.7143 | 0.7003 | 0.6905 | 0.6977 |
| Woman—dark hair | 0.8223 | 0.8558 | 0.8269 | 0.8324 |
| Barbara | 0.8526 | 0.8939 | 0.8185 | 0.8349 |

As far as human eye view recognition is concern the coefficient value what we have obtained is showing an efficient best output result. It is adequate to show the plot of Camera Man Image having PSNR and SSIM Vs bpp values, which are displayed in Fig. 3 and Fig. 4. In these graphical figure, the bpp values are varying from 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1and its obtained PSNR and SSIM for one image named as camera man image is depicted as follows.

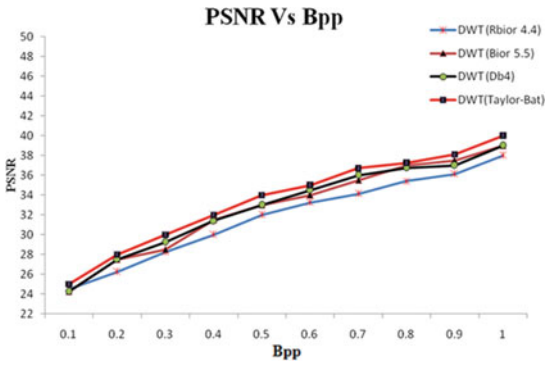



**Fig. 3** PSNR vs Bpp

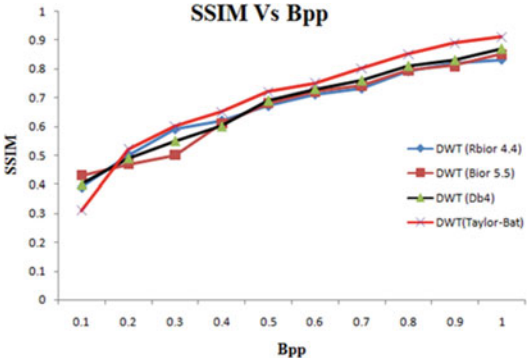



**Fig. 4** SSIM vs Bpp

## 5 Conclusion

The proposed method deals with optimization of Bior5.5 using its coefficients based on Taylor hybrid bat algorithm. It is seen from the experimental findings that the optimized filter performs well in compression image reconstructions using the PSNR and SSIM Vs bpp, relative to conventional wavelet filters. It is also noticed that the Taylor hybrid bat algorithm is failing to increase the PSNR fitness value after a specific number of iterations. In the future, it is proposed that either changing the parameters used for control or hybridizing Taylor-bat algorithm with some other existing algorithm can rectify this limitation.

## References

1. Joshi M, Raval M, Dandawate Y, Joshi K, Metkar S (2014) Image and video compression: fundamentals. Taylor and Francis Group, CRC Press
2. Setyaningsih E, Harjoko A (2017) Survey on hybrid image compression techniques. Int J Electr Comput Eng 7(4):2206–2214
3. Joshi N, Sarode T (2020) Validation and optimization of image compression algorithms. In: Information and communication technology for sustainable development. Springer, pp 521–529
4. Bento PMR, Pombo JAN, Calado MRA, Mariano SJPS (2019) Optimization of neural network with wavelet transform and improved data selection using bat algorithm for short-term load forecasting. Neurocomputing, 52–71
5. Nain G, Gupta A, Gupta R (2019) DWT based compression algorithm on acne face images. In: International conference on intelligent computing and smart communication. Springer, pp 985–993
6. Suzuki T (2019) Wavelet-based spectral spatial transforms for CFA sampled raw camera image compression. IEEE Trans Image Process 29:433–444
7. Moore F, Marshall P, Balster E (2005) Evolved transforms for image reconstruction. In: 2005 IEEE congress on evolutionary computation, vol 3. IEEE, pp 2310–2316
8. Peterson MR (2008) Evolutionary methodology for optimization of image transforms subject to quantization noise. Wright State University

9. Ravi RV, Subramaniam K, Roshini T, Muthusamy SPB, Venkatesan GP (2019) Optimization algorithms, an effective tool for the design of digital filters; a review. J Amb Intell Hum Comput 1–17
10. Bindusri M, Koteswara Rao S (2019) Sunspot data denoising using wavelet. Int J Innov Technol Expl Eng 8(4):230–236
11. Ali M, Ahn CW, Siarry P (2014) Differential evolution algorithm for the selection of optimal scaling factors in image watermarking. Eng Appl Artif Intell 31:15–26
12. Ravi RV, Subramaniam K (2020) Optimized two dimensional wavelet filter from BAT algorithm. In: Proceedings of ICSSS-2020. IEEE Explorer
13. Kaur M, Singh D, Sun K, Rawat U (2020) Color image encryption using non-dominated sorting genetic algorithm with local chaotic search based 5D chaotic map. Futur Gener Comput Syst 107:333–350
14. Di Martino F, Sessa S (2020) PSO image thresholding on images compressed via fuzzy transforms. Inf Sci 506:308–324
15. Chen YH, Chang CC, Lin CC, Hsu CY (2019) Content-based color image retrieval using block truncation coding based on binary ant colony optimization. Symmetry 11(1):21

# Optimization of Linguistic Techniques by Extracting Opinion in Text Summarization Using Transferable Neural Network


**S. B. Rajeshwari and Jagadish S. Kallimani**

**Abstract** In dense opinion mining, calculating opinion targets and expressions based on user defined texts is the core task towards constructing organized opinion review. To extract the aspect and opinion pairs, the phrase structure relations between both the terms plays an important role. The conditional adversarial domain network helps to understand the domain adaptation model's selective information, which is influenced by the anticipating classifiers. At last, integrate the RNN with an identifier on its top for term labeling which will help in textual understanding of the final classified terms. The existing researches used manual entries by adding more number of features to have better opinion extraction but the process consumed more time. Therefore, in order to reduce time complexity for feature learning, deep learning models are proposed for extraction of sentimental across various aspects.

**Keywords** Conditional random field · Cross domain neural network · Deep learning · Feature learning · Opinion mining · Recurrent neural network · Recursive neural network · Sentiment analysis · Single domain neural network · Syntactic structure · Transferable neural network

## 1 Introduction

The dense opinion mining aims to evoke crucial knowledge from the opinion terms. This consists of some sub-functions, including extraction of opinion and aspect terms. The extraction of opinion and aspect terms serves as a basic goal which involves the

Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore, India and affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

S. B. Rajeshwari · J. S. Kallimani (✉)
Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore, India
e-mail: jagadish.k@msrit.edu

S. B. Rajeshwari
e-mail: rajeshwari.sb@msrit.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_63

understanding of accurate features from user input along with opinion terms. For instance, in a cafeteria review: The coffee here has a bright flavor, here the flavor is the aspect term, and bright is the opinion term.

Most of the existing works regarding obtaining the opinion-aspect words were dependent on pre-established order relation mining rules with characteristic learning. However, most of the existing works focused on extraction of features in single domain which lead to adaption failure of other domains where unsupervised data are available. The syntactic association among opinion and aspect words to be obtained is crucial for both overcoming the limitation in different domains and for the individual domain.

To overcome the problem with extraction of cross domain aspects and opinions, RNN is constructed. Since there is huge bridge gap between initial domain and final domain, the available supervised approaches cannot be directly to it and only some of the unsupervised learning techniques are applicable for transferability of labeled domain to unlabeled domain. However, both the techniques work only if there are existing mined patterns of syntactic structures and existing knowledge of aspect and opinion vocabularies. Though the importance of syntactic relation between the opinion and aspect terms is realized, nonetheless, they fall short in terms of adaptability and error detection. RNN based dependency tree is formulated for framing the syntactic flows through data distribution. Considering this prior knowledge, construct a TRNN to propagate features across different domains.

## 2 Literature Review

This division will explain the basic information in the field of analyzing the extraction of aspect and opinion tokens. Firstly, describe the primary definition then the approaches in the extraction and lastly the different categories of extraction analysis that is required to understand the aspects and opinions. Figure 1 shows sample sentences to extract aspect and opinion pairs. Opinion mining examines the point of view of the people, evaluates them across various services, domains, problems, events, fields and their attributes [1]. It is used as a classifier to categorize the opinions into three classes namely positive, negative and neutral. Recently there has been

| **User Review of Hotel System** |
| --- |
| The food of this hotel is delicious but the hotel area is overcrowded. |
| **Aspect and Opinion Terms** |
| Aspect Terms: Food, Area |
| Opinion Terms: Delicious, Overcrowded |
| **Pairing the Aspect-Opinion Terms** |
| Food - Delicious |
| Area - Overcrowded |

**Fig. 1** An example of extraction of aspect and opinion term

huge research studies conducted in this field and various experiments are carried out to bring accuracy in the mining process.

The fundamental approach in which this problem can be considered is by classifying the user input in two ways and addressing this input is either by personal opinion of the user or it can be based on facts. If the user input is based on personal opinion then classifying it as positive, negative or neutral [2]. The identification of this difference between user inputs has to be considered as the basic task. Analyzing this task is complex as few of the user input can be of combination of both cases. Even after having such complexity, this process is still considered to be prominent in initializing the opinion mining process [3]. The opinion of the aspects which are to be classified into positive, negative and neutral can be defined over an interval period of positive and negative classes whereas the neutral can be taken as center of these classes.

Phrase level classification of opinion term involves analyzing each phrase of the record in terms of positive, negative or neutral opinion [4, 5]. The key role of this classification is the isolation of the factual opinions from that of the user's personal opinions [6]. The words in the phrases are semantically relied to each other [7]. More effective methods need to be formulated to solve the complications involved in extracting sarcastic phrases [8]. Aspect level extraction is the core exploration of the attributes of the terms involved in the phrases [9]. Hence to solve the complexity of the opinion mining, advanced machine learning algorithms are required [10, 11].

## 3 Problem Definition

### *3.1 Description*

Tagging is a challenge when it comes to collecting opinion and aspect words. The input is collection of terms denoted as $i = \{t_1, t_2, t_3, \ldots\ldots\ldots t_n\}$ with inserted words denoted as $w = \{w_1, w_2, w_3 \ldots\ldots\ldots w_n\}$, where each $w_i$ belongs to dependency of relation. The goal is to provide labels for each terms resulting in linear output $o = \{o_1, o_2, o_3, \ldots\ldots\ldots o_n\}$ with $o_i$ belonging to certain classes.

For better understanding, consider an instance of statements where each statement has a syntactic structure which is retrieved using dependency tree as shown in the above figure. A dedicated path is established between the root and child of the dependency tree relation named *psubj*.

### *3.2 Motivation*

The syntactic relationship between aspect and opinion words determines how they are extracted. For instance, consider the phrase: *The service in this cafeteria was so poor*

*to us*. The dependent link *psubj* exists amidst the term *service* which is aspect and the term *poor* which is an opinion. If the term *poor* has been extracted, and the dependent relation *psubj* is most often noted as opinion and aspect terms, the aspect term *service* should be calculated. The existing research demonstrated the relational encoding to the certain predefined rules which resulted into inflexibility. This limitation is formulated by applying syntactic structures into feature learning process. This model was named as neural network for single domain which is developed by traversing each phrase of a dependency tree. The left region of Fig. 3 demonstrates the recursive process of single domain and the right region of the Fig. 3 illustrates the cross domain extraction between different word pairs of opinion mining.

Unlike the phrase or record level information transfer, extraction of opinion and aspect words across different domains requires well-dense word-level alterations which are most difficult task. The existing researches stated some predefined rules and sentimental constant fields that were used as core elements. For instance, in Fig. 2, observe that the given user input review: *Their hardware with upgraded version is excellent* is initial domain whereas: *The service in this restaurant was very poor to us* is destination domain. If *poor* was obtained as a general opinion and *opinion–psubj–aspect* would be taken as a general relation as the initial domain, service can be reduced as a term of aspect with help of this relational structure.

Using the same thought some of the existing works formulated some relational structures between aspect and opinion terms as predefined rules embedded in these relational structures that were combined as co-functions into a neural network which is recurrent in its nature [8]. These researches stated the fundamental role of syntactic relation between the terms as a gateway between different fields [12]. However, the pitfall which is observed using this fundamental knowledge is the inflexibility of the rules that were designed manually.

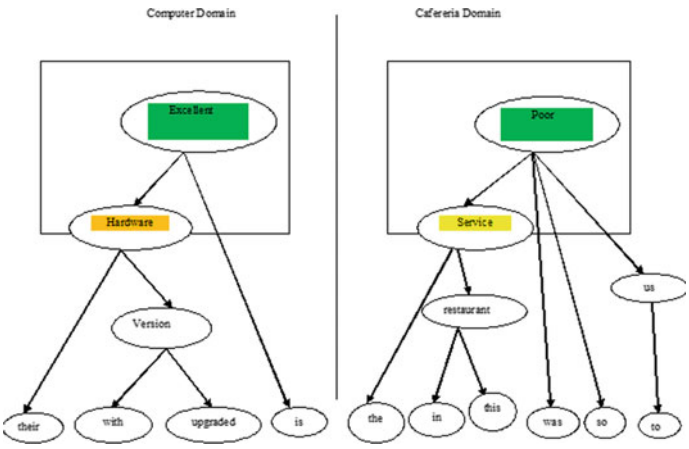



**Fig. 2** An example review of two different domains

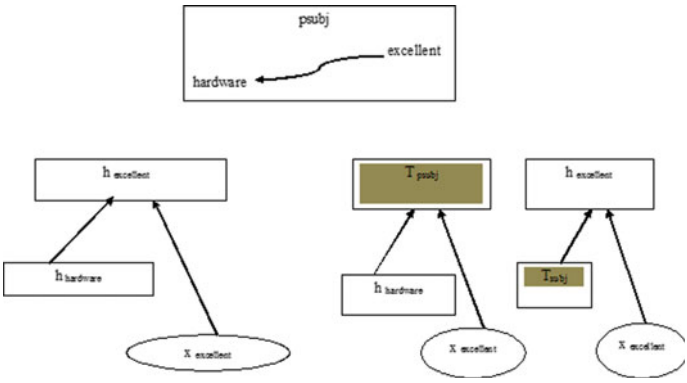



**Fig. 3** An example of two recursive structures

To address this, RNN is proposed which is transferable with the properties which are, rather than relying on manually embedded vital knowledge the model can learn the syntactic relations by itself using dependent recursive neural network using the tree structural patterns. Across different domains, syntactic relational structures are regarded as uniform [13]. When two pairs of words in the initial and final domains have the same dependence relational structures then they are said to be co-related and their functionalities are identical to each other.

To facilitate knowledge transfer, two reusable components that use dependency tree structures are formulated. The first component is a support function that predicts dependent relations between linked aspect and opinion pairs [14]. A domain adversarial network which is conditioned is the second component, which is used to highlight the output into a shared space denotes of syntactically related terms [15].

The TRNN which is deployed on the top of the recursive neural network is shown in the right region of the Fig. 3. Unlike the single domain neural network; the dispersed directions are encoded in the dispersed directional space. The system links each dispersed direction *rpsubj*, which is related as a relational direction mapping. The relational direction acts as an input to the parent word representation from the hidden model. Depending on this kind of deployment, the system applies an identifier on the *rpsubj* to analyze the syntactic relation to which the system belongs to. The identifier acts as a classifier which is trained using supervised learning.

## 4 Structure of the Proposed Model

In this section, SRNN is used for single-domain extraction of opinion and aspect terms and TRNN for cross-domain extraction. These components were obtained using dependency tree patterns using each phrase with their paired terms and computed in a looping manner in which the nodes present in higher levels of the tree structure

are obtained when the child nodes have been retrieved before them. TRNN encodes each dependent route into a dispersed direction that provides a gateway to transfer knowledge across different fields, while SRNN provides the terms with different structural relational patterns as a dependency matrix which invokes the calculation of top-down route of the nodes [16]. To compute final matrix, both these techniques make use of gated recurrent unit. The combined model utilizes the result of either SRNN or TRNN to classify the final opinion aspect paired terms. The following segments outline the various constructions.

## *4.1 Single Domain RNN*

A tree structure for neural networks that is recursive in nature is developed for opinion extraction and aspect term analysis for individual fields based on previous experiments. Unlike the existing researches work, this approach organizes the semantic patterns of phrases and converts them into a spontaneous depiction of the model architecture, allowing structural information to be extracted. Figure 4 gives an outline of the system layout. The module takes term *tn* for each term as input and outputs hidden variant *hn*. A parser produces a relational tree for each phrase, and is utilized to calculate the hidden values in a repeated way. The primary curved arrows in the below figure depicts the relational dependency of the opinion and aspect terms, and their connection flow displayed underneath the arrows. Each connection links a child

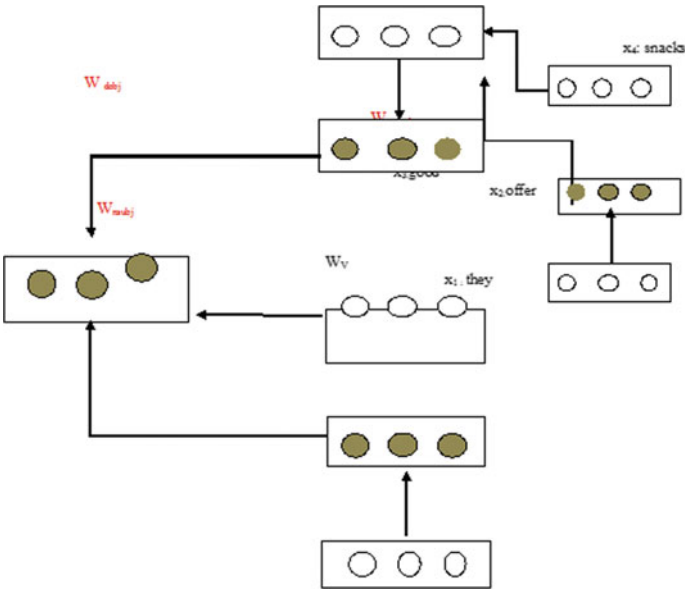



**Fig. 4** The dependency structure for the recursive tree neural network

node to a parent node, where each node has a single parent node. Build a relational tree based patterns depending upon its hierarchy based on which the structural model where each term represents an aspect or opinion in the phrase. The parent node term is derived before its children in the tree structure. The neural network computes the leaf terms to the root iteratively.

For the $n^{\text{th}}$ term, the hidden representation *hn* is calculated and obtain hidden representation for *h4* expressed as Eq. (1):

$$h_4 = f\,(W\,v\,x_4 + b + W_{amod}\,h_3) \tag{1}$$

Note that the final term for snacks is influenced by the feature for good, as well as the dependency matrix relation as per Eq. (1). When this relationship between feature and opinion terms happens regularly in the training module, the system may learn to understand it and use it to assist the results in the testing domain. The system design is however not applicable in a cross domain research. A basic goal is to distribute the arguments of the initial domain to the target results when provided with a labeled initial domain and given with unlabeled destination domain. The system still cannot guarantee the complete learning of the extraction terms. The resulting output vectors are not similar because the opinion and aspect word pairs from two domains have disjoint characteristics and distinct leaf nodes in the dependency tree. This triggers the creation of a novel system to assist information transfer, which is explored in the following segment.

## *4.2 Recursive Neural Network Model*

Provided an input phrase with well learned term encoding of *x*, the system repeatedly determines the hidden terms *h* for each term provides a result of Transferable neural networks that are recursive in their nature. The work of a domain adversarial networks that are conditioned is to understand the representations of the opinion aspect pairs. Transferable Recursive Neural Network produces a link route from a destination path by connecting its edges. The dependency tree is shown in Fig. 4. After that, an additional task uses a route path as feedback to define the relied relationship classes. An auto embedding cipher is used to create connected clusters as an embedding level to remove the unfavorable effects of distorted labeled terms provided by imperfect interpreter.

**Additional Tasks in Recursive Neural Networks.** Transferable Recursive Neural Networks are developed repeatedly from each phrases pre-produced tree. Each relation in the tree is dependent and is provided with a variable directed space that can be utilized as initialization to the additional function to make the domain model to be transferable. The calculation of the hidden representation of each term in the tree is done as a bottom up traversal which is from leaf to higher level nodes. Instead of

utilizing the characteristics of bottom leaves of a tree, utilize a directed route which is dependent syntactically to produce the hidden label values of the higher level nodes.

### *4.3 Predicting Temporal Sequences Using Cross Training*

The neural networks that are recursive predominantly reflects the syntactic associations present in the order, but they disregard temporal similarities. To overcome this constraint, two-part joint model is suggested: To investigate syntactic relationships between aspect and opinion terms, to investigate syntactic relationships amidst the opinion and aspect terms, the TRNN or SRNN is the prime part for single domain or cross domain environments, respectively. In the second part, the GRU is a recurrent neural network variant that is used to model contextual interactions among sentences. In contrast to traditional neural networks, GRU is capable of learning the dependencies by gating modules which is thus less vulnerable to over-fitting than LSTM structures. To build a joint model that integrates the information in particular order, a layered GRU which is a type of recurrent neural network is formed. The final attribute representation for each word is obtained by the equations using the parameters of gradient layer. They are jointly updated in a way that the additional task can be transferred to complete the task with improved performance.

## 5 Experimental Results and Analysis

The final output values for the reviews from three different domains are taken which are hotel, computer, and digital gadgets. The input data is collected from available datasets of these domains which are taken in the Table 1 as shown below.

The outcomes from SemEval 2014 task 4 subtask 1 (Pontiki et al. 2014) and SemEval 2015 task 12 subtask 1 were merged for the restaurant area. (Pontiki friends 2015). In the computer area, extract feedback for SemEval 2014 task 4 subtask 1. Huand Liu's consumer reviews can be found in the digital design domain (2004) which constitutes phrases from five digital systems.

Table 2 shows the composition for each domain, as well as the total number of sentences. To ensure the precision of the comparisons, replicate each experiment three times and aggregate the results. Three random splits are used to separate each

**Table 1** Distribution of data for specified domains

| Data set | Domain | Phrases | Train | Test |
|---|---|---|---|---|
| H | Hotel | 5,841 | 4,381 | 1,460 |
| C | Computer | 3,845 | 2,884 | 961 |
| G | Gadget | 3,836 | 2,877 | 959 |

**Table 2** Comparisons of aspect and opinion terms using baseline components [1]

| Models | Hotel AS | Hotel OP | Computer AS | Computer OS | Gadget AS | Gadget OS |
|---|---|---|---|---|---|---|
| **Proposed model (TRNN)** | **72.11** | **75.26** | **69.76** | **71.35** | **44.48** | **58.34** |
| CRF-embedding | 72.10 | 75.94 | 70.22 | 73.56 | 46.78 | 53.44 |
| SRNN | 68.15 | 68.08 | 64.10 | 62.37 | 38.53 | 50.83 |
| RNCRF | 76.78 | 79.20 | 72.28 | 73.74 | 49.40 | 63.77 |
| ***Wang* et al. *(2021)* SRNN-GRU** | **75.70** | **79.34** | **73.62** | **73.31** | **50.28** | **64.15** |

domain's data into a training set and a testing set, with a 3:1 training to testing ratio. Each domain's data is divided into a training set and a testing set using three random splits, with a 3:1 training to testing ratio. The training data is used to train the model on each split, and the model is then checked on the test range. Table 2 also displays the number of sentences in each preparation and study split.

The labeled data for training from the specified domain is used to train the component for the single domain query. Despite the fact that the sets of testing data is for unlabeled domain adaptation contain both initial and final training outcomes, in each transition experiment. The model is compared in the below table where the values are expressed in terms of baseline components.

The baselines used in the above tables are:

- CRF 1: Standard linear sequenced CRF
- CRF 2 is a syntactic extension of CRF-1 that adds dependency relations
- LSTM: A RNN
- CRF- embedding: As the input functions, use a standard linear-chain CRF.
- SRNN: Single domain RNN.
- RNCRF: The combined module of RNN and a CRF on it.
- SRNN GRU: The combined component for opinion mining of single domain

Table 2 shows the comparison values of proposed model with SRNN-GRU. To make the study compatible to compare with the previously proposed models, common domain contents are considered. Now, the obtained results are compared with similar work presented in [1], which also discusses about the significance of aspect-opinion pairing during text summarization. Comparisons with different baselines, comparison on each component of TRNN-GRU, comparisons with and without auto encoders on noisy dependency relations are presented in [1]. Domain distances, sensitivity and analysis on it and also number of relation groups are presented. On the contrary, in this paper, the baseline components are different. The proposed model yields similar aspect-opinion pair ratio for these different baseline components.

## 6 Conclusion

In this article, the performances of various types of dependent tree structures of neural networks that are recursive in nature were interrogated in single domain and cross domain mining of opinion and aspect pairs. The dependency-tree knowledge combined with a deep recursive framework, automated function learning can be related to syntactic constructs, which have been shown to be important for both single domain and cross domain learning aspects of opinion mining. The syntactic relationships between the aspect and opinion terms within each sentence are encoded by the proposed information distribution models. A dependency relation prediction auxiliary function is introduced to create structural relations for every word in between the source and target domains in order to transfer information through domains. A conditional domain adversarial network is used to gain invariant word features of the domain that are based on the syntactic structure of the terms at the same time. The proposed model will easily deal with noisy relation knowledge due to the use of an auto cipher. Experiments and research studies have been performed to measure and illustrate the effect of each component.

## References

1. Wang W, Pan SJ (2020) Syntactically meaningful and transferable recursive neural networks for aspect and opinion extraction
2. Ahmad Rana T, Cheah YN (2017) A two-fold rule-based model for aspect extraction. Expert Syst Appl 89:273–285
3. Wang J, Peng Y, Lin Y, Wang K (2017) Template Based industrial big data information extraction and query system. In: Second international conference (DMBD), international conference on data mining and big data. Springer, Cham, pp 247–254
4. Liu Q, Gao Z, Liu B, Zhang Y (2015) Automated rule selection for aspect extraction in opinion mining. In: Proceedings of the twenty-fourth international joint conference on artificial intelligence, Buenos Aires, Argentina, 25–31 July 2015, pp 1291–1297
5. Jochim C, Deleris LA (2017) Named entity recognition in the medical domain with constrained CRF models. In: Proceedings of the 15th conference of the European chapter of the association for computational linguistics: Volume 1, Long Papers, Valencia, Spain, 3–7 April 2017, pp 839–849
6. Zhou Y, Jiang W, Song P, Su Y, Guo T, Han J, Hu S (2020) Graph convolutional networks for target-oriented opinion words extraction with adversarial training. In: Proceedings of the 2020 international joint conference on neural networks (IJCNN), Glasgow, UK, 19–24 July 2020, pp 1–7
7. Chutmongkolporn K, Manaskasemsak B, Rungsawang A (2015) Graph-based opinion entity ranking in customer reviews. In: Proceedings of the 2015 15th international symposium on communications and information technologies (ISCIT), Nara, Japan, 7–9 October 2015, pp 161–164
8. Ansari G, Saxena C, Tanvir Ahmad M, Doja N (2020) Aspect term extraction using graph-based semi-supervised learning. Procedia Comput Sci 167:2080–2090
9. Hu J, Zheng X (2020) Opinion extraction of government microblog comments via BiLSTM-CRF model. In: Proceedings of the ACM/IEEE joint conference on digital libraries in 2020, Wuhan, China, 19–23 June 2020, pp 473–475

10. Contractor D, Patra B, Singla P (2020) Constrained BERT BiLSTM CRF for understanding multi-sentence entity-seeking questions. Nat Lang Eng 27:65–87
11. Zhao H, Huang L, Zhang R, Lu Q (2020) SpanMlt: a span-based multi-task learning framework for pair-wise aspect and opinion terms extraction. ACL. In: Proceedings of the 58th annual meeting of the association for computational linguistics, Florence, Italy, 5–10 July 2020, pp 3239–3248
12. Samha AK, Li Y, Zhang J (2015) Aspect-based opinion mining from product reviews using conditional random fields. In: Proceedings of the 13th Australasian data mining conference, Sydney, Australia, 8–9 August 2015, pp 119–128
13. Hu K, Ou Z, Hu M, Feng J (2019) Neural CRF transducers for sequence labeling. In: Proceedings of the ICASSP 2019–2019 IEEE International conference on acoustics, speech and signal processing (ICASSP), Brighton, UK, 12–17 May 2019, pp 2997–3001
14. Yin Y, Wei F, Li D, Xu K, Zhang M, Zhou M (2016) Unsupervised word and dependency path embedding for aspect term extraction
15. Wang W, Pan SJ, Dahlmeier D, Xiao X (2016) Recursive neural conditional random fields for aspect-based sentiment analysis. In: Proceedings of the 2016 conference on empirical methods in natural language processing, Austin, Texas, 1–5 November 2016. Association for Computational Linguistics, pp 616–626
16. Al-Smadi M, Talafha B, Al-Ayyoub M, Jararweh Y (2019) Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. Int J Mach Learn Cybernet 10:2163–2175

# Performance Analysis of a High Gain Quasi Z-Source Network Based Cascaded H-Bridge Multi-level Inverter


**Swathy Nair, K. T. Prajwal, S. Nagaraja Rao, and B. M. Kiran Kumar**

**Abstract** This paper presents a High Gain quasi-Z-Source Network (HG-qZSN) integrated with Cascaded H-Bridge Multi-level Inverter (CHB-MLI) used for renewable energy source applications. The proposed system consists of two quasi Z-source switched capacitor networks and five-level CHB-MLI with eight power switches. The proposed five-level CHB-MLI uses a two HG-qZSN for DC-DC conversion which gives a very high voltage gain or boost factor when compared to conventional boost converters. A five-level CHB-MLI can have the superior performance using sine Pulse Width Modulation (PWM) which is used for DC–AC conversion to provide an efficient conversion and hence suitable for renewable energy applications. In addition, the proposed HG-qZSN based CHB-MLI has modular circuit layout and can be easily extended to higher levels. Simulations are carried out for five-level HG-qZSN based CHB-MLI using MATLAB/Simulink for different duty cycles i.e., 0.2, 0.3 and 0.4. The presented results support the theoretical analysis with a high gain or boost factor of 3, 4.25 and 8 for the 0.2, 0.3 and 0.4 duty cycles respectively.

**Keywords** CHB-MLI · HG-qZSN · PWM · Renewable energy sources

## 1 Introduction

Multi-level Inverters (MLI) are inevitable in most of the high voltage and current applications like non-conventional energy sources, industrial applications, electric vehicles and so on. Also, the MLI are key component to integrate the generated

S. Nair · K. T. Prajwal · S. N. Rao (✉) · B. M. K. Kumar
M.S. Ramaiah University of Applied Sciences, Bangalore 560058, India
e-mail: nagarajarao.ee.et@msruas.ac.in

S. Nair
e-mail: swathynair.ee.et@msruas.ac.in

K. T. Prajwal
e-mail: prajwal.ee.et@msruas.ac.in

B. M. K. Kumar
e-mail: kirankumar.ee.et@msruas.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_64

energy from non-conventional energy sources with grid [1]. The MLI have been classified as Diode Clamped, Flying Capacitor and Cascaded H-bridge MLIs. These inverters are better in harmonics, voltage stress, switching losses and fault tolerance when compared with the conventional two level converters [2–4]. Among the said inverters, Cascaded H-bridge inverters are given priority over other as it can be used for high power application with less number of switches and low switching losses. Therefore, it has modular circuit layout and can be easily extended to higher levels. In this type of inverters, the power is stepped up gradually by series of capacitor cells. The advantage of this type of inverter is that it can give three voltage ranging starting from zero, positive DC voltage and negative DC voltage [5]. But, all these conventional MLI requires an additional DC-DC converter to obtain the desired output voltage [6] as most of the times the input to the MLI is connected from Renewable Energy Sources (RES) like solar or wind energy systems whose output is always lesser than the required voltage. The conventional DC-DC converter is required to get the desired DC link voltage as shown in Fig. 1. This results in limited boost factor, lower efficiency and makes the converter bulky and high cost [7, 8]. Also, these sort of converters suffers from EMI that leads to decrease in the life of capacitors used.

To overcome the drawbacks associated with conventional two stage MLI, an impedance source inverter has been proposed [7]. It is also named as Z-Source Inverter (ZSI) as the impedance is denoted by Z [9]. The ZSI is two port network made up of impedance that is connected between input and inverter bridge as shown in Fig. 2. In ZSI the conventional DC-link is substituted with the impedance source network [10].

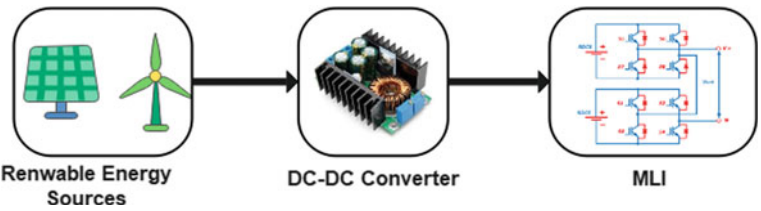



**Fig. 1** RES integrated with conventional DC-DC converter with CHB-MLI

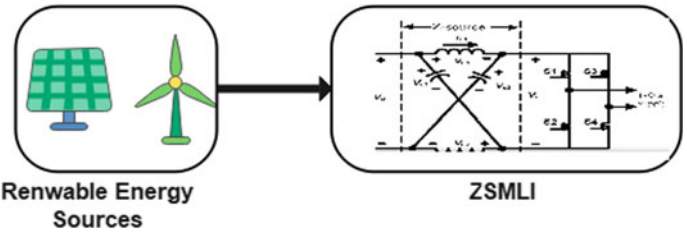



**Fig. 2** RES integrated with Z-source network based CHB-MLI

The main advantage of Z-source network based MLI over conventional boost converter based MLI is high voltage gain or boost factor, high reliability and buck or boost capability [11]. Also, it has its less voltage stress when compared with existing step-up DC-DC converters. Apart from all these advantages, the ZSI poses few disadvantages such as, high inrush current, discontinuous input current, missing of common ground between the input and main bridge [5, 6, 12–14]. To overcome the drawbacks associated with ZSI based MLI, this paper presents a HG-qZSN based CHB-MLI for RES like solar or wind energy systems.

## 2 Related Works

Impedance based converters or ZSI are gaining much importance after its invention as it has many advantages compared to the traditional inverter [17, 18]. The main advantage of Z-Source inverter is its high boost capability with Shoot—Through (ST) protection abilities. Another major advantage for the ZSI is it has the ability to convert the power in single stage, thereby reducing the number of components and thus reducing the cost of the system [19, 20]. The qZSN which is used in this paper has the ability to provide high conversion ratio with the limited number of switches. The components used in the qZSN also has low power ratings compared to the other inverters. The structure of HG-qZSN is shown in Fig. 3.

Quasi Z-source based MLIs are more efficient to overpower current and voltage ripples since the impedance network does the function of a second order filter. It has smaller inductance and capacitance requirement when compared to an inverter interfaced with a conventional boost converter [21]. The circuit also provides buck-boost action. These types of inverters are a better choice for renewable energy integration.

The classical MLIs requires more power switches i.e., '2(m – 1)', therefore, the classical MLIs increases the hardware complexity, bulky and additional space is required for installation [5]. Comparatively diode clamped and flying capacitor MLIs have certain limitations [1, 2], the diode clamped MLI requires excess diodes

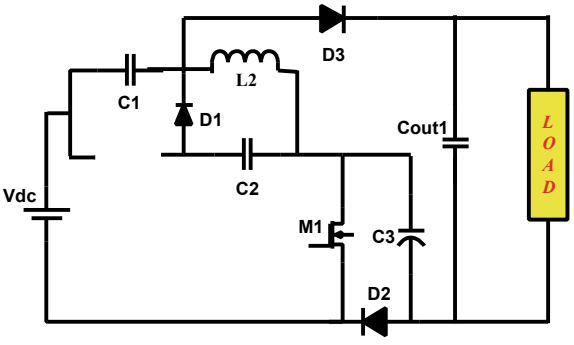



**Fig. 3** Structure of HG-qZSN

whereas, the flying capacitor MLI requires excess capacitors in addition to power switches with increase in levels, in addition, the flying capacitor MLI creates voltage-balancing issues [3]. The main advantage of having HG-qZSN based CHB-MLI over conventional ZSI is more output voltage gain, truncated passive components ratings, continuous input current, low voltage stress on the capacitors and presence of a common ground between source and the main bridge [15, 16]. Another major advantage for the HG-qZSN based CHB-MLII is requires less number of components and thus reducing the cost of the system and space requirement.

## 3 Proposed System

The proposed HG-qZS based CHB-MLI is shown in Fig. 4. Like any general DC-DC converter, the HG-qZSN also has two stages, when MOSFET '$M_1$' is ON and OFF [19, 20]. The topology gives higher voltage gain and reduced capacitor stress in comparison with the conventional topology discussed in previous section. The configuration also provides low Total Harmonic Distortion (THD) and a lesser amount of Electro Magnetic Interference (EMI) which makes it a better configuration for PV integration.

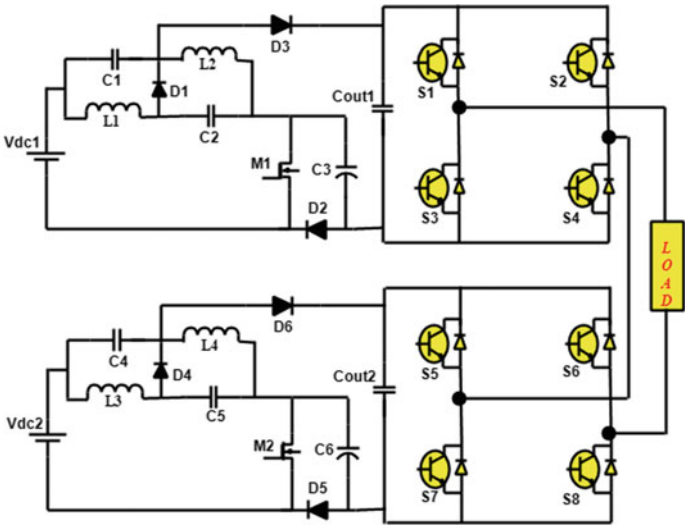



**Fig. 4** Configuration of five-level HG-qZSN based CHB-MLI

### 3.1 Steady State Analysis

The proposed topology of CHB-MLI uses HG-qZS Network for DC-to-DC conversion. A HG-qZSN is shown in Fig. 3. The steady state analysis of the network gives the boost factor or gain of the network. The network consists of four capacitors, two inductors and three diodes. Assuming continuous conduction, the network has two modes of operation as follows:

When switch, $M_1$ is ON, diodes $D_1$ and $D_2$ are reverse biased due to the charge in capacitor plates and diode, $D_0$ is forward biased as shown in Fig. 5(a). The inductor, $L_1$ charges from $V_{dc}$ and $C_2$, so that $V_{L1} = V_{dc} + V_{C2}$. At the same time, inductor $L_2$ charges from $V_{dc}$ and $C_1$. Hence $V_{L2} = V_{dc} + V_{C1}$ and the output voltage across the load resistor, $V_0 = V_{dc} + V_{C1} + V_{C3}$.

When switch $M_1$ is OFF, diodes $D_1$ and $D_2$ are forward biased and $D_0$ is reverse biased as illustrated in Fig. 5(b). Capacitor, $C_1$ gets charged from $L_1$, $C_2$ gets charged from $L_2$, $C_3$ gets charged from the series combination of $V_{dc}$, $L_1$ and $L_2$. The load resistance is supplied by $C_{out1}$. Hence, $V_{L1} + V_{C1} = 0$, $V_{L2} + V_{C2} = 0$ and $V_{dc} = V_{L1} - V_{C2} + V_{C3}$.

At steady state, average voltage across an inductor is zero. Hence, the following equations are arrived as a function of duty ratio, $D = \frac{T_{ON}}{T}$.

At steady state, average voltage across an inductor is zero. Hence the voltage across the capacitors by considering the duty ratio, $D = \frac{T_{ON}}{T}$ are given in Eqs. (1) and (2) respectively:

$$V_{C1} = V_{C2} = \frac{D}{1-2D} V_{dc} \tag{1}$$

$$V_{C3} = V_i + V_{C1} + V_{C2} = \frac{1}{1-2D} V_{dc} \tag{2}$$

Therefore, the output voltage is given by the Eq. (3):

$$V_0 = V_{dc} + V_{C1} + V_{C3} = \frac{2-D}{1-2D} V_{dc} \tag{3}$$

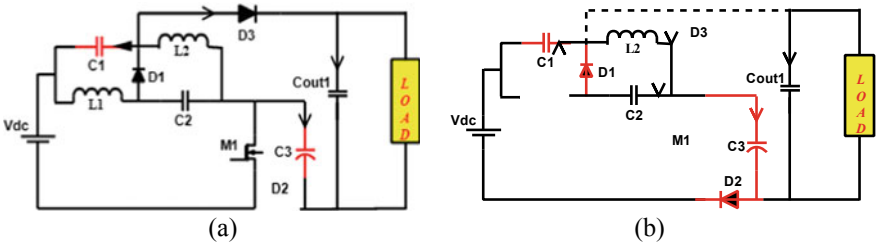



**Fig. 5** Different modes of operation of HG-qZSN

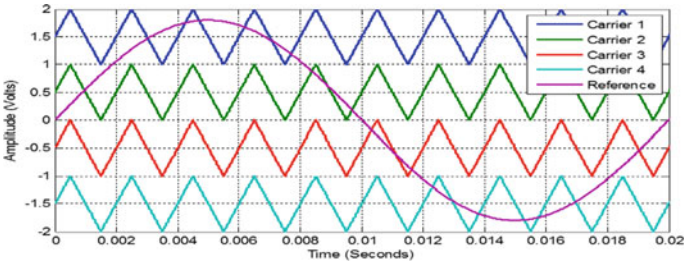



**Fig. 6** Sine PWM generation

**Table 1** Switching sequence for single phase five-level CHB-MLI

| Sl. No | Switches in conduction | Voltage |
|---|---|---|
| 1 | $S_1, S_4, S_5, S_8$ | $+2V_{dc}$ |
| 2 | $S_1, S_4, S_7, S_8$ | $+V_{dc}$ |
| 3 | $S_3, S_4, S_7, S_8$ | 0 |
| 4 | $S_2, S_3, S_7, S_8$ | $-V_{dc}$ |
| 5 | $S_2, S_3, S_6, S_7$ | $-2V_{dc}$ |

Hence, the boost factor of the network, $k$ is given by the Eq. (4):

$$k = \frac{V_0}{V_{dc}} = \frac{2 - D}{1 - 2D} \tag{4}$$

For a duty ratio of 0.3, the boost factor of a HG-qZSN is 4.25, whereas a conventional boost converter has a boost factor of 1.43 [21]. This network provides very high boost factor when compared to a conventional boost converter. The boosted output is given to a five-level CHB- MLI. Sine PWM technique is used as to provide signals for switches in the MLI circuit [22]. The sine PWM technique is typical conventional technique which is used in all MLI for switching pulses. In this research, four carrier waves of triangular shape are compared with the sine reference wave to generate switching pulses as shown in Fig. 6. The switching sequence of the inverter for single phase CHB-MLI is shown in Table 1 to generate five-levels. Full shoot through in a single-phase CHB inverter leads to short circuit of the phase which is not preferable. So, during the preferred non-ST, the switching of single-phase switches along with other phase switches are shown in the Table 1 [22].

## 4 Simulation Results and Discussion

The proposed HG-qZSN based five-level CHB-MLI topology is simulated using MATLAB/Simulink using the simulation parameters are given in Table 2.

**Table 2** Simulation parameters

| Parameters | Value/Range |
|---|---|
| Input voltage of each HG-qZSN | 20 V |
| Switching frequency | 5 kHz |
| $L_1$ and $L_2$, of HGQ-ZSI | 220 μH |
| $C_1$,$C_2$ and $C_3$ of HGQ-ZSI | 330 μF |
| Modulation index | 0.85 |
| D (duty-ratio) | 0 to 0.4 |
| Single-phase 'R' load | 10 Ω |

The output voltages of the HG-qZSN for various duty ratios i.e., 0.2, 0.3 and 0.4 are shown in Fig. 7, Fig. 8 and Fig. 9 respectively along with voltage across capacitors $V_{c1}$, $V_{c2}$ and $V_{C3}$. It is proved that, boost factor of HG-qZSN for the output voltage is 3, 4.25 and 8 times for 0.2, 0.3 and 0.4 the duty cycles respectively. Therefore, the

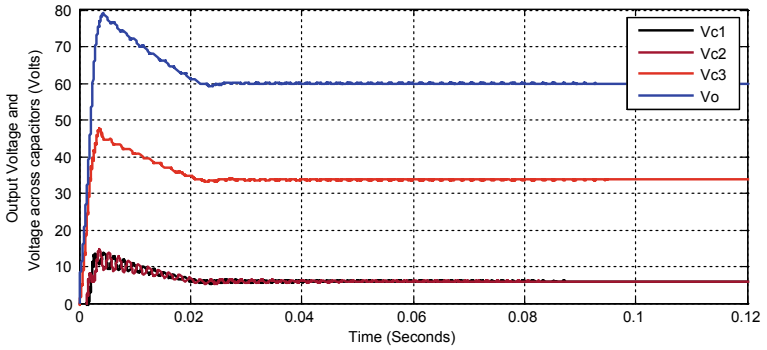



**Fig. 7** Output voltages of HG-qZSN for D = 0.2

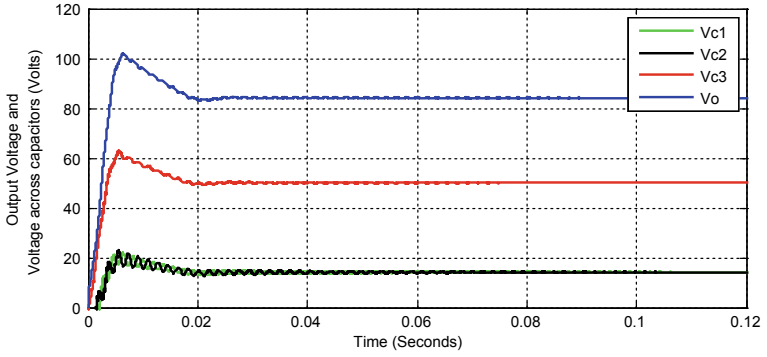



**Fig. 8** Output voltages of HG-qZSN for D = 0.3

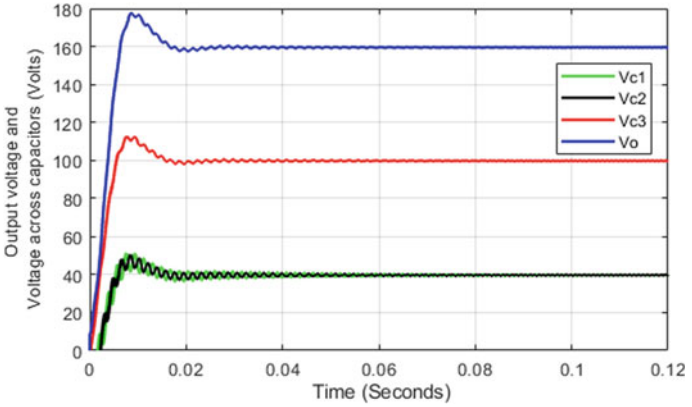



**Fig. 9** Output voltages of HG-qZSN for D = 0.4

**Table 3** Boost factor comparison of HG-qZSN with traditional converters

| Converter | Voltage gain expression | Optimum boost factor |
|---|---|---|
| Conventional boost [21] | $G = \frac{1}{1-D}$ | 2 |
| CLC boost converter [8] | $G = \frac{3}{1-D}$ | 3 |
| Three-level boost converter [16] | $G = \frac{3}{1-D}$ | 3 |
| Proposed HG-qZSN | $G = \frac{2-D}{1-2D}$ | 8 |

output of HG-qZSN is increased to 60 V for 0.2 duty ratio, 85 V for 0.3 duty ratio and 160 V for 0.4 duty ratios.

Analysis of HG-qZSN with the comparison of existing conventional and Z-Source based DC-DC converters are listed in Table 3 to highlight the enhanced gain of proposed HG-qZSN. It can be noticed that the boost factor of HG-qZSN is very high when compared boost converters.

The output of five-level CHB inverter for an input voltage of 20 V is shown in Fig. 10. The peak output voltage is up to 310.8 V in combination with a HG-qZSN for D = 0.4. The FFT analysis is shown in Fig. 11. THD of HG-qZSN based CHB-MLI is obtained as 27.08%.

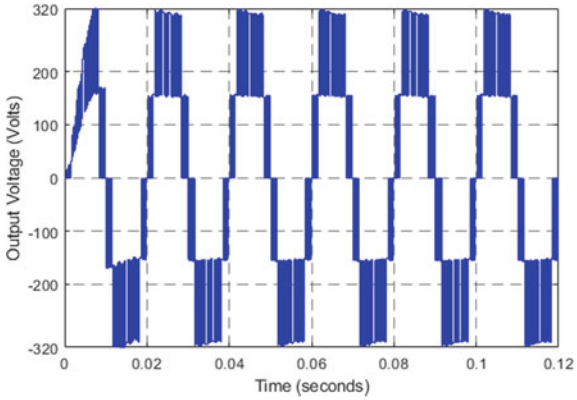



**Fig. 10** Output waveform of proposed HG-qZSN based CHB-MLI

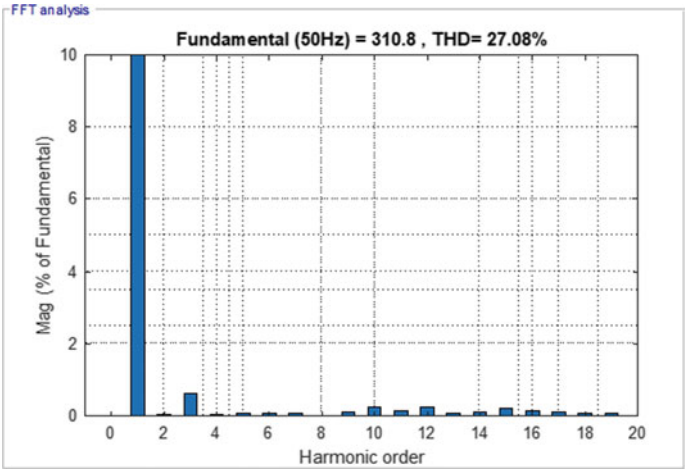



**Fig. 11** FFT analysis of proposed HG-qZSN based CHB-MLI

## 5 Conclusion and Future Scope

A five-level high gain quasi- Z-source based cascaded H-bridge multi-level inverter is presented in the paper for renewable energy applications. The proposed system uses a HG-qZSN for DC-DC conversion which provides a boost factor of 3, 4.25 and 8 for the 0.2, 0.3 and 0.4 duty cycles respectively. Whereas, the conventional boost converter gain is only two due to optimum duty ratio of 0.5. Therefore, the obtained high gain DC output is applied to a five-level CHB-MLI for conversion to AC output. Steady state analysis is done for the network and expression for boost factor is derived. Simulations are done for different duty ratios of qZS network and observed the output voltages. The simulation results support the theoretical calculation. The AC output

of HG-qZSN based CHB-MLI has shown a THD of 27.08%. The proposed system provides high gain and hence it is very efficient for renewable energy applications. Sine PWM is considered in the presented work to analyze the performance of five-level HG-qZS based CHB-MLI. The proposed work can also be extended to space vector pulse width modulation to reduce THD and to improve the voltage gain.

**Acknowledgements** Authors would like to express their sincere thanks to the VC and Management of RUAS, for providing all the facilities needed to carry out this research work.

## References

1. El-Hosainy A, Hamed HA, Azazi HZ, El-Kholy EE (2017) A review of multilevel inverter topologies, control techniques, and applications. In: 2017 nineteenth international middle east power systems conference (MEPCON). IEEE, pp 1265–1275
2. Vijeh M, Rezanejad M, Samadaei E, Bertilsson K (2019) A general review of multilevel inverters based on main submodules: structural point of view. IEEE Trans Power Electron 34(10):9479–9502
3. Omer P, Kumar J, Surjan BS (2020) A review on reduced switch count multilevel inverter topologies. IEEE Access 8:22281–22302
4. Vemuganti HP, Sreenivasarao D, Ganjikunta SK, Suryawanshi HM, Abu-Rub H (2021) A survey on reduced switch count multilevel inverters. IEEE Open J Industr Electron Soc
5. Misal S, Rao SN (2015) Comparative analysis of a 3-phase, 3-level diode clamped ZSI based on modified shoot through PWM techniques. In: 2015 IEEE international conference on signal processing, informatics, communication and energy systems (SPICES). IEEE, pp 1–5 (2015)
6. Zhu X, Zhang B, Qiu D (2018) A new nonisolated quasi-Z-source inverter with high voltage gain. IEEE J Emerg Sel Topics Power Electron 7(3):2012–2028
7. Kumar A, Wang Y, Raghuram M, Naresh P, Pan X, Xiong X (2020) An ultra high gain quasi z-source inverter consisting active switched network. IEEE Trans Circuits Syst II Express Briefs 67(12):3207–3211
8. Kiran Kumar BM, Indira MS, Nagaraja Rao S (2021) Performance evaluation of solar PV using multiple level voltage gain boost converter with C-L-C cell. In: Mekhilef S, Favorskaya M, Pandey RK, Shaw RN (eds) Innovations in electrical and electronic engineering. Lecture notes in electrical engineering, vol 756. Springer, Singapore
9. Honarbari S, AlizadehBidgoli M (2020) Designing a quasi-z-source inverter with energy storage to improve grid power quality. IETE J Res 1–9
10. Nagaraja RS, Rohith HM, Kiran Kumar BM (2020) Reduced switch z-source multilevel inverter using constant boost control methods for renewable energy sources. J Phys Conf Ser 1706(1):012104
11. Ho AV, Chun TW (2017) Single-phase modified quasi-Z-source cascaded hybrid five-level inverter. IEEE Trans Industr Electron 65(6):5125–5134
12. Rajasegharan VV, Premalatha L, Rengaraj R (2018) Modelling and controlling of PV connected quasi Z-source cascaded multilevel inverter system: an HACSNN based control approach. Electr Power Syst Res 162:10–22
13. Rao SN, Veerabhadra, Kumar P (2020) Performance analysis of Z-source inverter topologies for renewable energy sources and fuel cell applications. In: 2020 IEEE international conference on distributed computing, VLSI, Electrical circuits and robotics (DISCOVER). IEEE, pp 165–170
14. Rao SN, Kumar P, Veerabhadra (2021) An enhanced boost Z-source Inverter topology for electrical vehicle applications. In: Advances in energy technology, Lecture notes in electrical engineering, vol 766. Springer, Singapore

15. Zhu X, Zhang B, Qiu D (2019) A high boost active switched quasi-Z-source inverter with low input current ripple. IEEE Trans Industr Inf 15(9):5341–5354
16. Rao SN, Kumar DV, Babu CS (2018) Implementation of multilevel boost DC-link cascade based reversing voltage inverter for low THD operation. J Electr Eng Technol 13(4):1528–1538
17. Rao SN, Kumar BM, Nair S (2021) Three-level modified capacitor assisted extended boost Z-source multilevel inverter for photovoltaic applications. In: Advances in energy technology, Lecture notes in electrical engineering, vol 766. Springer
18. Aleem Z, Yang HK, Ahmed HF, Park JW (2020) Quasi-clamped ZSI with two transformers. IEEE Trans Ind Electron
19. Fang X, Tian Y, Ding X, Ma B (2020) Series-type switched-inductor Z-source inverter. CES Trans Electr Mach Syst 4(1):53–60
20. Zhu X, Zhang B (2017) High step-up quasi-Z-source DC-DC converters with single switched capacitor branch. J Modern Power Syst Clean Energy 5(4):537–547
21. Prabaharan N, Palanisamy K (2016) Analysis and integration of multilevel inverter configuration with boost converters in a photovoltaic system. Energy Convers Manage 128:327–342
22. Reddy VNB, Rao SN, Babu CS (2013) Emphasis of modulated techniques for cascaded multilevel inverters fed drive using FPGA. In: 2013 IEEE international conference on power, energy and control (ICPEC), pp 686–692. IEEE

# Performance Analysis of Classification Models for Liver Disease Diagnosis


**Anusha Marouthu, V. Srikanth, Hari Krishna Deevi, and Siva Krishna Kalluri**

**Abstract** Liver disease prediction system is developed for early diagnosis of human liver disease. For humans, liver is a main organ, which lies inside the abdomen's right upper quadrant. Preliminary diagnosis of liver disease is critical for successful care. It is a difficult job as medical science, which requires a great deal of expertise and experience to diagnose the disease in the early stages using conventional methods. Machine learning algorithm is used to solve this problem. Machine learning algorithms also be used to find hidden diagnostic information and to make effective decisions. This paper also targets to compare different algorithms of machine learning. The system interface is designed using python that will be ready for use in the medical center. After collecting the data, missing data values are filled by estimating median or mean of the attributes and then the feature selection is performed by using principal component analysis. Finally, data classification is accomplished using random forest classifier.

**Keywords** Data pre-processing · Liver disease diagnosis · Machine learning approaches · Principal component analysis · Random forest classifier

## 1 Introduction

Liver disease is one of the main causes of death that exists in both the worldwide and India. Many of the major causes that cause damage to the liver are infections, smoking, drinking etc. In some cases, the condition can also be attributed to inheritance. Diarrhea, loss of appetite, fatigue and many more are basic signs of liver

A. Marouthu (✉) · V. Srikanth
Department of CSE, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Andhra Pradesh, India
e-mail: anushaaa9@kluniversity.in

V. Srikanth
e-mail: vsrikanth@kluniversity.in

H. K. Deevi · S. K. Kalluri
Sri Mittapalli College of Engineering, Guntur, Andhra Pradesh, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_65

failure. Now a day, medical data is generated more and more, so content extraction from the biological datasets is a challenging task in the today's world [10, 16]. Information on the web is also of different types like structured and unstructured kind of records, homogeneous, heterogeneous and mixed varieties of data and current websites present a larger wide variety of difficulties and complexities than traditional ones [14]. A lot of data is produced everyday through health care systems regarding the liver diseases. But most of the data are not used in a proper way. Efficient use of the medical data might return in a better output of determining the cause and prediction of all type of diseases like liver, Heart disease and cancers etc. [11]. Data Mining plays a vital role to classify this data and use the available data in a successful manner [3].

In recent studies, many machine learning models have been established to help the physicians to identify liver diseases in the medical field [2]. The proposed method performs prediction on the available data using classification model [7]. There are several algorithms are used for data classification. In this paper, six algorithms are compared, which are random forest, naïve-Bayes, decision tree, logistic regression, Support Vector Classification (SVC), and K-nearest neighbors [8, 13]. Among these algorithms, random forest achieved better classification performance.

## 2 Related Works

Dhamodharn [1] checked the algorithm for the classification method of liver disease condition using data mining techniques. Author says from his analysis that Naïve Bayes is stronger than FT growth using machine learning by using WEKA Software, Naïve Bayes i.e., 75.54% provides more precision than the FT growth 72.66%. The comparison had taken place across 29 datasets, with 12 specific attributes. Saritha et al. [4] developed an innovative method for the early diagnosis of patient's liver diseases. Isolation of points using plane algorithm was implemented to separate stable patients from poor patients and helped in liver disease diagnosis. The Separation algorithm had been implemented for the classification of liver functional results [12]. The data had been obtained from a Hyderabad hospital. As a consequence, separation algorithm is able to diagnose the liver condition with great precision and the overall time required to complete the training and testing is just a second.

Rajeswari et al. [5], suggested a BUPA dataset for early detection of the condition of liver disorder. Classification algorithms such as FT tree, Naïve Bayes and K-star are used for assessment using tenfold cross validation to forecast the liver disease condition. Then they equate the findings obtained by using these algorithms. Ramana et al. [6], suggested different datasets for liver disease like AP liver dataset and UCLA dataset and then tested the efficiency of classification techniques from precision, consistency, specificity and responsiveness. Throughout his research, Sug [9], usefully compensated for the insufficiency of data on liver disease condition, he said a system focused on oversampling throughout lesser groups. Both CART and C4.5 algorithms produce reasonable results with data set oversampling for liver

disease condition, however in future research, the modest class increase may be decreased to a smaller range [19, 20]. Singh et al. [22] performed experiment on Indian liver patient data that includes the attributes like total bilirubin, globulin ratio, age, sgot, direct bilirubin gender, Alkphos, albumin, and sgpt. In this literature, the experiment was carried out using several classification algorithms with and without feature selection models.

## 3 Proposed Model

### *3.1 Dataset Description*

The dataset used in this research is obtained from the UCI repository. This dataset contains of 7 attributes. They are MCV (Mean corpuscular volume), ALKPHOS, SGPT, SGOT, Gamma-GT, drinks, selector. A clear explanation of attributes is shown in Table 1.

**Table 1** Dataset description

| Sl. no. | Attribute | Medical name | Description |
|---|---|---|---|
| 1 | MCV | Mean corpuscular volume | Refers to the red blood cells count or rate in body |
| 2 | ALKPHOS | Alkaline phosphate | It denotes the amount of alkaline phosphate enzyme present in bloodstream |
| 3 | SGPT | Serum glutamic pyruvic transaminase | It measures an enzyme called alanine transaminase (ALT) in the blood |
| 4 | SGOT | Serum glutamic oxaloacetic transaminase | It determines the functionality of liver by measuring levelsof aspartate aminotransferase |
| 5 | Gamma GT | Gamma glutamic transferase | It measures the amount of enzyme GGT in the blood stream |
| 6 | Drinks | Alcohol | It denotes the amount of alcoholic beverages consumed by patient per day |
| 7 | Selector | – | It denotes the presence or absence of disease in patient liver |


**Fig. 1** Number of missing values in collected data

### *3.2 Data Pre-processing*

The dataset available to process in today's world contain mostly raw data. Preprocessing techniques are used to transform this raw data into efficient and useful format.

**Handling Missing Data.** The data set may sometimes contain null data that refers to missing of data [17]. Handling of such missing and noisy data can be processed and managed. The missing data can be filled manually by calculating mean or median of the attributes or by dropping the missed values. Dropping the missing values might sometime lead to data loss. So it would be better not to choose the dropping process.

As the dataset does not contain any missing values, there is no need of performing any techniques to handle the data. Having no missing values in data is an advantage for the system developed because any change in data is not requires which would result in good accuracy to predict the disease [18, 21]. Figure 1 represents the output generated while finding out the presence of missing values in our data.

**Correlation.** Correlation is a preliminary analysis, which is used to find the association of relationship between two variables or more. In this paper the correlation matrix is used to construct the heat map. The highly correlated attributes are identified and removed using the heat map. In Fig. 2, it is find that the two attributes have highly correlated values. Hence, one of those two are removed and made the dataset more efficient for further algorithm processing.

### *3.3 Feature Selection*

After data pre-processing, feature selection is carried out by using principal component analysis. The $i^{th}$ sample principal component of a vector $x = x_1, x_2, ..x_p$ is denoted in Eq. (1).

$$e_i Z = e_{i1} Z_1 + e_{i2} Z_2 + \ldots e_{ip} Z_p, i = 1, 2, 3, ..p \tag{1}$$

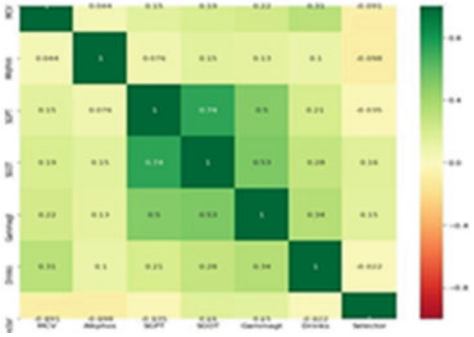


**Fig. 2** Correlation heat map

where, $e_i Z = \left(e_{i1}, e_{i2}, ....e_{ip}\right)$ states $i^{th}$ Eigen value and $Z = Z_1, \ldots . Z_p$ states standardized vector observation. Mathematically, it is represented in Eq. (2).

$$Z_k = \frac{x_{k-\dot{x}_k}}{\sqrt{v_{kk}}}, k = 1, 2, 3, \ldots p \tag{2}$$

where, $x_k$ indicates sample mean and $v_{kk}$ denotes sample variance of $x_k$.

## *3.4 Classification*

The classifiers used in this research are given below;

**Decision Tree.** It is one of the supervised machine learning algorithms that constructs a decision tree through which analysis of data is done and output is predicted.

**SVC.** It works better in low-dimensional data, but a preprocessing phase is needed if the data is high-dimensional. It makes no claim about the number or form of classes in the results.

**Random Forest.** It is an analysis of data through a collection of decision trees. The output will be produced from each tree and the most repeated output will be considered as final output of classification.

**Naïve Bayes.** It is one of the classification techniques which depends on Bayes theorem. The Naive Bayes claims that the presence of a certain function in a class is irrelevant to some other feature.

**KNN.** It is used for both regression and classification problems. It is a basic algorithm that stores all existing cases and through a plurality vote of its neighbors, classifies new cases.

**Logistic Regression.** It is used to measure a discrete range of independent set of variable values. In plain terms, it forecasts the probability of an input by fitting data to a logit function. Out of these classifiers used in this project, Random forest resulted in best accuracy. So, the random forest is implemented for the development and thus presence or absence of disease is predicted.

## 4 Result Analysis

### *4.1 Experimental Setup*

In this study, the proposed model is evaluated on a computer with Intel(R) Core™2 i7-5500U CPU @ 2.40 GHz, 4 MB (Megabyte) Cache memory, 8 gigabyte (GB) RAM (Random Access Memory). In addition, process base frequency is 2.40 GHz, Bus Speed is 5 GT/s DBI2. Windows 10 Home with 64 bit Operating System and × 64-based processor is used to study the performance and comparing the accuracy of classification algorithms. Algorithms were implemented.

### *4.2 Splitting of Dataset*

The dataset need to be split into two parts namely training set and testing set. The division should be done as 80 percentage of the data into training set and the remaining into testing set, which is graphically given in Fig. 3.

From Fig. 3, it is clear that chance for predicting the presence or absence of liver disease is more accurate with our study. Because the accuracy using Random Forest 84% is highest among all compared base papers. Figure 4 and Table 2 represents the graphical comparison of accuracy of all the algorithms implemented in this system. Random Forest stands high among all with accuracy of 84.05%.

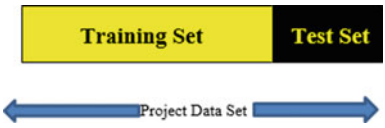



**Fig. 3** Graphical representation of training and testing set

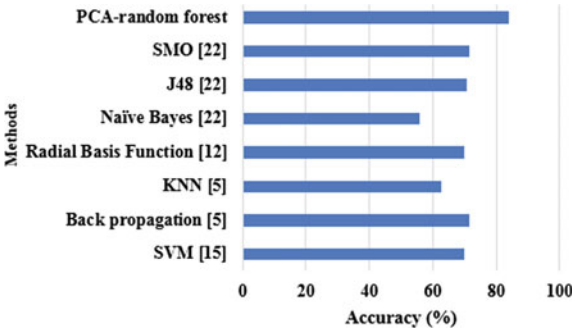



**Fig. 4** Graphical representation of comparison results

**Table 2** Comparison between proposed and existing models

| Method | Accuracy (%) |
|---|---|
| SVM [15] | 70 |
| Back propagation [5] | 71.5 |
| KNN [5] | 62.8 |
| Radial basis function [12] | 70 |
| Naïve Bayes [22] | 55.90 |
| J48 [22] | 70.67 |
| SMO [22] | 71.36 |
| PCA-random forest | 84.05 |

## 5 Conclusion

In this research study, among six algorithms like decision tree, random forests, naive Bayes, KNN, SVC and logistic Regression in-order to predict presence or absence of liver disease using machine learning. The accuracy varies for different algorithms. The accuracy for decision tree algorithm is 71.01%. The accuracy for SVC algorithm is 73.91% and the accuracy for logistic regression is 66.6%. The accuracy for naive Bayes algorithm is 59.42%m, and the accuracy for KNN algorithm is 78.20%. The highest accuracy is given when we have used random forest algorithm, which is nearly 84.05% in liver disease detection.

## References

1. Dhamodharan S (2014) Liver disease prediction using bayesian classification. In: 4th national conference on advanced computing, applications & technologies
2. Jacob J, Mathew JC, Mathew J, Issac E (2018) Diagnosis of liver disease using machine learning techniques. Int Res J Eng Technol 5:4011–4014

3. Sireesha M, Vemuru S, TirumalaRao SN (2018) Coalesce based binary table: an enhanced algorithm for mining frequent patterns. Int J Eng Technol 7(15):51–55
4. Saritha B, Ramana SV, Manaswini N, Priyanka R, Hiranmayi D, Eswaran K (2017) Classification of liver data using a new algorithm‖. In: 4th international conference on new frontiers of engineering, science, management and humanities
5. Rajeswari P, Reena GS (2010) Analysis of liver disorder using data mining algorithms. Glob J Comput Sci Technol 10:48–52
6. Ramana BV, Babu MSP, Venkateswarlu NB (2011) A critical study of selected classification algorithms for liver disease diagnosis. Int J Database Manag Syst 3(2):101–114
7. Sireesha M, Tirumala Rao SN, Vemuru S (2018) Frequent itemset mining algorithms: a survey. J Theor Appl Inf Technol 96(3):744–755
8. Sireesha M, Vemuru S, Tirumala Rao SN (2020) Classification model for prediction of heart disease using correlation coefficient technique. Int J Adv Trends Comput Sci Eng 9(2):2116–2123
9. Sug H (2012) Improving the prediction accuracy of liver disorder disease with oversampling. In: Applied mathematics in electrical and computer
10. Prakash KB, Rangaswamy D (2016) Content extraction of biological datasets using soft computing techniques. J Med Imaging Health Inf 6(4):932–936
11. Sireesha M, Tirumala Rao SN, Vemuru S (2019) Optimized feature extraction and hybrid classification model for heart disease and breast cancer prediction. Int J Recent Technol Eng 7(6):1754–1772
12. Olaniyi EO, Adnan K (2013) Liver disease diagnosis based on neural networks. In: Advances in computational intelligence
13. Shinde SA, Rajeswari PR (2018) Intelligent health risk prediction systems using machine learning: a review. Int J Eng Technol 7(3):1019–1023
14. Prakash KB (2018) Information extraction in current Indian web documents. Int J Eng Technol 7:68–71
15. Hashem EM, Mabrouk MS (2014) A Study of support vector machine algorithm for liver disease diagnosis. Am J Intell 4:9–14
16. Prakash KB, Rangaswamy MAD, Raman AR (2010) Text studies towards multi-lingual content mining for web communication. In: 2nd international conference on trendz in information sciences and computing, pp 28–31
17. Ahammad SH, Rajesh V, Venkatesh KN, Nagaraju P, Rao PR, Inthiyaz S (2019) Liver segmentation using abdominal CT scanning to detect liver disease area. Int J Emerg Trends Eng Res 7(11):664–669
18. Gandey S, Aparna V, Kandarapu R (2020) Preparation and optimization of silybin sustained release microspheres for the treatment of liver disorders. Res J Pharm Technol 13:2742–2748
19. Bommadevara HSA, Sowmya Y et al (2019) Heart disease prediction using machine learning algorithms. Int J Innov Technol Explor Eng 8:270–272
20. Jonnavithula SK, Jha AK, Kavitha M, Srinivasulu S (2020) Role of machine learning algorithms over heart diseases prediction. In: AIP conference proceedings, vol 2292
21. Kavitha M, Gnaneswar G, Dinesh R, Sai YR, Suraj RS (2021) Heart disease prediction using hybrid machine learning model. In: International conference on inventive computation technologies, pp 1329–1333
22. Singh J, Bagga S, Kaur R (2020) Software-based prediction of liver disease with feature selection and classification techniques. Procedia Comput Sci 167:1970–1980

# Performance Estimation of ML Techniques for Pancreatic Tumor Classification in PET/CT Images


**A. Sindhu and V. Radha**

**Abstract** Pancreatic cancer is one of the deadliest diseases and detecting the cancer at an early stage is the difficult task and the reason that 'cancer' symptoms happen only at a problematical period. The reliable screening tool to spot high-risk patients is lacking. Machine Learning (ML) algorithms were utilized in Positron Emission Tomography (PET)/Computed Tomography (CT) pancreatic images. Combination of PET/CT is a current mechanism, which combines both the functional information of 'PET' with the anatomic features of 'CT'. Incorporated 'PET/CT' devices show 'PET' with contrast medium enhanced 'CT' images of the complete body in single condition. Classification techniques have a significant function in 'Computer Aided Diagnostic (CAD)' systems in identifying the malicious (cancerous) area from any kind of medical images. Supervised ML techniques, for example K-Nearest Neighbor (KNN), Decision Tree (DT) and Support Vector Machine (SVM) have been accepted as strong ideal in classifying features for medical images. Classification using SVM achieved 80.7% accuracy; KNN achieved 90.8% accuracy and decision tree achieved 89.9% accuracy in detecting pancreatic cancer. From the above deliberations, this paper attempts to evaluate the latest 'ML' techniques for 'Pancreatic Cancer' classification using second order statistical features from 'PET/CT' images. The results indicate that KNN can be achieved the maximum accuracy than SVM and DT algorithms.

**Keywords** Computed tomography · Grey level run length matrix method · K-nearest neighbor · Machine learning algorithms · Pancreatic cancer

A. Sindhu (✉) · V. Radha
Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India
e-mail: sindhunair1982@gmail.com

V. Radha
e-mail: radhasrimail@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_66

## 1 Introduction

Pancreatic cancer is a common malignant cells form in the muscles of pancreas. Its 'frightful' also 'mortality' has expanded recently. Pancreatic threat is a heterogeneous group of 'neoplasms'. 85% of malignant tumors are of exocrine inception including 'ductal adenocarcinoma' and the uncommon 'acinar' cell carcinoma (1%–2%) [1]. Most customary strategies for automated classification of pancreatic tumors cannot work in a start to finish way. Regularly, the measure of fluorine-18 'FDG: fluorodeoxyglucose' take-up in typical pancreatic parenchyma is inconsequential contrasted in the liver. In any case, both 'malignant & benign' pancreatic states can exhibit extreme 'FDG' take-up. The Medical Image Data is obtained from the Biomedical Devices such as CT MRI, Ultrasound Imaging, PET, PET/CT. Medical image Modeling methods have been developed for decades but many other techniques have been developed for imaging modalities such as X-ray, CT, CAD, MRI, and Ultra sound. PET/CT imaging only increased the research focus for cancer diagnosis for 5 years [2]. CAD system can assist physicians with enhancing analytic efficiency and exactness that does not rely on the specialist's summary judgment and awareness. Existing pancreatic tumors classification techniques experience the ill effects of the matter of fractional mechanization and disregard spatial and fleeting qualities. ML techniques give methods to improve the expectation of clinical execution. Classification algorithm assumes a significant job in CAD frameworks in distinguishing the malicious areas from PET/CT scans. Medical image classification is a big challenge and it is related to the use of methods, identifying the patterns, exploiting the image processing results and validating the classification results [14]. The main objective of the image classification is to not only reach the high accuracy but also identify the part of the body, which is affected by the disease. In the field of medical research, digital imaging data consist of specific characteristics, which are formed by pixels that correspond to a physical object because of imaging modalities [15]. Eventually, three notable Supervised ML systems were chosen for pancreatic tumor classification, specifically KNN, decision tree and SVM. The arrangement of the research paper is as per the following.

## 2 Related Works

Sindhu and Radha [16] performed pancreatic tumor segmentation on the basis of optimized K-means clustering [13] and saliency map model. Obaid and Mohammed [17] proposed a method on classifying the Wisconsin breast cancer from ultrasound scan images. They evaluated three types of classifiers (KNN, DT and SVM) with 32 features and two classes and the performance of these classifiers and the three parameters are evaluated and compared. Based on their research work by applying the machine-learning algorithm, SVM shows better accuracy of 98.1% when compared with other algorithms. Punithavathy and Poobal [18] developed CAD system based

on texture features and fractal analysis and found that it was useful for detecting and classifying in dual modality images. They used 3 texture features and 10 fractal features to detect lung cancer from PET/CT images. SVM classifier is used for lung cancer diagnosis and shows 98.1% of accuracy when compared with KNN and DT. Nandpuru et al. [3] employed on 'MRI' brain cancer classification with SVM (Support Vector Machine). In this research work texture, grey scale and symmetrical features are extracted from MRI brain Images for classification and achieved good result. Padma [4] utilized SVM Classification technique with sentimental Tissues in 'CT' Brain Images. The author used 'Wavelet Based Dominant Grey Level Run Length Texture' for feature extraction. It required highlighted method of medical 'CT' imaging is used and applied techniques, which is reliable for detecting the cancer using SVM and reached 98% accuracy. Ubaidillah et al. [5] utilized Artificial Neural Network (ANN) and SVM classification system for cancer detection. A Comparative analysis was conducted in that research. Using 'SVM' and 'ANN' classifiers, the author evaluated the performance on four different cancer functions. Finally, the ANN classification algorithm achieved best classification results on the extracted features, which have greater input characteristics (prostate & ovarian cancer characteristics). SVM also showed fine performance on features by means of lower input features (breast & liver cancer), and finally SVM classification is approached to improve the result for classifying the cancer Sonawane and Prakash [6] employed the classical KNN. Initially, a 'Dual Tree M-Band Wavelet Transform' is utilized to select the discriminative features. The performance is authorized with 'tenfold' cross justification. The classification exactness has achieved nearly 90%. Minimum numbers of features are misclassified. Grossmann [7] utilized new decision tree ensemble classifier named as Adatree algorithm which uses the characteristics of decision tree. It has new kind of ensemble learning framework, which integrated the advantages of 'ensemble learning algorithm' and 'decision tree'. Also, in that classification system, a replacement splitting criteria has been developed. Couple of feature datasets are utilized to found the performance of the proposed classification algorithm Estruch [8] utilized a formation called 'decision multi-tree' and the experimental assessment has showed the efficiency of 'decision multi-tree'.

## 3 Methodology

Feature extraction may be a general phase utilized in ML system and the set of the features are extracted from the segmented images for subsequent process of learning algorithm [9]. The grey level distribution in a picture is evaluated from the first-order statistic features. Second order statistics features are that they considered pixels in pairs [10].

### 3.1 GLRM

The 'second order statistical' features are computed by utilizing GLRM system. At this point, two approaches are utilized for calculating the grey level pixel-run length. At first phase, a vector-examining pixel runs that made by the function q (L, D, and T). In this function, 'L' means length of the pixel run, 'D' means direction of the pixel run and 'T' means the threshold. D value is calculated almost same in the 'GLCM' method. T value for pixels to be integrated into the pixel-run that is indicated by the user. The process of constructing the pixel-runs is as pursues: every pixel row of segmented image at 'D' is scanned. therefore, the 1st pixel of the row is around to be the initial pixel-run with length 1 and similar grey value 'I' due to the initial pixel; then subsequent pixel within the row is searched; if | I - In | ≤ T ('In' means 'grey' value of consequent pixel), consequent pixel is combined into the 'pixel-run', else, a replacement pixel-run is done then the pointer is navigated to consequent pixel". This procedure is carried out until the searching of the whole row is completed then an alternate row is initiated [11].

The grey level runs are categorized with the grey value of the run in the GLRM process and it reforms the length along with path of the run [12]. Here, 'P (i, j)' means the run length matrix array. It contains the aspects with the grey value features "i" a run length "j". Textural features are identified from the constituents of the array, which does not know the natures of the textures of the input image. Several numeric texture determinations are often calculated since the first run length matrix is p (i, j). The seven authentic features of grey level 'run lengths statistics' feature are described in the Eqs. (1–7).

$$Short\ Run\ Emphasis\ (SRE) = \frac{\sum_i \sum_j p(i,\ j)}{j^2} \tag{1}$$

$$Long\ Run\ Emphasis\ (LRE) = \frac{\sum_i \sum_j j^2 p(i,\ j)}{\sum_i \sum_j p(i,\ j)} \tag{2}$$

$$Gray\ Level\ Non-uniformity(GLN) = \frac{\sum_j \left(\sum_i p(i,\ j)\right)^2}{\sum_i \sum_j p(i,\ j)} \tag{3}$$

$$Run\ Length\ Non-uniformity\ (RLN) = \frac{\sum_i \left(\sum_j p(i,\ j)\right)^2}{\sum_i \sum_j p(i,\ j)} \tag{4}$$

$$Run\ Percentage\ (RP) = \frac{\sum_i \sum_j p(i,\ j)}{n} \tag{5}$$

$$Low\ Gray\ Level\ Run\ Emphasis(LGRE) = \frac{\sum_i \sum_j \frac{p(i,j)}{i^2}}{\sum_i \sum_j p(i,\ j)} \tag{6}$$

$$High\ Gray\ Level\ Run\ Emphasis\ (HGRE) = \frac{\sum_i \sum_j i^2 p(i,j)}{\sum_i \sum_j p(i,j)} \quad (7)$$

where 'Pi'j means the number of runs with intensity 'I' also length 'j'. The GLRM features which obtained are used for the machine learning algorithms for better accuracy in terms of classification and detection.

## 4 Results and Discussion

MATLAB 2018B is used and three notable supervised ML systems have been applied in this study such as KNN, SVM and DT for pancreatic tumor classification. These three classifiers have been applied on the datasets of pancreatic cancer with seven features and two classes (benign, malignant). The below figure shows the proposed method of pancreatic cancer diagnosis. The Pancreatic PET/CT datasets are utilized for assessing the machine learning algorithms, which is acquired on a PET/CT scanner. 119 PET/CT images of 50 to 71-year-old patients confined from 2016 to 2018 were used to evaluate the ML methods. During this research, two equally exclusive labels were considered (i) normal, i.e. images that do not present any radiological abnormalities; (ii) abnormal, i.e. images reported as malicious constituents. 'PET/CT' images are of dimension 512 × 512 × 178. The PET/CT images are in 'Digital Imaging and Communications in Medicine (DICOM)' format that images were transformed into 'JPEG' images of dimension 256 × 256 by appropriate tool and after that it is utilized for this research. Therein dataset, ML algorithms has been tested on MATLAB 2018B runs on Intel i7 CPU with the 2 TB hard disc, 8 GB RAM with Windows 10 Operating Systems. GLRM technique is employed in feature extraction process to extract optimal feature from the segmented images. This process is engaged to enhance the classification process of machine leaning model. These extracted features are often used for training machine-learning algorithms. Some sample imaging features extracted in this paper are listed in Table 1.

From the tables, GLRM features were extracted from the pancreatic PET/CT datasets, which are used as an input files for SVM, KNN and DT machine learning techniques, as stated in Fig. 1.

### *4.1 Performance Evaluation*

The ML algorithms are compared by measuring the subsequent evaluation parameters [15]. Accuracy' value is the proportion of the accurate number of predictions. Can be determined using Eq. (8). Precision' is the ratio of predicted positive examples which really are positive, as shown in Eq. (9). Recall' also called hit rate or sensitivity; it measures how much a classifier can recognize positive examples, as shown in Eq. (10).

**Table 1** Sample features form pancreatic PET/CT images

| IMAGE | SRE | LRE | GLN | RLN | RP | LGRE | HGRE |
|---|---|---|---|---|---|---|---|
| Image1 | 1.1804 | 1.8802 | 18.91 | 8.9203 | 19.701 | 0.0142 | 0.20236 |
| Image2 | 9.0903 | 1.7302 | 24.803 | 9.2903 | 22.3879 | 0.01432 | 0.23063 |
| Image3 | 1.0804 | 1.6902 | 43.9292 | 6.2603 | 54.2042 | 0.01868 | 0.34041 |
| Image4 | 1.2004 | 1.4302 | 45.6037 | 4.7803 | 1.1002 | 0.02376 | 0.45658 |
| Image5 | 1.0004 | 1.4502 | 70.6459 | 2.0903 | 3.2702 | 0.0471 | 0.55732 |
| Image6 | 1.2604 | 1.7502 | 49.0583 | 5.9103 | 1.2202 | 0.02095 | 0.459 |
| Image7 | 1.2704 | 1.5702 | 68.4265 | 4.5303 | 1.3802 | 0.02699 | 0.46452 |
| Image8 | 1.2704 | 1.6802 | 65.7542 | 5.3703 | 81.9681 | 0.02116 | 0.40188 |
| Image9 | 1.2204 | 8.803 | 1.1002 | 8.9303 | 2.8202 | 0.06626 | 0.08727 |
| Image10 | 1.3904 | 8.0903 | 1.0902 | 9.9403 | 2.1102 | 0.06027 | 0.07795 |

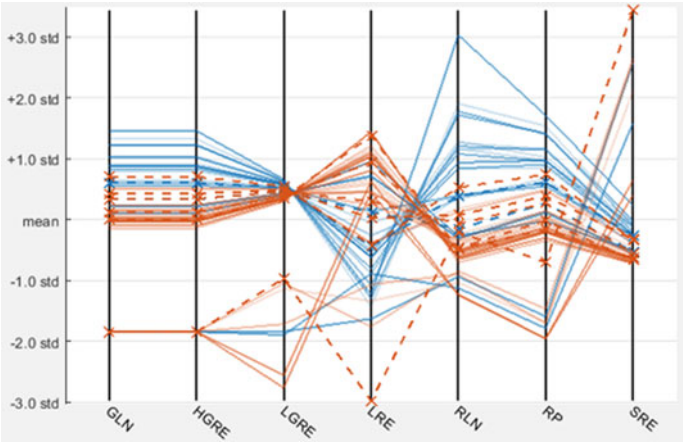



**Fig. 1** Parallel coordinators plot of KNN with GLRM features

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{8}$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{9}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{10}$$

F1_Score is the 'Harmonic Mean' of recall with precision, which is represented in Eq. (11). PPV (positive predictive value)' is the percentage of cases with a positive diagnosis, determined using Eq. (12). 'NPV (negative predictive value)' is the

**Table 2** Performance evaluation

| | DT [16] | SVM [18] | KNN |
|---|---|---|---|
| Precision | 0.9 | 0.92 | 0.96 |
| Recall | 0.7 | 0.85 | 0.9 |
| F1-Score | 0.78 | 0.86 | 0.92 |
| PPV | 0.71 | 0.75 | 0.79 |
| NPV | 0.69 | 0.79 | 0.82 |
| G_Mean | 0.79 | 0.88 | 0.92 |

percentage of cases with a negative diagnosis, determined using Eq. (13). G-mean is the arithmetical mean of recall with precision, as shown in Eq. (14).

$$F1_{Score} = \frac{(2 \times Precision \times Recall)}{(Precision + Recall)} \tag{11}$$

$$PPV = \frac{TP}{(TP + FP)} \tag{12}$$

$$NPV = \frac{TN}{(TN + FN)} \tag{13}$$

$$\text{G-mean} = \sqrt{\text{Precision} \times \text{Recall}} \tag{14}$$

From the above equations, 'TP' and 'TN' stands for True Positive & True Negative values moreover 'FN' and 'FP' stands for False Negative & False Positive. The performance evaluation of DT, SVM, and KNN is represented in Table 2.

Figure 2 shows the KNN algorithm has better value of Precision, Recall and F1-Score in comparison to other ML algorithms. The ML algorithm provided by the MATLAB machine learning toolbox is employed for assessing the trouble of the ML methodologies.

## *4.2 Confusion Matrix*

To describe the performance of the classification algorithms, confusion matrix has been used in this research. It allows visualization of the performance of an algorithm. It also allows easy identification of confusion between classes. For example, one class is commonly mislabeled as the other.

Out of 119 cases, 48 cases were benign and predicted correctly. The remaining 8 cases were benign but could not be predicted accurately as malignant. 48 cases were malignant and predicted correctly. The remaining 15 cases were malignant, but could not be predicted accurately as benign. Overall, 80.7% of the predictions are correct, remaining 19.3% could not be predicted accurately, as shown in Fig. 3.

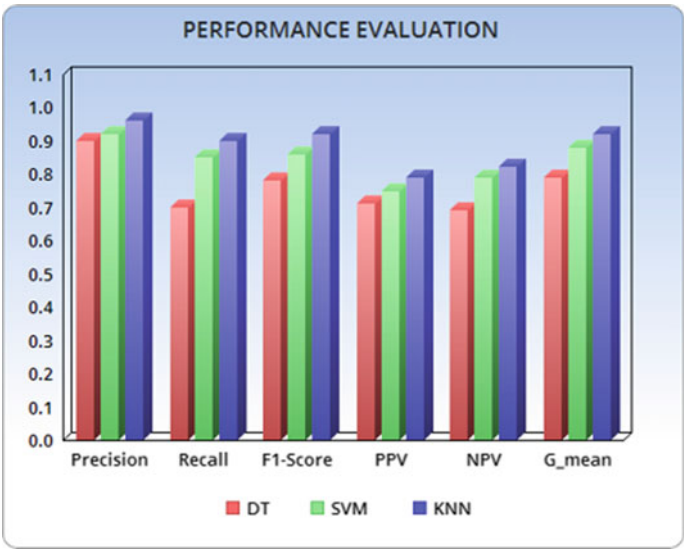



**Fig. 2** Performance evaluation

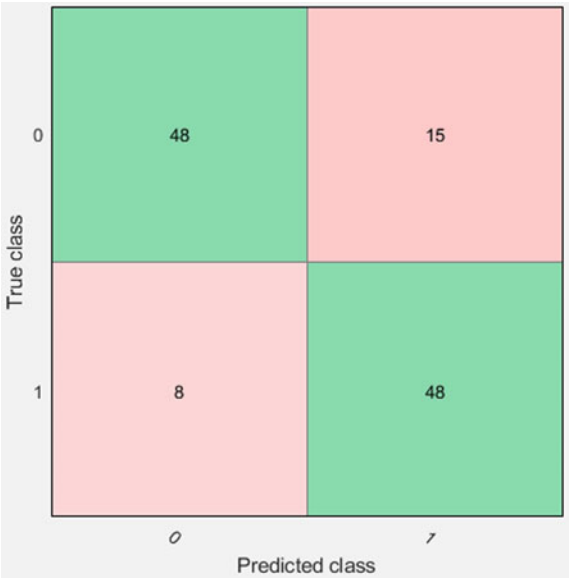



**Fig. 3** Confusion matrix with 80% data for SVM

Out of 119 cases, 54 cases were benign and predicted correctly. The remaining 2 cases were benign but could not be predicted accurately as malignant. 54 cases were malignant and predicted correctly. The remaining 9 cases were malignant, but could

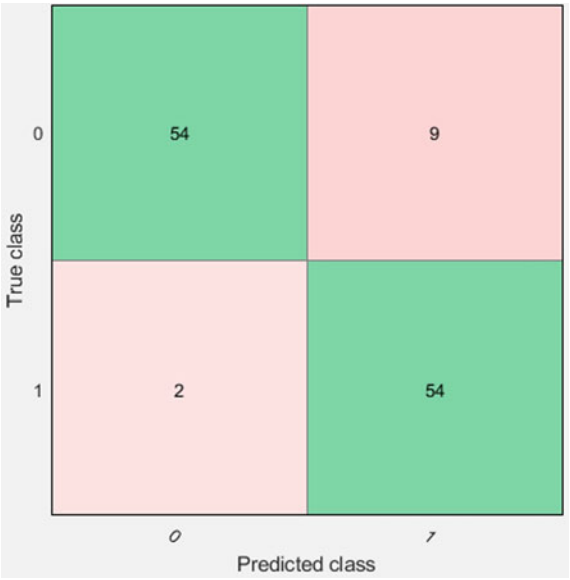



**Fig. 4** Confusion matrix with 80% data for KNN

not be predicted accurately as benign. Overall, 90.8% of the predictions are correct, remaining 9.2% could not be predicted accurately, which is depicted in Fig. 4.

Out of 119 cases, 55 cases were benign and predicted correctly. The remaining 4 cases were benign but could not be predicted accurately as malignant. 52 cases were malignant and predicted correctly. The remaining 8 cases were malignant, but could not be predicted accurately as benign. Overall, 89.9% of the predictions are correct, remaining 10.1% could not be predicted accurately, which is depicted in Fig. 5.

### *4.3 Evaluation of Accuracy*

The accuracy of classification techniques has been evaluated from the pancreatic image features which is extracted by GLRM feature extraction algorithm in the proposed system. From the validation of classification process, KNN predicts less number of false negative cases (misclassification) when compared with SVM [18] and DT [16] from the following Table 3 and Fig. 6, the KNN algorithm has showed highest accuracy rate.

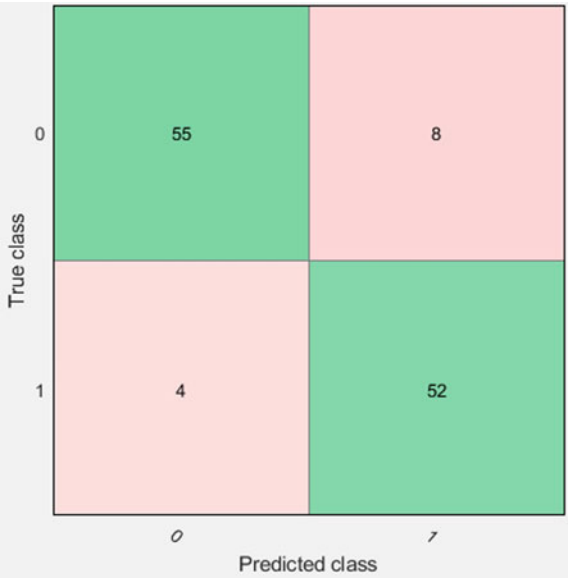



**Fig. 5** Confusion matrix with 80% data for DT

**Table 3** Comparison of ML algorithms using accuracy value

| Sl. No | Algorithms | Accuracy obtained |
|---|---|---|
| 1 | KNN | 90.8% |
| 2 | DT [16] | 89.9% |
| 3 | SVM [18] | 80.7% |

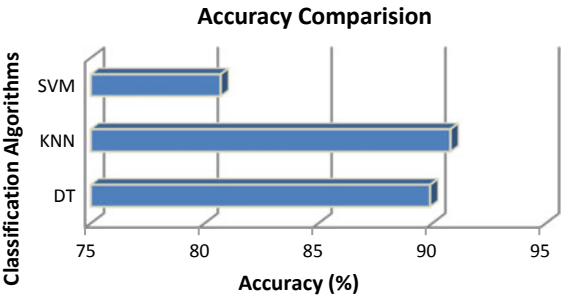



**Fig. 6** Performance evaluation with accuracy

## 5 Conclusion

Detecting and classifying the 'Pancreatic cancer' is a vital role in medical field for saving the human's life. Despite using common diagnosis tools, most of the researchers nowadays have an interest in using ML classification algorithms to classify cancer tumor. This research is conducted to match the performance of three ML

classification techniques, which are DT, SVM and KNN in classifying carcinoma in PET/CT images. All techniques are often effective tools to classify cancer but compare to other classification techniques, KNN shows the higher accuracy because the number of false predictions are low when compared with SVM and DT. In the future scope, novel machine learning algorithm can be implemented to minimize the false predictions (misclassification) and enhance the performance of the classification.

## References


1. American Cancer Society (2017) Cancer Facts & Figures 2017. American Cancer Society, Atlanta, GA
2. Dibble EH, Karantanis D, Mercier G, Peller PJ, Kachnic LA, Subramaniam RM (2012) PET/CT of cancer patients: part 1, pancreatic neoplasms. AJR 199:952–967
3. Nandpuru HB, Salankar SS, Bora VR (2014) MRI brain cancer classification using support vector machine. In: 2014 IEEE students 'conference on electrical, electronics and computer science. 978-1-4799-2526-1/14/$31.00 ©2014 IEEE
4. Padma A, Sukanesh R (2013) SVM based classification of soft tissues in brain CT images using wavelet based dominant gray level run length texture features. Middleeast J Sci Res 13(7):883–888
5. Ubaidillah SHSA, Sallehuddin R, Ali NA (2013) Cancer detection using artificial neural network and support vector machine: a comparative study. Jurnal Teknologi 65(1):73–81
6. Sonawane JM, Gaikwad SD, Prakash G (2017) Microarray data classification using dual tree m-band wavelet features. Int J Adv Sig Img Sci 3(1)
7. Grossmann E (2004) Adatree2: boosting to build decision trees or Improving Adatree with soft splitting rules. NIPS
8. Estruch V (2004) Bagging decision multi-trees. MCS
9. Sonawane JM, Gaikwad SD, Prakash G (2017) Microarray data classification using dual tree M-Band wavelet features. Int J Adv Signal Image Sci 3(1):19–24
10. Srinivasan GN, Shobha G (2008) Statistical texture analysis. In: Proceedings of world academy of science. Engineering and technology, vol 36
11. Bharathi PT, Subashini P (2013) Texture feature extraction of infrared river ice images using second-order spatial statistics. World Acad Sci Eng Technol 74:195–205
12. Yang P, Chen Y (2017) A survey on sentiment analysis by using machine learning methods. In: IEEE 2nd information technology, networking, electronic and automation control conference, pp 117–121
13. Ayyad SM, Saleh AI, Labib LM (2019) Gene expression cancer classification using modified K-Nearest Neighbors technique. BioSystems 176:41–51
14. Shanta R, Shobha G, Sandeep RV, Raj K (2014) Comparative study of decision tree classifier with and without GA based feature selection. Int J Adv Res Comput Commun Eng 3(1)
15. https://medium.com/thalus-ai/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b
16. Sindhu A, Radha V (2019) Pancreatic tumor segmentation based on optimized K-means clustering and saliency map model. lnt J Recent Technol Eng 8:9276–9280
17. Obaid OI, Mohammed MA (2018) Evaluating the performance of machine learning techniques in the classification of Wisconsin breast cancer. Int J Eng Technol 7(4.36):160–166
18. Punithavathy K, Sumathi P (2019) Performance evaluation of machine learning techniques in lung cancer classification from PET/CT images. FME Trans 47:418–423

# Power Quality Analysis of High-Voltage Gain Switched LC Z-Source Inverters


**P. Kannan, P. Bhuvaneswari, K. Prabhu Chandran, P. Ebby Darney, K. Lakshmi Narayanan, and R. Santhana Krishnan**

**Abstract** The regular Switched L and Z-source inverters are tested which accomplishes greater-voltage gain and great in version of power operation at reduced values of shoot-through values when contrasted with the unadventurous ZSI. Since the SL-ZSIs have increased number of inactive components, mass, volume and misfortunes of the framework builds that prompts the drop of efficiency. To conquer the problems of traditional switched inductor Z-source inverter a double single phase switched inductance capacitance later on acquainted to accomplish high-voltage gains at lower estimations of duty ratio using lesser count of inactive components while contrasted to SL-ZSIs. While these single-phase inverters have low estimations of duty ratio, modulation index M gives way to deal with greater qualities that occurs in enhanced current yield at lessened distortion of harmonics. The single-stage inverters are supplanted with three-stage inverters in this anticipated project which accomplishes higher efficiency. The power quality analysis of LC Z-source inverters which is three phase switched whose voltage gain is high are executed and the execution of the proposed inverters are analyzed.

**Keywords** Duty ratio · Efficiency · Harmonic distortion · Switched L-Z-source inverter · Three-stage inverters · Voltage gain

P. Kannan (✉) · K. L. Narayanan
Department of ECE, Francis Xavier Engineering College, Tirunelveli, India
e-mail: kannamuthu1987@gmail.com

P. Bhuvaneswari
Department of ECE, Sri Venkateswara College of Engineering and Technology, Chittoor, Andhra Pradesh, India

K. P. Chandran
Department of ECE, Sreenivasa Institute of Technology and Management Studies, Chittoor, Andhra Pradesh, India

P. E. Darney
Department of EEE, SCAD College of Engineering and Technology, Cheranmahadevi, Tirunelveli, India

R. S. Krishnan
Department of ECE, SCAD College of Engineering and Technology, Tirunelveli, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_67

## 1 Introduction

The superior VSI and CSI are generally essential in numerous mechanical applications such as, UPS systems, dispersed power frameworks, servomotor drives and hybrid electric vehicles [1]. ZSIs are novel electronic converters of power with both capacities to buck and boost voltage that are anticipated for usage in propellant energy transformation systems as well as engine drives with a fore front rectifier [2]. A DC voltage source which is upheld using a moderately expansive capacitor supports the primary inverter circuit [3]. The DC voltage source that must be utilized can be a battery, energy component count, diode rectifier, as well as capacitor. The exemplary ZSIs can create the coveted yield voltage by simply changing the shoot-through period in two power switches in a leg [4]. In any case, an unbounded gain can't be gotten inferable from the dependent impacts of components. Along these lines, higher lifts in the exemplary ZSIs prompt expanding power misfortunes and inconsistency [5].

An effective composition of the Z-source inverter and SL structure, known as a switched inductor ZSI (SL-ZSI) gives a solid step-up inversion that beats the enhancement restrictions of the great ZSI [6]. The prohibited shoot-through period in the regular VSI choose the mandatory to improve the gain of the ZSI. The values of D in the shoot-through and a 3-stage ZSI systems which is associated with modulation index under the imperative of $(M + D \leq 1)$ [7]. Z-source inverter overcomes the theoretical/conceptual barriers and restrictions of the conventional voltage-source current-source converter. Also generates a novel power conversion concept to describe the operating principle and control [8–10].

This paper is systematized as described below. Section 2 gives a brief summary of the existing related works. Section 3 elaborates about the proposed work which is the design of Type 1 and Type 2 three-phase switched LC Z-source inverters. Section 4 comprises about the results of the proposed inverters. The proposed inverters are implemented and analyzed using MATLAB/SIMULINK and the results are validated.

## 2 Related Works

Ding et al. [11] has presented a Cockcroft-Walton Multiplier Voltage (CWMV) qZSI has obvious advantages in terms of the voltage and current stresses on power electronic components. Even though, if k is large, the slope is too steep at less shoot-through duty cycle which brings difficulty in circuit controlling. Hou et al. [12] has demonstrated a fuzzy Proportional Complex Integral (PCI) method to further optimize the control effect of the quasi-Z source grid-connected solar inverter. This method can eliminate the steady-state error and has the characteristic of zero steady-state error adjustment. However, the capacitor voltage of the ZSI is too high, resulting in strong voltage resistance and large capacitance volume that increases

the price. Belila et al. [13] has proposed the control and implementation of a Z-source inverter for hybrid solar-diesel generator energy storage system. This method is used more advantageously without any harmful effects on the Z source inverter operation. In critical situations, when a peak load occurs, at that time the peak DC-link voltage becomes uncontrollable. Faruqui, and Anwer [14] has demonstrated Z-source inverter and voltage source inverter for renewable energy applications to achieve shoot through pulses. But at the same time, it destroying the nature and creating problems in the environment. Palanikumar et al. [15] has presented Modern control scheme for Z-Source inverter based PV power generation systems to improve the power quality features. Then the capacitor voltage is compared with the reference voltage which has to be maintained across the DC link of the inverter bridge. But this method is suitable for specific applications only.

## 3 Proposed Methodology

To propose this research, a harmonic analysis of two three phase switched LC Z-source Inverters are designed and analyzed. With the end goal to acquire increased voltage gain with limited count of inactive components when contrasted with the prevailing SL-ZSIs, that anticipated and is gotten from L-ZSI which includes capacitors, some diodes and 2-switches. The proposed inverters possess lower voltage stress and current stress compared to SL-ZSIs which is similar to i/o voltages. Figures 1 and 2 illustrates the circuit diagram for Type-1, Type-2 SLC ZSI respectively.

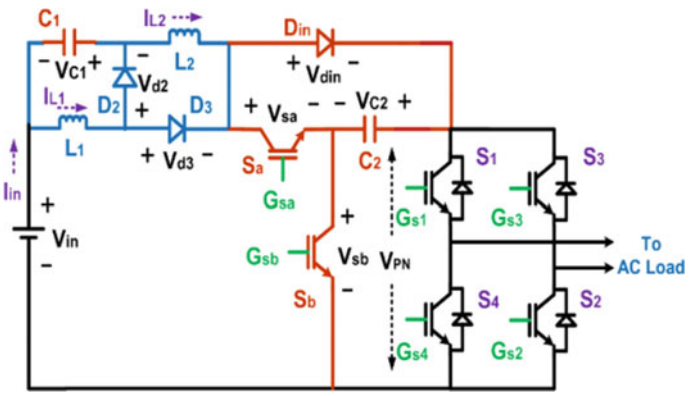



**Fig. 1** Type 1 SLC-ZSI

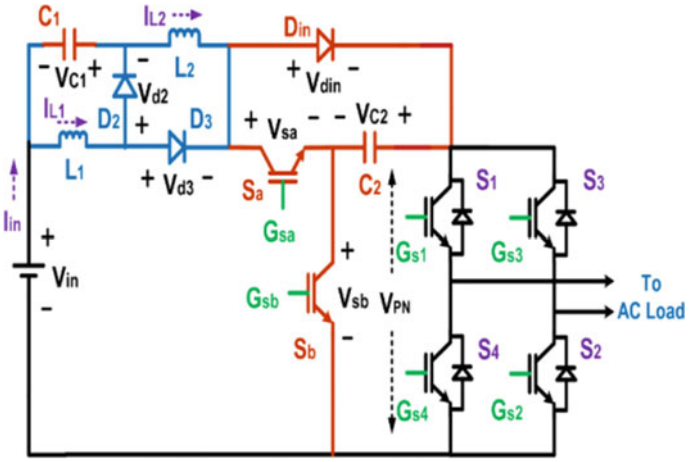



**Fig. 2** Type 2 SLC-ZSI

## *3.1 SLC-ZSI Operation-Type 1*

With the help of detailed investigation, the single switch s is changed with the help of H-bridge inverter circuit. The operation of the Type 1 SLC-ZSI and shoot-through and non-shoot through periods are explained.

**Shoot-Through Period.** Diodes $D_1$ and $D_2$ which are forward biased and $D_2$ and $D_{in}$ are made reverse biased by turning on the switches S and $S_a$ and turning off the switch S. The capacitor *C* discharges and the DC source and capacitor provides energy to inductors with help of source S and $S_a$. Figure 3 shows the Shoot-through interval of Type 1 SLC-ZSI.

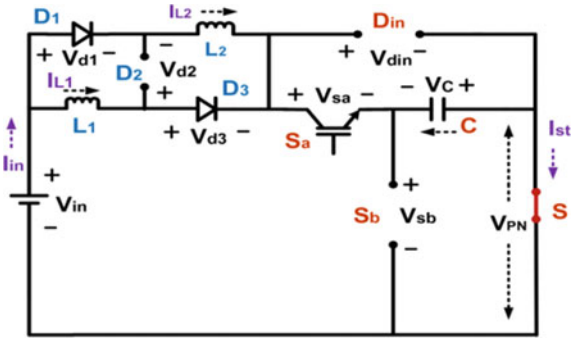



**Fig. 3** Shoot-through interval of Type 1 SLC-ZSI

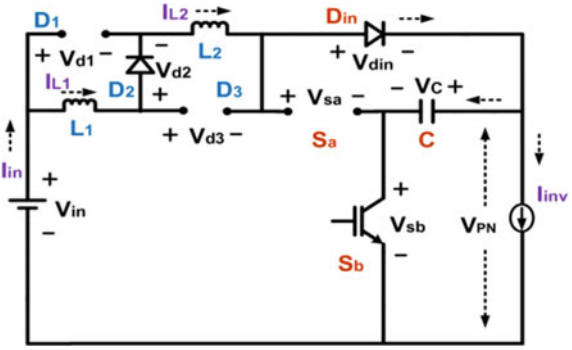



**Fig. 4** Non-shoot-through interval of Type 1 SLC-ZSI

**Non-shoot-through Period.** The diodes $D_1$ and $D_3$ are made to inverse biased and $D_2$ and $D_{in}$ are made to forward biased which makes $S$ and $S_a$ to turn OFF and $S_b$ is turned ON.

Figure 4 illustrates the Non-shoot-through interval of Type 1 SLC-ZSI. With these conditions, where $L_1$ and $L_2$ are held in series and the energy of DC power source and the stored magnetic energy is provided to the circuit which is connected to the inverter along with capacitor to be charged.

## *3.2 Process of Type 2 SLC-ZSI*

With the simple investigation, a single mode switch $S$ is substituted for the H-bridge inverter circuit. The shoot-through and non-shoot through periodic operations are described.

**Shoot-Through Interval**

Figure 5 shows the Shoot-through interval of Type 2 SLC-ZSI, While the switches $S$ and $S_a$ are kept in the ON an $S_b$ is kept in OFF position, diodes $D_2$ which is forward biased and $D_2$ and Din which are made to backward bias. The capacitors in the circuit ($C_1$ and $C_2$) runs down of its charge and the inductors helps to store energy from DC and S and $S_a$ are charged through capacitors.

**Non-shoot-through Interval**

Figure 6 shows the non-shoot-through interval of Type 2 SLC-ZSI. At the same time, if $S$ and $S_a$ are kept in OFF position and $S_b$ is kept in ON position, $D_3$ is made to backward bias and $D_2$ along with Din are made to forward bias. The energy along with DC source is sued to store up magnetic energy from the inductors and the inverter receives it. The capacitors are charged equally.

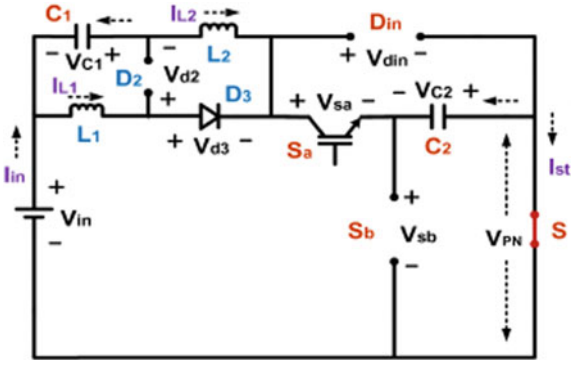



**Fig. 5** Shoot-through interval of Type 2 SLC-ZSI

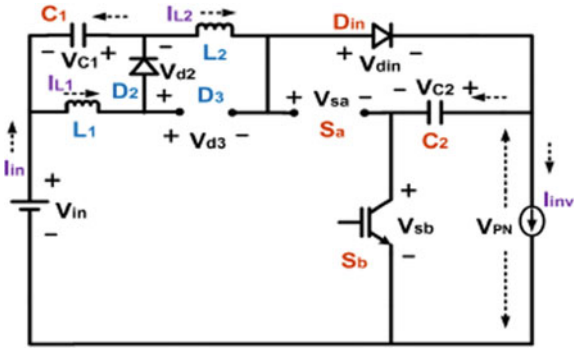



**Fig. 6** Non-shoot-through interval which is of Type 2 SLC-ZSI

## 4 Results and Discussions

The Simulation model for the 3 phase high voltage gain switched Inductance (L), Capacitance (C) Z-source inverter.

### *4.1 Output Analysis*

The voltage analysis and current analysis of the proposed 3φ high gained voltage switched LC Z-source inverter are given below. Steady state investigational results of SLC-ZSIs are given for alternate values of duty ratio D which helps to manage AC voltage levels. The proposed inverters use the Modulation index (M) and Duty ratio (D) which are dependent in shoot-through unipolar PWM technique.

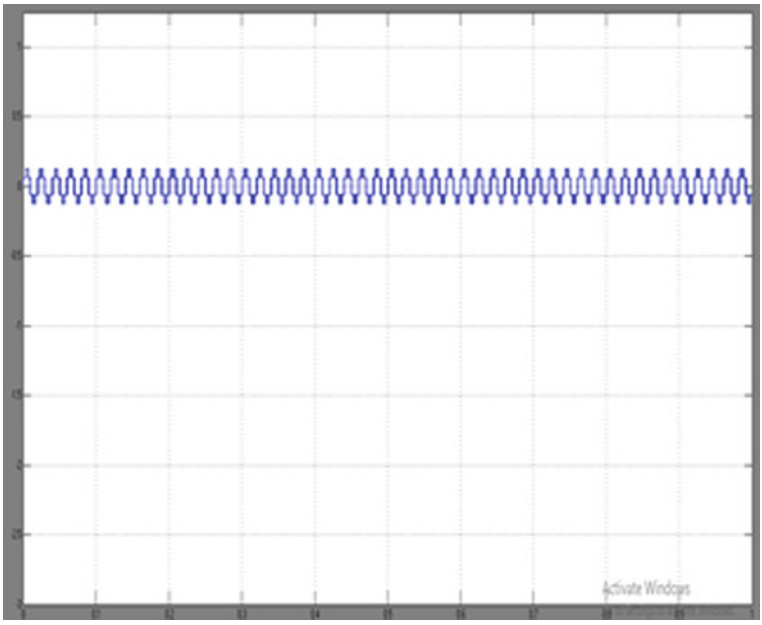


**Fig. 7** Voltage analysis of 3-phase switched LC Z-source inverters

Figures 7 and 8 describes the voltage and current analysis of the 3φ switched LC Z-source inverter. The input voltage that is provided as the input is boosted with the help of the proposed 3-phase Z-source inverter. The input current which is given as the input is boosted through this proposed three-phase Z-source inverter.

## *4.2 Output for Type 1 3-Phase Switched LC Z-Source Inverter*

Figure 9 illustrates the Simulink model for Type 1 switched LC Z-source inverters. Switch S replaces the H-bridge inverter circuit. It describes the overall model for which the type 1 switched LC inverters are designed. This Simulink model is designed in such a way to easily identify the shoot-through and the non-shoot-through interval.

This Fig. 10 shows the input that is given for the Type 1 three-phase switched LC Z-source inverter. The input is given in the form of DC voltage.

This Fig. 11 illustrates about boosted DC voltage of the Type 1 switched LC Z-source inverter and DC voltage is given as input is boosted.

This Fig. 12 demonstrates the o/p waveform of the 3-phase switched LC Z-source inverter in the form of load voltage and load current.

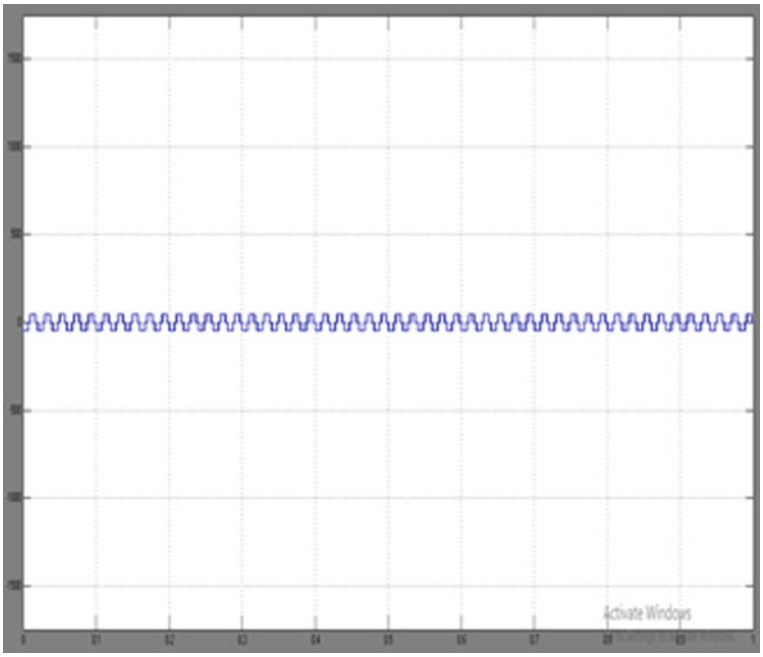


**Fig. 8.** 3-phase switched LC Z-source inverters-current analysis

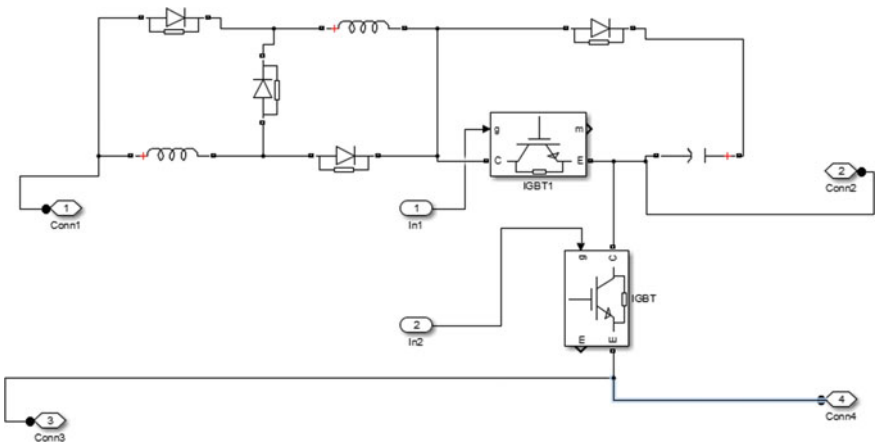



**Fig. 9** Simulink model for Type 1 switched LC Z-source inverters

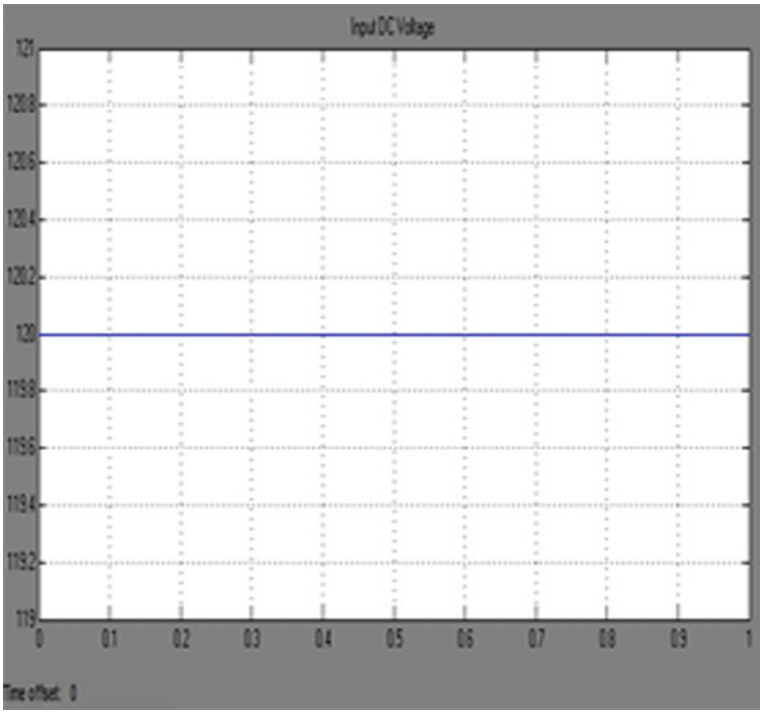



**Fig. 10** Input voltage for Type 1 3-phase switched LC Z-source inverter

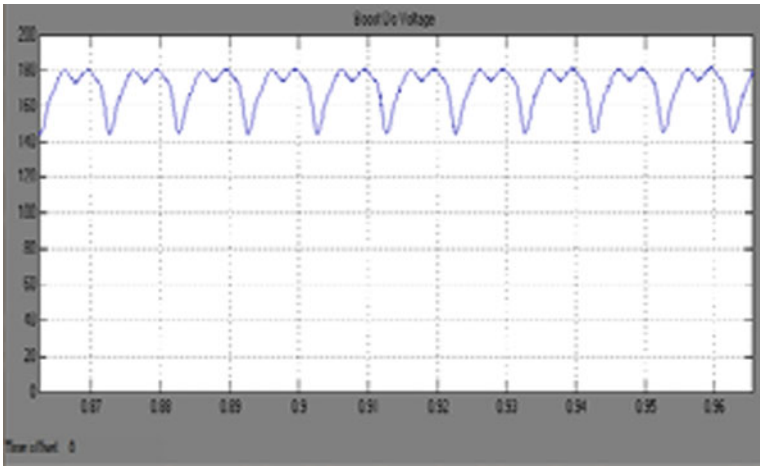



**Fig. 11** Boost DC voltage of Type 1 three-phase switched LC Z-source inverter

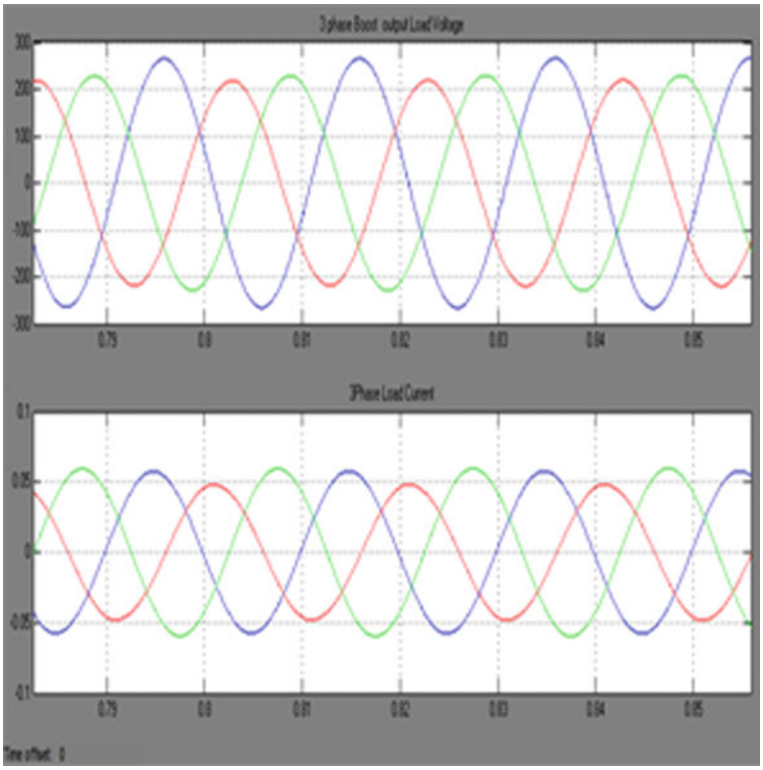



**Fig. 12** Output waveform of 3-phase Z-source inverter load voltage and load current

This Fig. 13 explains the Gate Pulse waveforms of the proposed inverters.

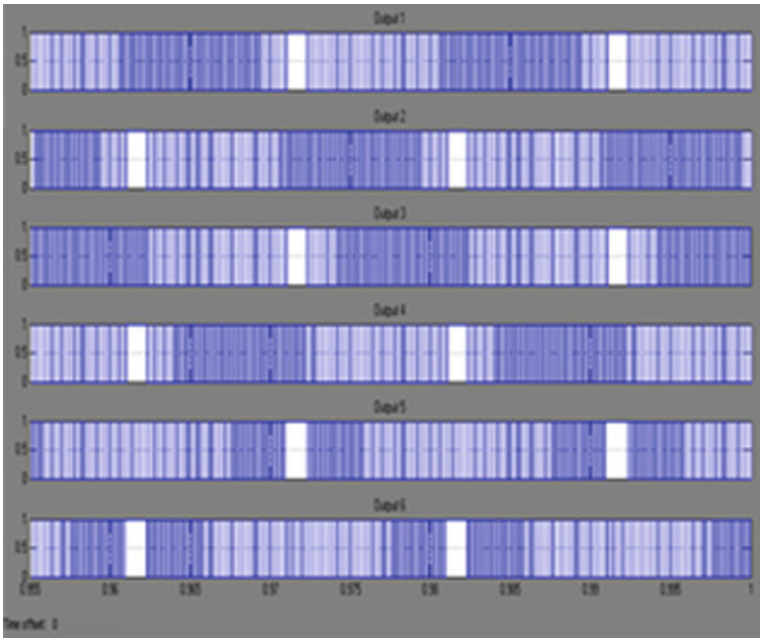


**Fig. 13** Gate Pulse Output Waveform

## *4.3 Output Parameters*

Table 1 tabulates Parameters derived from SLC-ZSI and SLC-ZSI of type 1 and type 2. From the Table 1, it clearly indicates that SLC-ZSI-TYPE 2 achieves the high voltage gain of 1.83 (G) which is much better than SLC-ZSI-TYPE 1.

Table 2 tabulates the comparative analysis of THD. From the Table 2, it indicates that proposed SLC-ZSI-TYPE 2 achieves less THD in both voltage and current analysis. The proposed SLC-ZSI-TYPE 2 attained 3.69 and 2.52% of THD while the existing Z-source inverter [14] has reached 4.09 and 2.93% only.

**Table 1** Parameters derived from SLC-ZSI and SLC-ZSI of type 1 and type 2

| Parameters | SLC-ZSI-TYPE 1 | SLC-ZSI-TYPE 2 |
|---|---|---|
| DC input voltage (V) | 120 | 120 |
| DC boost voltage (V) | 200 mean | 220 mean |
| AC output voltage (V) | 200 avg | 220 avg |
| Voltage gain factor (G) | 1.66 | 1.83 |

**Table 2** Comparative analysis of THD

| Performance | THD analysis | |
|---|---|---|
| | Z-source inverter [14] (%) | Proposed SLC-ZSI-TYPE 2 (%) |
| THD at output voltage | 4.09 | 3.69 |
| THD at output current | 2.93 | 2.52 |

## 5 Conclusion

In this research, the two 3-phase SLC-ZSIs are proposed. These are capable of providing optimum gains of voltage at low shoot-through values that has reduced number of inactive components while comparing with the Z-source inverters. At allowable THD, modulation index can be made optimum by the presence of low values of *D,* which provides high AC o/p voltages. Moreover, the proposed inverters are more efficient because of its utilization of minimal number of components and its size and heaviness when compared with the present system SL-ZSIs. From the simulation results, it demonstrates that proposed SLC-ZSI-TYPE 2 accomplished THD of 3.69% and 2.52% which is better than existing methods. A complete study and steady-state observation that anticipated both the types of SLC-ZSIs is performed and the expected results of the invertors which are already proposed are verified via investigation and outcomes at different values of *D*.

## References

1. Ahmad A, Busssa VK, Singh RK et al (2017) Single-phase high-voltage gain switched LC Z-source inverters. IET Power Electron 11:796–807
2. Axelrod B, Berkovich Y, Ioinovici A (2008) Switched-capacitor/switched-inductor structures for getting transformer less hybrid DC-DC PWM converters. IEEE Trans Circuits Syst I Fund Theory Appl 55(2):687–696
3. Barathy B, Kavitha A, Viswanathan T (2014) Effective space vector modulation switching sequence for three phase Z source inverters. IET Power Electron 7(11):2695–2703
4. Carrasco JM, Franquelo LG, Bialasiewicz JT et al (2006) Power-electronic systems for the grid integration of renewable energy sources: a survey. IEEE Trans Ind Electron 53(4):1002–1016
5. Dong S, Zhang Q, Cheng S (2016) Analysis and design of snubber circuit for Z source inverter applications. IET Power Electron 9(5):1083–1091
6. Liu P, Liu HP (2012) Permanent-magnet synchronous motor drive system for electric vehicles using bidirectional Z-source inverter. IET Electr Syst Transp 2(4):178–185
7. Loh PC, Vilathgamuwa DM, Gajanayake CJ et al (2007) Transient modeling and analysis of pulse-width modulated Z-source inverter. IEEE Trans Power Electron 22(2):498–507
8. Peng FZ (2003) Z-source inverter. IEEE Trans Ind Appl 39(2):504–510
9. Rajakaruna S, Jayawickrama L (2010) Steady-state analysis and designing impedance network of Z-source inverters. IEEE Trans Ind Electron 57(7):2483–2491
10. Ramasamy BK, Palaniappan A, Yakoh SM (2013) Direct-drive low-speed wind energy conversion system incorporating axial-type permanent magnet generator and Z-source inverter with sensorless maximum power point tracking controller. IET Renew Power Gener 7(3):284–295

11. Ding X, Liu Y, Zhao D, Wu W (2019) Generalized Cockcroft-Walton multiplier voltage Z-source inverters. IEEE Trans Power Electron 35(7):7175–7190
12. Hou T, Zhang C-Y, Niu H-X (2020) Quasi-Z source inverter control of PV grid-connected based on fuzzy PCI. J Electron Sci Technol 2020:100021
13. Belila A, Berkouk E-M, Benbouzid M, Amirat Y, Tabbache B, Mamoune A (2020) Control methodology and implementation of a Z-source inverter for a stand-alone photovoltaic-diesel generator-energy storage system microgrid. Electric Power Syst Res 185:106385
14. Faruqui SNA, Anwer N (2019) Performance evaluation of Z-source inverter and voltage source inverter for renewable energy applications. Int J Energy Water Resour 3(1):43–53
15. Kumar SP (2018) Modern control scheme for Z source inverter-based PV power generation systems. Int J Eng Technol 7(1.3):166–170

# Prediction Scheme Using Fuzzy Logic System to Control the Congestion in Wireless Sensor Network


**Zainab G. Faisal, Maysam Sameer Hussein, and Amany Mohammad Abood**

**Abstract** A key challenge of Wireless Sensor Networks (WSNs) is congestion, this is because of the resource-constrained of WSN. The nodes that have one hop to the sink (node near the sink) have a chance of congestion because of the heavy data traffic through it to reach the sink when an event is detected. So the soft computing based on fuzzy logic is proposed. The proposed approach is called Fuzzy Logic Congestion Controller (FLCC) divided into 4 stages, (1) detecting and estimating the congestion as congestion level indications, (2) control the traffic rate using a fuzzy logic system, (3) adjust the rate of active nodes and (4) notify it by its new rate. With several qualities of service factors, such as packet loss ratio, remaining energy, buffer utilization, and throughput, FLCC was applied to determine proper weights among those factors. According to the simulation results, the proposed controller can avoid and estimate the congestion in WSN and show superior performance in terms of increased throughput, reduced packet loss, and balancing the overall energy consumption.

**Keywords** Congestion control · Energy consumption · Fuzzy logic · Soft computing · Wireless sensor network

## 1 Introduction

Although the resource-constrained of the sensor nodes it involved other functions, including sensors, computing, storage and transmission, but internal battery supply shortens the overall network life [1, 2]. Wireless sensor networks comprise of a large number of sensor nodes when an event is detected, many sensor nodes will be active

Z. G. Faisal (✉) · M. S. Hussein · A. M. Abood
Computer Engineering Techniques Department, Al Esraa University College, Baghdad, Iraq
e-mail: zainab.ghazi@esraa.edu.iq; zainabghazi89@gmail.com

M. S. Hussein
e-mail: maysam@esraa.edu.iq

A. M. Abood
e-mail: amany@esraa.edu.iq



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_68

and begin to send its data through intermediate nodes to reached sink [3]. Because the data is not fairly distributed over the all node, the sensor nodes that have one hop away from the sink node (node near sink) may be bottleneck and lead to lose the data packet [4]. The impact of congestion is more regrettable as congestion in wireless sensor network led to packet drop and energy utilization due to a more sensor is active to send or resend the packet. Indeed, even the general throughput of the network conveyance proportion debases because of congestion [5]. The utilization of fuzzy logic is suitable to enhance the performance of WSN [6]. Thus the approach (FLCC) has been proposed to estimate and control congestion at the network by adjusting the transition rate of the active sensor, and then the non-linearity in the traffic load can be reduced. The FLCC approach needs to be simple enough to run unhindered on individual sensor nodes without overly taxing the system resources.

## 2 Related Work

Many researchers have study congestion control in WSNs and try to reduce packet loss in different ways. So a fuzzy based congestion detection and alleviation is proposed by taking some metrics like Buffer occupancy, traffic rate, contenders, number of packets sent by sender sensor node, number of packets received by the receiver sensor node and level of congestion. All these metrics help to estimate congestion level in the sensor network and also minimizes the congestion by adjusting rest of the other metrics. The performance of the proposed is compared with existing schemes with respect to various qualities of service parameters [7]. The network life time is improved by control the rate process of the active nodes based on cluster routing. As a result, congestion is reducing throughout the network [8]. Also in [9], proposed transmission rate control to avoid the congestion by adjusted the transmission rate at active node based on its packet loading information, congestion degree and queue length. In [10], a method based on fuzzy is proposed to manage the sleeping time of the sensor node in smart home. On the other hand, in [4], the Proportional Integral Derivative controller is proposed to control the congestion by using queue length of messages in the nodes. In addition, a fuzzy logic control is proposed to solve the problems of slow parameter optimization of traditional PID controller so the proposed mechanism is combined the fuzzy with Proportional Integral Derivative (PID) to adjust queue length of messages in node. In [11], a fuzzy sliding mode controller is proposed, that combine between fuzzy logic control and sliding mode control, which regulates the queue length in the congested nodes and significantly reduces the uncertain disturbance of the traffic. Also in [12], congestion controller based on neural network was proposed at sink node to reduce congestion, buffer occupancy used as input to the neural network and according to the decision of controller, the transmission rate of active sensor node is adjustable.

## 3 Methods

This section is divided in two sections. The first section explains the network model, the second section describes the proposed control for congestion estimation and avoidance.

### *3.1 The Model of the Network*

The model of network consists of a set of N sensor nodes, randomly distributed over coverage area and one sink node in center of the area. Two activities have been in the sensor network node, nodes can generate data traffic and route it to other nodes [1–3]. In general, the event-driven sensor networks is acceptable load condition [10] if events detected, the sense's node has been become active and encode the information into packets and forwarded it via multiple hops to a sink using shortest path routing. Bit Rate BR(Si) is the transmit rate used at each sensor Si to forward data to the sink. It must be acceptable to the available capacity of the channel in the network. The Eq. (1) explained it [9].

$$\sum_{i=1}^{N} BR(S_i) \leq Channel\ \ capacity \tag{1}$$

### *3.2 The Proposed Fuzzy Logic Congestion Control*

The components of FLCC are four units. Fuzzy Congestion Detection Unit (FCDU) which is used to detect and estimate if the congestion has occurred, Fuzzy Logic Control Unit (FLCU) used to control the congestion using fuzzy logic member ship, Fuzzy Rate Adjustment Unit (FRAU) calculates the new rate for all active nodes and the new rate is sent to the Fuzzy Rate Notification Unit (FRNU) which is notify all active nodes of the new rate [13]. Figure (1) illustrates the block diagram of the proposed FLCC approach that designs each node near the sink.

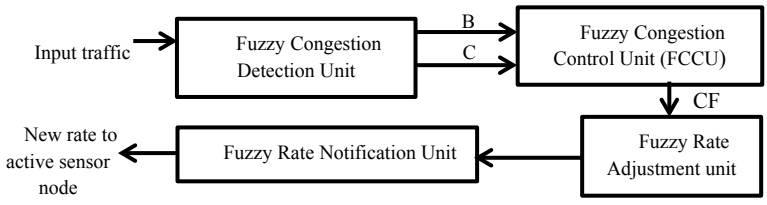



**Fig. 1** The block diagram of the FLCC

**Fuzzy Congestion Detection Unit (FCDU).** Accurate and efficient detection unit is played an important role to control the congestion in the sensor networks. To detect the congestion, the level of it should be quantified to obtain a fine-grained congestion control. When the detection unit detects congestion, it refers to identification of possible events that may build-up congestion in the network.

Congestion detection is done by calculating the Buffer occupancy $B_O$ at time $t$ and Congestion indication $C_I$ of the traffic that follow through the node. The latter metric $C_I$ reflects the packet loss ratio [9].

The packet loss ratio value is the ratio between numbers of the losses packet at considered node to the numbers of packets generated by sensing nodes which is sends their data through the considered node to reach the sink. $B_O$ is calculated using accumulation packets in the buffer of that considered node. The function $B_{ON}$ in Eq. (2) denoted the amount of buffer occupancy at node N [9].

$$B_{ON} = \frac{B_{OI}}{B_{OT}} \tag{2}$$

where, $B_{OI}$ represents the number of packets that reached to queue buffer at node N, and $B_{OT}$ represents the buffer size at the same node. Clearly, that the value of $B_{ON}$ is always in the range of [0, 1]. In the worst case scenario, $B_{ON}$ is 1. Hence, it means $B_{OI} = B_{OT}$.

For more explanation, Fig. 2 show nodes 1, 2 and 3 that are the active nodes to forward their data packet to the sink through node 4. Fuzzy Congestion Detection Unit in node 4 calculates $C_{I4}$ (t) and $B_{O4}$ (t) where the value of $C_{I4}$ is defined in Eq. 3.

$$C_{I4}(t) = \frac{number\ of\ packets\ loss\ at\ node\ 4\ in\ network\ (PL_4(t))}{number\ of\ packets\ generated\ by\ sensing\ nodes\ (1, 2, 3)\ at\ time\ t} \tag{3}$$

and $B_{O4}$ (t) is occupancy of buffer at node 4 at time t.

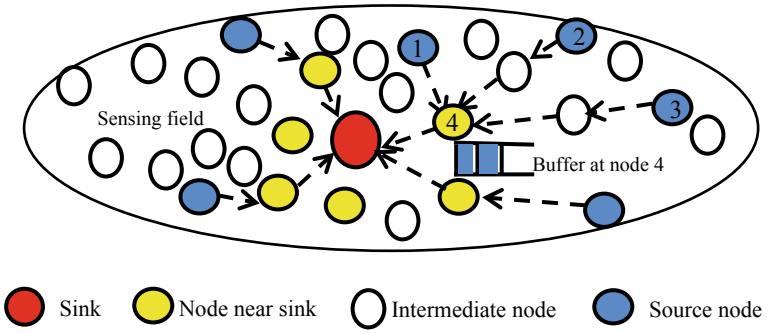
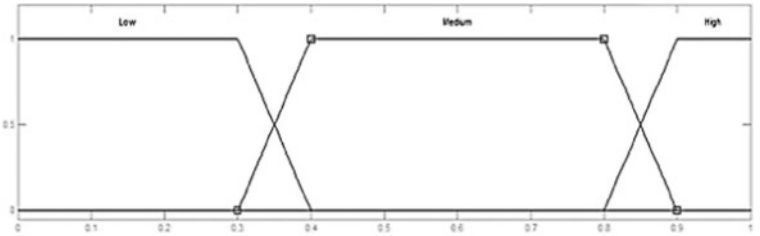



**Fig. 2** Structure of WSN show traffic flow through node 4

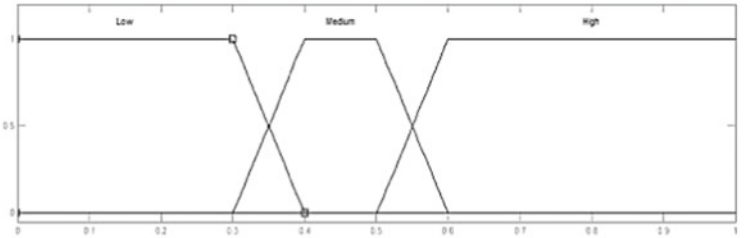
**Fuzzy Congestion Control Unit (FCCU).** The BO and CI of each node near the sink are taken as input parameters to FCCU. The Membership Function (MF) of BO is chosen to explain the selection of the Membership Function (MF) in this unit. The [0,1] is the range of BO. The greater BO refers to the node becomes congested. When the value of BO exceeds 0.9, the congestion probability in that node is 1. The linguistic values for inputs MFs are [L, M, and H] representing Low, Medium, High values, respectively. Similarly, the membership function of $C_I$ can be obtained. Figure 3 shows the input and output value membership functions.

The linguistic values of output MFs are [L, M, H] representing Low, Medium, High, respectively. Those linguistic values of the Fuzzy Congestion Level (FLC) increase from left to right [6]. Table 1 is show a typical set of FCL evaluation rules.

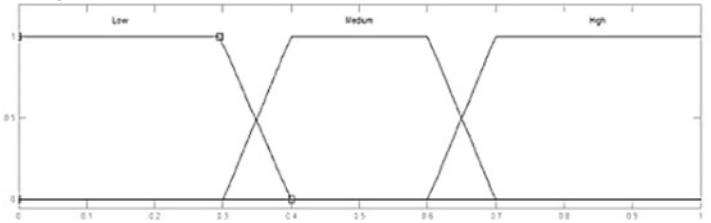
The table is therefore the fuzzy table to determine the level of congestion which takes the values of $B_O$ and $C_I$ as input and Fuzzy Congestion Level (FCL) as a single output, the defuzzified value of which determines the level of congestion in sensor nodes.




Buffer occupancy




Congestion indication




Fuzzy Congestion Level

**Fig. 3** The input and output membership functions

**Table 1** Fuzzy rules

| Fuzzy congestion level FCL | 1. Buffer occupancy $B_O$ | | |
|---|---|---|---|
| | L | M | H |
| Congestion indication $C_I$ | L | M | M |
| | M | H | H |
| | M | H | H |

The process of Center of Gravity (CoG) defuzzifier is applied to get the crisp FCL in the proposed FLC. Finally, obtaining the value of FCL is in the range of [0–1], higher values representing more congestion in wireless sensor networks.

Three levels that Congestion state is divided where two threshold boundaries $B_{min}$ and $B_{max}$ ($0 < B_{min} < B_{max} < 1$) are involved. The following regulation explains it where CL is congestion level.

$$CF = \begin{Bmatrix} -1 & 0 < CL < B_{min}(nocongestion) \\ 0 & B_{min} < CL < B_{max} \\ 1 & B_{max} < CL < 1(congestion) \end{Bmatrix}$$

where CF is the Congestion Flag used in Fuzzy Rate Adjustment Unit (FRAU). The CF value helps making forwarding decisions at the nodes, especially at times of incipient congestion. Matlab is suitable to implement the proposed controller and clarity its effect on quality of service of network [14, 15].

**Fuzzy Rate Adjustment Unit (FRAU).** When CF is 1, FRAU understands the level of congestion is very large so it attempts to decrease sending rate BR of each active node Si and when the value is −1, there is no congestion occur but FRAU is tried to make the incoming traffic almost equal to effective buffer utilization thr.BS so it tries to make the sending rate of each active node increased. Moreover, no congestion has occurred in case the value of CF is 0 but the rate of active nodes remains the same because the current buffer occupancy is critical so it leaves relatively enough space available in the buffer if the node suddenly becomes active.

The sending rate of active nodes should be adjusted immediately to obviate congestion problem at current node and this is done by Fuzzy Rate Adjustment Control. The FRAC fairly apportions rate between active nodes by Eq. (4).

$$R(S_i(t+1)) = thr.BS.\frac{R(S_i(t))}{\sum_{i=1}^{n} R_i} \tag{4}$$

where the number of active nodes is ***n*** that is send their data through nodes near to the sink. This value allows the node to take an informed choice of forwarding a packet to the sink node keeping in mind both the aspects of fairness and load balancing.

**Fuzzy Rate Notification Unit (FRNU).** In FRNU, each active node obtained its new rate from FRAU that is piggybacked with the ACK massage. The active node adjusts

its rate according to the received value taking into consideration the maximized and minimized sending rate.

## 4 Results

In order to solve the problem of packet loss due to congestion in WSN, this manuscript introduced FLCC to mitigate and estimate the amount of packet that passing through the node near the sink.

The proposed controller FLCC is compared with [12] Modified Neural Network Wavelet Congestion Control MNNWCC based on neural network that have wavelet as activation function and Neural Network Congestion Control NNCC based on neural network that have sigmoid as activation function. The MNNWCC and NNCC were implemented at a sink node to avoid and estimate the congestion.

In the comparison experiment, the simulation parameter was mentioned in Table 2 to implement the controllers for accurate and fair simulation comparison. The simulation has been conducted in MATLAB program.

The model of the network comprises of 100 sensor nodes randomly deployed in the coverage area and the sink in the center of it. Different situation is used to examine the performance of the FLCC and compare it with network Without Congestion Controller WCC, MNNWCC and NNCC.

Figure 4 Show the number of packets that was reached to the sink along the simulation time. It is clear; the amount of traffic is change, sometime low and anther high, according to the Event occurrence at the coverage area.

**Table 2** Simulation parameters

| | |
|---|---|
| Total area | 100 × 100 m |
| Nodes | 100 node |
| Sinks | 1 |
| Active sensor node | 10–40 node |
| Buffer size of sink | 250 packet |
| Maximum Threshold of sink buffer | 225 packet |
| Buffer size of node | 50 packet |
| Maximum Threshold of node buffer | 45 packet |
| Data packet size | 512 byte |
| Simulation time | 120 ms |
| Channel Bandwidth | 2(Mega bit per msec.) |
| Traffic type | No priority |
| Routing | Shortest path algorithm |
| Data packet generation rate | 10–30 (packet / msec.) |

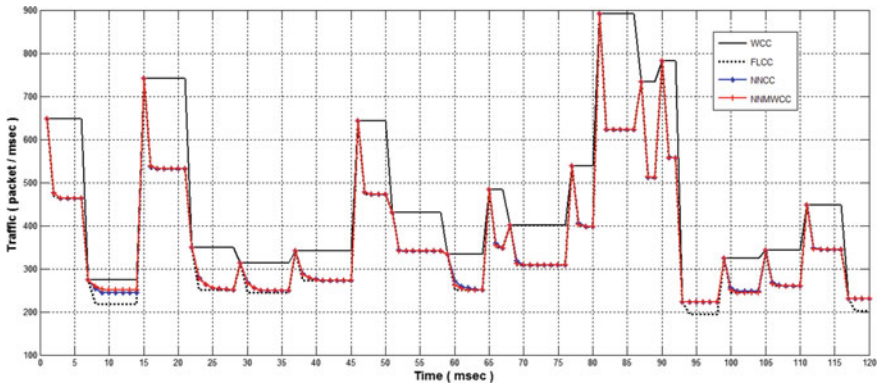



**Fig. 4** The traffic reached to the sink node at specific time

Also notice from Fig. 4, the proposed controller was sensed this existence of this hustle better than MNNCC and NNCC and has modified the sending rate of the active node to acceptable value to avoid congestion.

Figure 5 clearly shows the performance of the FLCC in term of packet loss ratio is better than network without controller because the intelligent methodologies of the FLCC. Also the proposed controller attempt to avoid packet loss faster the other controller previously applied.

Figure 6 clarifies the buffer utilization ratio for the network with FLCC and compare the performance with WCC, MNNCC and NNCC. Figure 7 and Fig. 8 shows the Network Throughput Ratio and Network Energy Consumption Ratio of the network respectively. All the results obtained from the figures shows that the network with FLCC tries to improve the performance by reduced packet loss,

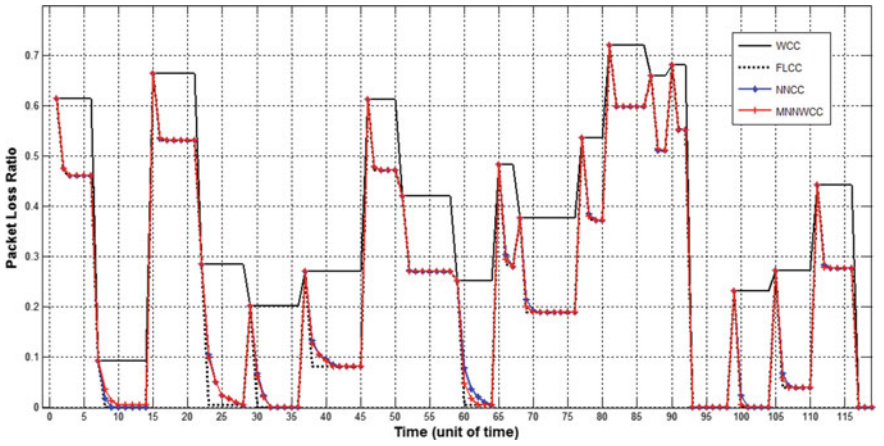



**Fig. 5** Packet loss ratio of the network

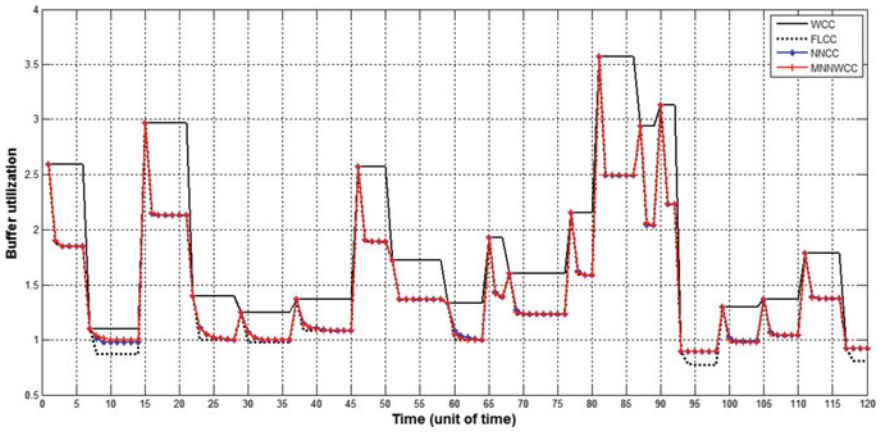



**Fig. 6** The buffer utilization of network

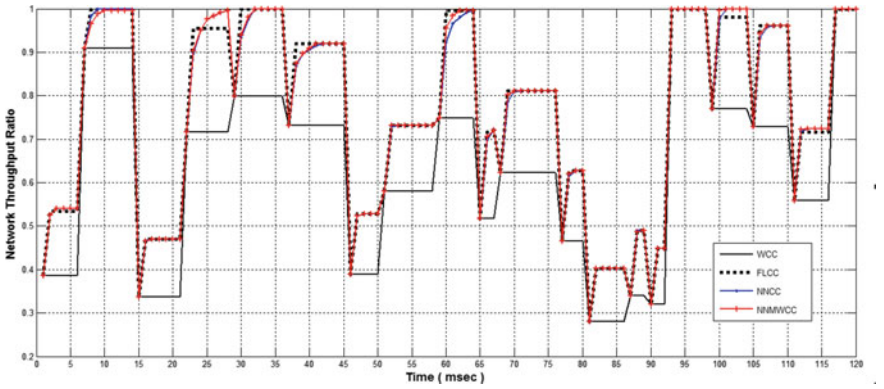



**Fig. 7** The network throughput ration of network

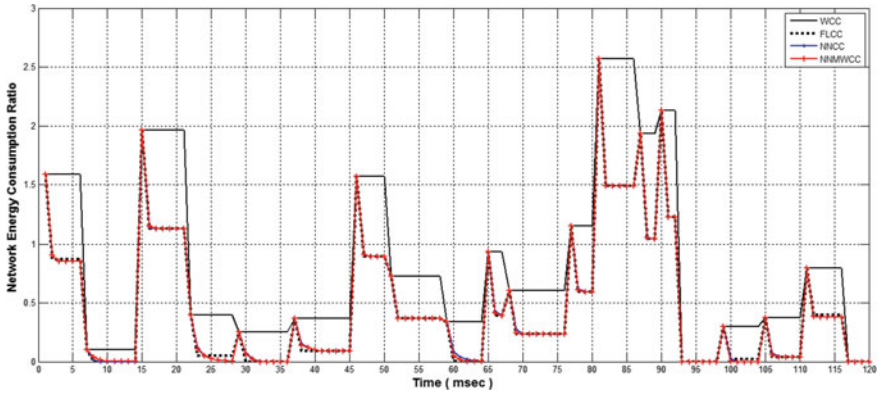



**Fig. 8** The network energy consumption ration of the network

improved buffer utilization, efficiently network throughput and energy consumption at sensor is reduced.

## 5 Discussion

From the result, the proposed controller FLCC tries to avoid losing data with the least possible means and as quickly as possible to conserve the energy of the sensor. The MNNCC and NNCC tried to prevent congestion in WSNs by controlling the traffic flow in the sink node and adjusting the sending rate of the active nodes fairly but without taking into account possible congestion in other sensors. The proposed FLCC approach provides better quality of service than the other approaches, by controlling the sending rate of active sensor nodes and estimated the network traffic efficiently so the actual traffic faster modified to reaches the desired buffer occupancy; as a result, it will improve the buffer utilization.

## 6 Conclusion

In this paper, the proposed approach was introduced based on intelligent algorithm that tried to prevent and estimate congestion in WSNs. Buffer occupancy and congestion indication was tacked in account to control the traffic flow through the node near the sink and fairly adjusting the sending rate of active nodes. This leads to making the network traffic at an acceptable level to fit the available space of the sink buffer. Finally, the FLCC is used to optimize the packet loss ratio globally.

In the future, designing a cluster network and implementing the proposed controller on each cluster head and also more energy efficient congestion control based on intelligent algorithm can be applied and compare the results.

## References

1. Singh AP, Luhach AK, Gao XZ, Kumar S, Roy DS (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
2. Barath SR, Usha BU, Arpitha GB, Kumar V, Yaseen SM (2020) Strategies for congestion control in wireless sensor networks a survey. In: International Conference on Smart Systems and Inventive Technology (ICSSIT), IEEE. https://doi.org/10.1109/ICSSIT46314.2019.8987931
3. BenSaleh MS, Saida R, Kacem YH, Abid M (2020) Wireless sensor network design methodologies: a survey. Hindawi J Sens 2020:1
4. Lin L, Shi Y, Chen J, Ali S (2020) A novel fuzzy PID congestion control model based on cuckoo search in WSNs. In: MDPI, 27 March 2020
5. Prabhu S, Chandrasekar C (2020) Congestion control in wireless sensor networks using TOPSIS and percolate congestion control algorithm (PCCA). Int J Adv Sci Technol 29(5):13573–13575

6. Mishra PK, Verma SK (2021) FFMCP: feed-forward multi-clustering protocol using fuzzy logic for wireless sensor networks (WSNs). In: MDPI, Energies
7. Mishra P, Jaiswal S, Agarwal N (2019) Expanding lifetime of wireless sensor network using fuzzy logic. In: First International Conference on Secure Cyber Computing and Communication (ICSCCC), IEEE
8. Srivastava V, Tripathi S, Singh K, Son LH (2020) Energy efficient optimized rate based congestion control routing in wireless sensor network. J Ambient Intell Humaniz Comput 11:1325
9. Kazmi HS, Javaid N, Awais M, Tahir M, Shim SO, Zikria YB (2019) Congestion avoidance and fault detection in WSNs using data science techniques, WILEY
10. Pau G, Salerno VM (2019) Wireless sensor networks for smart homes: a fuzzy-based solution for an energy-effective duty cycle. In: MDPI Electronics
11. Qu S, Zhao L, Xiong Z (2020) Cross-layer congestion control of wireless sensor networks based on fuzzy sliding mode control. Neural Comput Appl 32:1355
12. Aboubakar M (2020) Efficient management of IoT low power networks, thesis, Université deTechnologie de Compiègne, France
13. Nawaz B, Mahmood K, Khan J, ul Hassan M, Shah AM, Saeed MK (2019) Congestion control techniques in WSNs: a review. Int J Adv Comput Sci Appl 10:194–197
14. Zhao L, Qu S, Huang X, Luo J (2019) Congestion control of wireless sensor networks using discrete sliding mode control. In: Chinese Control and Decision Conference (CCDC) IEEE. https://doi.org/10.1109/CCDC.2019.8833135
15. Faisal ZG, Hussein MS, Abood AM (2019) Design and realization of motion detector system for house security. Telecommun Comput Electron Control 17(6)

# Real-Time Heel Strike Parameter Estimation for FES Triggering


**Haaris Rahman, Ashwij Kumbla, V. N. Megharjun, and Viswanath Talasila**

**Abstract** For patients with certain types of movement disorders, electrical stimulation is used to facilitate the rehabilitation process. Typically, rehabilitation specialists manually control the triggering of the electrical pulse at desired points in the gait cycle, and this manual intervention can be erroneous. In this paper, we develop a real time gait parameter estimation model using artificial neural networks to activate electrical stimulation at a desired point in the gait cycle. The weights obtained from the neural network model are trained across a sample of similar aged population. The predicted output obtained from the real-time model were compared with the output of the offline analysis. We further compare results from a specific model trained with data of a single individual with that of a general model trained with data of eight individuals. A part of this work was funded by a BIRAC project BT/AIR0945/PACE-19/19.

**Keywords** Assistive devices · Gait analysis · Heel strike · Inertial measurement unit · Neural networks · Real-Time gait detection · Wearable sensors

## 1 Introduction

Human gait analysis involves the understanding of various patterns of body movement. These patterns vary among individuals depending on their age, gender, weight, and height. The understanding of these patterns helps in clinical rehabilitation of patients with locomotor disabilities [1] as well as those with stroke and spinal cord

H. Rahman · A. Kumbla
Electronics and Instrumentation Engineering, M S Ramaiah Institute of Technology, Bangalore, India

V. N. Megharjun
University of Pennsylvania, Philadelphia, PA 19104, USA
e-mail: manvn@seas.upenn.edu

V. Talasila (✉)
Electronics and Telecommunication Engineering, Center for Imaging Technologies, M S Ramaiah Institute of Technology, Bangalore, India
e-mail: viswanath.talasila@msrit.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_69

injuries [2]. Additionally, gait analysis is used in fall risk estimation [3] in the elderly and used in sports training [4] to identify shortcomings in athletes' movement that may hinder their performance.

In the medical field, gait analysis is generally performed by trained clinicians through observation, or in gait laboratories with specialized cameras and equipment [5]. However, these tests are expensive and access to such equipment is limited [6]. Further, for cerebral palsy patients with foot drop, functional electrical stimulation (FES) is used to facilitate the rehabilitation process [7]. Typically, this requires precise hand–eye coordination to trigger the electrical pulse at the accurate time of the gait cycle. An incorrect trigger instance resulting from a lack of concentration could risk the fall of these patients.

In this paper, we develop a real-time gait parameter estimation device as a portable and cost-efficient alternative to activate electrical stimulation at the optimal point in the gait cycle. An Inertial Measurement Unit (IMU) is used to collect spatiotemporal data and a Force Sensitive Resistor (FSR) is utilized to determine when a subject's foot encounters the ground, determining his/her full gait cycle. The data obtained from a sample similar-aged population is used to train an artificial neural network to predict one of the phases of the gait cycle, i.e., heel strike, which is then implemented on a microcontroller for real-time estimation.

## 2 Related Works

A human walk consists of gait cycles; a gait cycle is completed when there is a successive contact of the same foot with the ground [8, 9]. Several studies have proposed using inertial measurement units (IMU) as a reliable method for gait analysis [10, 11]. In [12], the use of long-short term memory networks (LSTM) is used to identify foot contact from motion capture systems, and in [13], the use of IMUs and thresholding algorithms is used to identify gait parameters for triggering the functional electrical stimulus. Due to the time-critical nature of triggering a FES at the appropriate time in the gait cycle, the use of deep neural networks, such as pre-trained models or LSTMs, particularly on embedded systems for real-time gait parameter estimation is impractical due to large computational times of the network. Further, thresholding algorithms are not as versatile as machine learning algorithms in capturing unique patterns of the gait cycle.

As described above and to the knowledge of the authors, there is no work on using a standalone embedded system for gait parameter estimation in real-time. Further, the low computational times of traditional neural networks ensure the time-critical operation of triggering a FES.

# 3 Experimental Setup

## *3.1 Sensors Used and Placement*

The wearable device consists of an ESP8266 microcontroller, with a full TCP/IP stack and IEEE 802.11 b/g/n WLAN, interfaced with a BNO-055 inertial measurement unit (IMU), shown in Fig. 1(a). The microcontroller and BNO-055 communicate via bi-directional I2C and UART interfaces. The IMU provides a 9-axis absolute orientation consisting of a triaxial 14-bit accelerometer, a triaxial 16-bit gyroscope with a range of $\pm 2000°$ per second, and a tri-axial geomagnetic sensor. This setup is placed on the metatarsal portion of the subject's shoe as shown in Fig. 1(b). The IMU is used to collect acceleration and gyroscope values in the three axes and quaternions. A force resistive sensor (FSR) of dimension $2 \times 3 \times 0.2$ in., shown in Fig. 1(c), is placed under the heel of the foot. The output of the FSR reaches a maximum value of 1024 when a large force is applied. The FSR is used to give the ground truth for the heel strike event [10]. The data obtained from the FSR and IMU are sampled at 100 Hz [14] which is transferred via a UDP packet as shown in Fig. 2. This data is stored

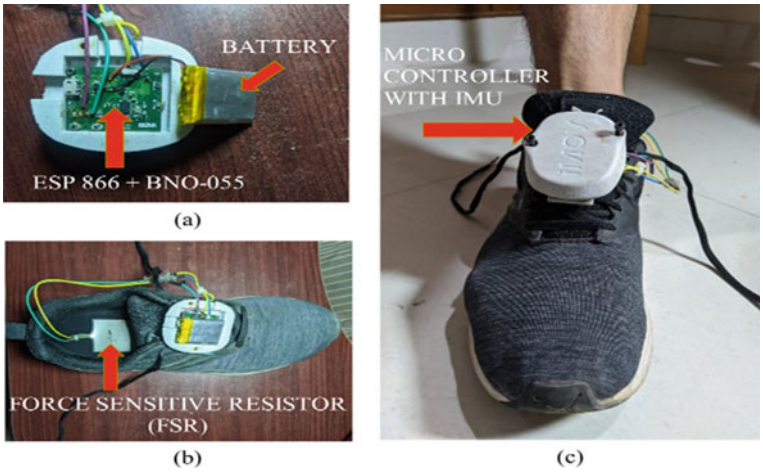



**Fig. 1** Hardware Implementation **a** Microcontroller interfaced with IMU **b** FSR placed at heel of the shoe **c** Setup placed at metatarsal position of shoe

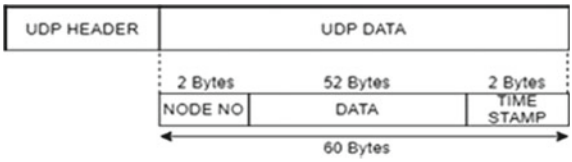



**Fig. 2** UDP frame format

in a tabular form that serves as input to the training algorithms. The data is then used to train a neural network model that is loaded onto the microcontroller giving a real-time portable model for the prediction of heel strike.

### *3.2 Experimental Protocol*

In this study, for the collection of data to train the neural network, 8 healthy participants (6 males and 2 females, age 21 ± 1 years, height 1.71 ± 0.51 m, weight 78.67 ± 10.8 kg) were considered with no impairment to their gait. Informed consent has been obtained from all subjects before the day of data collection.

The subjects were made to walk on a treadmill with the setup mentioned in the hardware, for a total period of 8 min and 20 s, giving approximately 350 strides [15], which forms a dataset of 50,000 samples. The data was collected and transmitted over a Wi-Fi hotspot connection to the PC where it was stored. Before each trial, proper placement of the FSR, under the heel, was ensured. The sensor could be placed in any orientation on the foot as any offset angles were corrected by global transformations; described in Sect. 4.1.

The data collection process would be stopped, after the specified period, after which the participant would stop the treadmill, ending the experiment.

## 4 Methods

### *4.1 Global Transformations*

IMU is sensitive to minute changes in the mounting position of the setup on the subject and therefore it is necessary to eliminate these errors before performing further computations. Global transformations convert the angles received by the sensor from a local frame to a global frame. This results in a constant frame of reference for the data collected [16] and compensates for any orientation errors in the placement of the sensor on the foot.

To achieve this, 100 samples are taken to be stationary so that a frame can be defined. Then it is converted as follows:

1. Let each sample be represented by a quaternion Q defined in Eq. 1.

$$Q = \left[ Q_w \, Q_x \, Q_y \, Q_z \right] \tag{1}$$

   where $Q_w$ is the magnitude of rotation about a certain axis and $Q_x$, $Q_y$ and $Q_z$ are the vector components of rotation about the respective axes.
2. The average of the first 100 quaternion values to get $Q_{avg}$ is explained in Eq. 2.

$$Q_{avg} = \begin{bmatrix} Q_{w_{avg}} \\ Q_{x_{avg}} \\ Q_{y_{avg}} \\ Q_{z_{avg}} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{100} Q_{w_i}/100 \\ \sum_{i=1}^{100} Q_{x_i}/100 \\ \sum_{i=1}^{100} Q_{y_i}/100 \\ \sum_{i=1}^{100} Q_{z_i}/100 \end{bmatrix} \tag{2}$$

3. Calculate the inverse of $Q_{avg}$ to get $Q_{body}^{world}$ using the formula is expressed in Eq. 3.

$$Q_{body}^{world} = \begin{bmatrix} Q_{w_{body}^{world}} \\ Q_{x_{body}^{world}} \\ Q_{y_{body}^{world}} \\ Q_{z_{body}^{world}} \end{bmatrix} = \frac{Q_{avg}}{Q_{w\,avg}^2 + Q_{x\,avg}^2 + Q_{y\,avg}^2 + Q_{z\,avg}^2} \tag{3}$$

4. Then for the subsequent quaternion values $Q^{body}$, multiply the quaternions with $Q_{body}^{world}$ to get $Q^{world}$ expressed in Eq. 4.

$$Q^{world} = \begin{bmatrix} Q_{w^{world}} \\ Q_{x^{world}} \\ Q_{y^{world}} \\ Q_{z^{world}} \end{bmatrix} = \begin{bmatrix} Q_{w^{body}} & -Q_{x^{body}} & -Q_{y^{body}} & -Q_{z^{body}} \\ Q_{x^{body}} & Q_{w^{body}} & Q_{z^{body}} & -Q_{y^{body}} \\ Q_{y^{body}} & -Q_{z^{body}} & Q_{w^{body}} & Q_{x^{body}} \\ Q_{z^{body}} & Q_{y^{body}} & -Q_{x^{body}} & Q_{w^{body}} \end{bmatrix} * \begin{bmatrix} Q_{w_{body}^{world}} \\ Q_{x_{body}^{world}} \\ Q_{y_{body}^{world}} \\ Q_{z_{body}^{world}} \end{bmatrix} \tag{4}$$

## 4.2 Neural Network Architecture

For gait parameter prediction, a supervised learning approach was employed. The neural network model shown in Fig. 3 utilizes a feedforward backpropagation algorithm. It consists of 3 layers, i.e., the input layer, hidden layer, and output layer. A

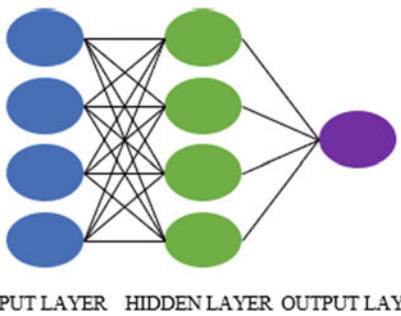



**Fig. 3** Neural network architecture

preliminary analysis was conducted to determine the appropriate number of nodes for the hidden layer. The input and hidden layer consist of four nodes each while the output layer is a single node.

The quaternion values are normalized is expressed in Eq. 5:

$$Q_{normalize} = \frac{Q - W_{avg}}{W_{sd}} \tag{5}$$

where $W_{avg}$ and $W_{sd}$ are $1 \times 4$ vectors of the average and standard deviation of all quaternion values in the dataset, respectively. $Q_{normalize}$ is the input to the neural network and the model is trained. The output of the model is a prediction of the probability of the heel strike between 0 and 1.

Let $x_t$ be the input quaternion vector of size 4 to the input layer, $W_1$ be the $4 \times 4$ weight matrix between input and hidden layer and $b_1$ be $1 \times 4$ bias vector. The Rectified Linear Unit (ReLu) activation function is used for the hidden layer. This is shown in Eq. (6), (7) and (8).

$$a_1 = x_t \tag{6}$$

$$z_2 = W_1 * a_1 + b_1 \tag{7}$$

$$a_2 = max(0, z_2) \tag{8}$$

where $a_1$ is the input layer consisting of four nodes, $z_2$ is the hidden layer nodes before the data passes through the activation function and $a_2$ is the hidden layer nodes after the activation function is applied.

Let $W_2$ be the $4 \times 1$ weight matrix between the hidden and output layer and $b_2$ be its respective $1 \times 1$ bias. The sigmoid activation function is used for the output layer. This is given in Eq. (9) and Eq. (10).

$$z_3 = W_2 * a_2 + b_2 \tag{9}$$

$$a_3 = \sigma(z_3) \tag{10}$$

where $z_3$ is the output layer nodes before the data passes through the activation function and $a_3$ is the output layer nodes after the activation function is applied.

The binary cross-entropy loss function is used to minimize the error where Y is the truth of heel strike and $Y_{pred}$ is the prediction of the model which is expressed in Eq. 11.

$$Loss = Y * -(logY_{pred}) + (1 - Y) * -(log(1 - Y_{pred})) \tag{11}$$

The network is optimized using the Adam version of stochastic gradient descent.

### 4.3 Real-Time Neural Network

The neural network model, mentioned above, was loaded onto the microcontroller. The trained weights of the neural network were sent from the PC to the microcontroller wirelessly before the start of each experiment. This provided a real-time portable gait analysis/heel strike analysis model that could predict the gait cycle/heel strike of an individual.

In the real-time implementation, the following steps were performed in the microcontroller. First, the weight matrices of the trained neural network $W_1, b_1, W_2\, and\, b_2$, shown in (7–10), and the normalization matrices $W_{avg}$ and $W_{sd}$, shown in (5), were transferred over Wi-Fi from the computer script to the microcontroller. Then, 100 samples were collected when the body is in a static position and global transformation was carried out, described in Sect. 4.1, where only role and pitch Euler angles were used to represent the quaternions as shown in Eq. (12) below.

$$(Roll_i, Pitch_i, Yaw_i = 0) = (Q_{w_i}, Q_{xi}, Q_{y_i}, Q_{z_i}) \tag{12}$$

Next, each input quaternion $Q_i$ is converted to the global frame as described in Sect. 4.1 to obtain $Q_i^{world}$. This is then normalized to obtain $Q_{i\,normalize}^{world}$ before passing it to the neural network which separate equations are expressed in Eq. 13, 14, 15, 16, 17, and 18.

$$Q_{i\,normalize}^{world} = \frac{Q_i^{world} - W_{avg}}{W_{sd}} \tag{13}$$

In the neural network, the following computations were performed.

$$a_1 = Q_{i\,normalize}^{world} \tag{14}$$

$$z_2 = W_1 * a_1 + b_1 \tag{15}$$

$$a_2 = max(0, z_2) \tag{16}$$

$$z_3 = W_2 * a_2 + b_2 \tag{17}$$

$$a_3 = \sigma(z_3) \tag{18}$$

where $a_3$ is the probability of the heel strike event for the particular input sample.

### 4.4 Event Detection Algorithm

The threshold value for a heel-strike event (HSE) was set at the average of the highest and lowest value given by the FSR. The heel-strike event was considered if the given sample was above the threshold provided the previous sample was below the threshold value is expressed as Eq. 19.

$$HSE(t) = \begin{cases} 1, sample_{t-1} \langle threshold \, and \, sample_t \rangle \, threshold \\ 0, \; otherwise \end{cases} \tag{19}$$

## 5 Results

### 5.1 Importance of the Various Neural Network Models

To enable the activation of FES for the self-rehabilitation of a patient, a specific neural network (SNN) model was developed which is trained with data of a single subject, who would ideally be the patient.

To test the hypothesis of whether a neural network model could generalize the gait cycle of a large population, so that repeated data collection of every patient is avoided, a generalized neural network model (GNN) was developed. The GNN-X model is trained with X number of subjects.

### 5.2 Specific Neural Network Model

A specific neural network (SNN) model was trained on 5 datasets of an individual, who will be addressed as the main subject henceforth. The weights of the SNN were transferred to the microcontroller for real-time implementation, after which the main subject was made to walk for a final trial. The SNN was then retrained with the dataset obtained from the final trial as part of the offline analysis. The probability of the heel strike event obtained from the offline analysis is compared with the predicted output of the real-time model or online analysis as shown in Fig. 4(a) and the time difference between the predicted output and the true heel strike for both the analysis is shown in Fig. 4(b). The negative time difference indicates that the prediction was made before the true heel strike which can be used to compensate for time latencies in activating electrical stimulation.

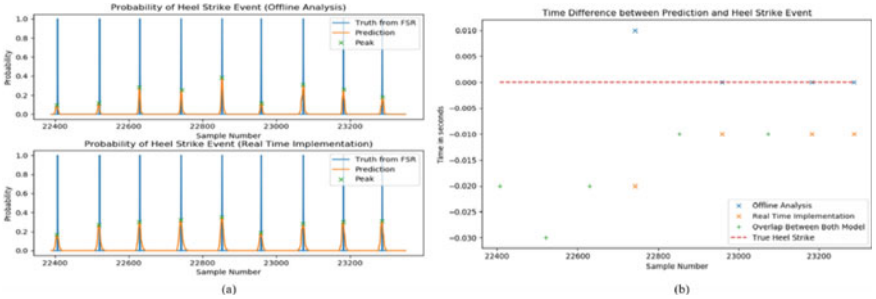



**Fig. 4** **a** Probability of heel strike event from offline analysis (Top) and real time implementation (Bottom) for SNN, **b** Time difference between predicted output and heel strike event for SNN

## *5.3 General Neural Network Model*

For GNN-2 model, the SNN was retrained with 5 datasets collected from a second subject and for GNN-8 model, the SNN was retrained with 5 datasets each of 7 subjects. After the models were retrained, the GNN-X weights were transferred to the microcontroller for real-time implementation after which the main subject was made to walk for a final trial. The GNN-X was then retrained with the dataset obtained from the final trial as part of the offline analysis. The probability of the heel strike event obtained from the offline analysis is compared with the predicted output of the real-time model for GNN-8 as shown in Fig. 5(a) and the time difference between the predicted output and the true heel strike for GNN-8 is shown in Fig. 5(b).

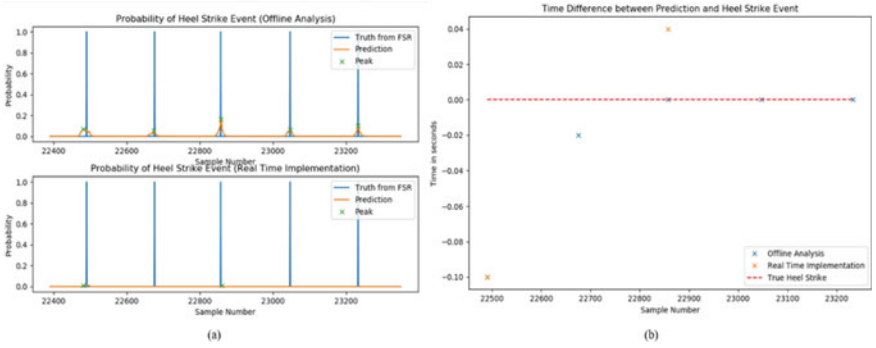



**Fig. 5** **a** Probability of heel strike event from offline analysis (Top) and real time implementation (Bottom) for GNN-8, **b** Time difference between predicted output and heel strike event for GNN-8

## 6 Discussion

A comparative analysis of the offline analysis and real-time implementation of all the models is shown in Table 1. For SNN, it is observed that the probability of HSE in real-time implementation is greater but almost equal to the offline analysis. This can be attributed to precision errors in the data from the real-time implementation when compared to the offline analysis and randomness in convergent states of the SNN.

For GNN-2, the real-time implementation has a significantly low probability of heel strike event when compared to the offline analysis and the real-time implementation of SNN and has failed to predict the HSE in certain instances. The average and standard deviation in time difference between prediction and true heel strike are also higher when compared to SNN as shown in Table 1. This presents the initial evidence that it is not possible to create a neural network model that can generalize the gait cycle of several individuals.

In GNN-8, the real-time implementation has failed to predict most heel strike events, and even if there was a prediction, the average probability of occurrence was close to zero. The probability of a heel strike event from the offline analysis was significantly higher. The results obtained from the GNN-8 model provides substantial evidence to disregard the hypothesis that an artificial neural network model cannot generalize the gait cycle of a large population.

Table 2 compares the accuracies of all the three models. It is seen that the accuracy of the SNN is the highest in the real time implementation and decreases with

**Table 1** Comparison between offline analysis and real time implementation for all models

| Statistical parameters | Offline analysis | | | Real-time implementation | | |
|---|---|---|---|---|---|---|
| | SNN | GNN-2 | GNN-8 | SNN | GNN-2 | GNN-8 |
| Average probability of HSE | 0.1924 ± 0.1085 | 0.1220 ± 0.0923 | 0.0789 ± 0.0430 | 0.2164 ± 0.0818 | 0.0268 ± 0.0283 | 0.0111 ± 0.0042 |
| Average time difference between prediction and true heel strike (in seconds) | –0.0027 ± 0.0153 | 0.0038 ± 0.0348 | –0.0289 ± 0.0357 | –0.0150 ± 0.0086 | –0.0462 ± 0.0282 | –0.0115 ± 0.0474 |

**Table 2** Comparison of accuracies between models

| Model | Total number of heel strike events | Offline analysis | Real-time implementation |
|---|---|---|---|
| SNN | 361 | 361 | 361 |
| GNN-2 | 343 | 343 | 325 |
| GNN-8 | 367 | 367 | 94 |

increasing number of subjects. This again provides more evidence that the gait cycle of a large population cannot be generalized using artificial neural networks.

## 7 Conclusion

From the statistical analysis of the prediction models carried out in the paper, we realize that the SNN had better accuracy and lower time differences between predictions and true heel strikes when compared to GNNs. Therefore, the highly accurate prediction of heel strike event made by the specific neural network model on the embedded system can be used to trigger electrical stimulation for real-time interventional gait control which will help in the automation of rehabilitation of patients with movement disorders. We also conclude that each person has a unique gait pattern, and the prediction model generalization is not practical.

## References

1. Baker R (2006) Gait analysis methods in rehabilitation. J Neuroeng Rehabil 3(1):4
2. Sale P, Franceschini M, Waldner A, Hesse S (2012) Use of the robot assisted gait therapy in rehabilitation of patients with stroke and spinal cord injury. Eur J Phys Rehabil Med 48(1):111–121
3. Shahzad A, Ko S, Lee S, Lee JA, Kim K (2017) Quantitative assessment of balance impairment for fall-risk estimation using wearable triaxial accelerometer. IEEE Sens J 17(20):6743–6751
4. Prakash C, Kumar R, Mittal N (2018) Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. Artif Intell Rev 49(1):1–40
5. Sant'Anna A, Wickström N (2010) A symbol-based approach to gait analysis from acceleration signals: identification and detection of gait events and a new measure of gait symmetry. IEEE Trans Inf Technol Biomed 14(5):1180–1187
6. Raghavendra P, Sachin M, Srinivas PS, Talasila V (2017) Design and development of a real-time, Low-Cost IMU based human motion capture system. In: Vishwakarma H, Akashe S (eds.) Computing and Network Sustainability. LNNS, vol 12, pp 155–165. Springer, Singapore. https://doi.org/10.1007/978-981-10-3935-5_17
7. Raghavendra P, Talasila V, Sridhar V, Debur R (2017) Triggering a functional electrical stimulator based on gesture for stroke-induced movement disorder. In: Vishwakarma H, Akashe S (eds.) Computing and Network Sustainability. LNNS, vol 12, pp 61–71. Springer, Singapore. https://doi.org/10.1007/978-981-10-3935-5_7
8. Mijailovic N, Gavrilovic M, Rafajlovic S, Đuric-Jovicic M, Popovic D (2009) Gait phases recognition from accelerations and ground reaction forces: application of neural networks. Telfor J 1(1):34–36
9. Gujarathi T, Bhole K (2019) Gait analysis using IMU sensor. In: 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp 1–5. IEEE, July 2019
10. Zhang H, Guo Y, Zanotto D (2019) Accurate ambulatory gait analysis in walking and running using machine learning models. IEEE Trans Neural Syst Rehabil Eng 28(1):191–202
11. Hori K et al (2020) Inertial measurement unit-based estimation of foot trajectory for clinical gait analysis. Front Physiol 10:1530

12. Kidziński Ł, Delp S, Schwartz M (2019) Automatic real-time gait event detection in children using deep neural networks. PLoS ONE 14(1):e0211466
13. Schicketmueller A, Rose G, Hofmann M (2019) Feasibility of a sensor-based gait event detection algorithm for triggering functional electrical stimulation during robot-assisted gait training. Sensors 19(21):4804
14. Li X, Xu H, Cheung JT (2016) Gait-force model and inertial measurement unit-based measurements: a new approach for gait analysis and balance monitoring. J Exerc Sci Fit 14(2):60–66
15. Meyer C et al (2019) Familiarization with treadmill walking: how much is enough? Sci Rep 9(1):1–10
16. Parthasarathy A, Megharjun VN, Talasila V (2020) Forecasting a gait cycle parameter region to enable optimal FES triggering. IFAC-PapersOnLine 53(1):232–239

# Role of Routing Techniques in Wireless Sensor Networks – A Survey


**Md. Shahid Thekiya and Mangesh D. Nikose**

**Abstract** Wireless sensor network (WSN) plays a significant role in various commercial, industrial and agriculture sector in wide ranges of applications. Routing is necessary to maintain the reliable communication between different nodes, cluster heads, and base station; however, performance of routing mechanism is challenging due to network lifetime, dynamic nodes, higher packet drop, scalability issue, limited adaptability and environmental conditions. This paper presents various routing techniques used in WSN that covers the brief overview of flat routing, hierarchical routing, location-based routing and bio-inspired routing techniques such as Ant Colony Optimization (ACO) and Artificial Bee Colony Optimization (ABC) algorithms. It focuses on the routing mechanism, routing scenario, mobility, scalability, energy consumption, data aggregation, performance evaluation metrics, challenges and constraints of the routing in WSN. This comprehensive survey provides the future direction for the improvement in routing mechanisms in WSN.

**Keywords** Bio-inspired routing · Clustering · Flat routing · Hierarchical routing · Routing · Wireless sensor network

## 1 Introduction

WSN is a collection of sensor nodes dispersed over the plane to monitor the physical and environmental parameters. WSNs are generally used for data processing, analysis, mining, and storage of data [1]. Due to the vast industrial revolution, WSNs plays vital role in industrial automation, Internet of Things (IoT), environmental monitoring, weather forecasting, landslide detection, flood detection, tsunami alert system, earthquake detection, transportation, military applications, wildlife monitoring, logistics etc. [2]. WSNs are categorized in to structured and unstructured

Md. S. Thekiya (✉) · M. D. Nikose
Sandip University, Nashik, India
e-mail: shahidthekiya@gmail.com

M. D. Nikose
e-mail: mangesh.nikose@sandipuniversity.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_70

WSN based on placement of sensors over the plane. In the structured WSNs, the sensor nodes are installed in the structured way over the plane. Whereas, the sensors are randomly placed in the unstructured WSNs. Depending upon the applications the WSNs are categorized in to terrestrial, underground, underwater and mobile WSN [3]. Sensor node of WSN consists of sensor or transducer that senses the physical or environmental action and converts it to electrical signal, processor to process the signal, communication model that encompasses transmitter and receiver, and battery as the power source. Routing is the process of the establishment of the path from the sensor node to other sensor node or cluster head (CH) or base station (BS) for the transmission and reception of the information. Because of the short life of the battery, energy efficient routing techniques are essential. Routing in WSN is an important aspect of data transmission. Routing techniques are broadly grouped into the network structure and protocol operation based routing. Flat, hierarchical, and location based routing protocol are subtypes of structure based routing techniques whereas protocol operation based routing techniques are categorized into negotiation, multipath, QoS, query, and coherent based routing protocol [4, 5].

This paper presents the detailed survey of the various routing techniques in WSNs. It provides brief about the major routing techniques covering flat routing, hierarchical routing, location-based routing and bio-inspired routing techniques. It focuses on the various issues of the routing techniques such as throughput, energy efficiency, scalability, mobility, data aggregation, etc.

The remaining paper is structured as follow: Sect. 2 provides the detailed survey of traditional flat, hierarchical, location-based, and operation-based techniques. Section 3 gives details survey of two popular bio-inspired routing techniques such as ant colony optimization and artificial bee colony optimization. Section 4 gives the discussion on various routing mechanisms. Finally, Sect. 5 concludes the paper and provides the challenges and constraints of the current routing mechanisms which open the paths for future work in the same domain.

## 2 Review on WSN Routing

Extensive work has been carried in the past on the different WSN issues like routing, clustering, data aggregation, security, etc. Different sorts of algorithms ranging from statistical, machine learning based algorithms to bio-inspired algorithms have been presented by researchers. The categorization of the WSN routing protocol is shown in Fig. 1.

### *2.1 Structure Based Routing*

Flat routing is also called a data centric routing approach in which each node plays a vital role to sense the data and send it to another node. Because of the large number

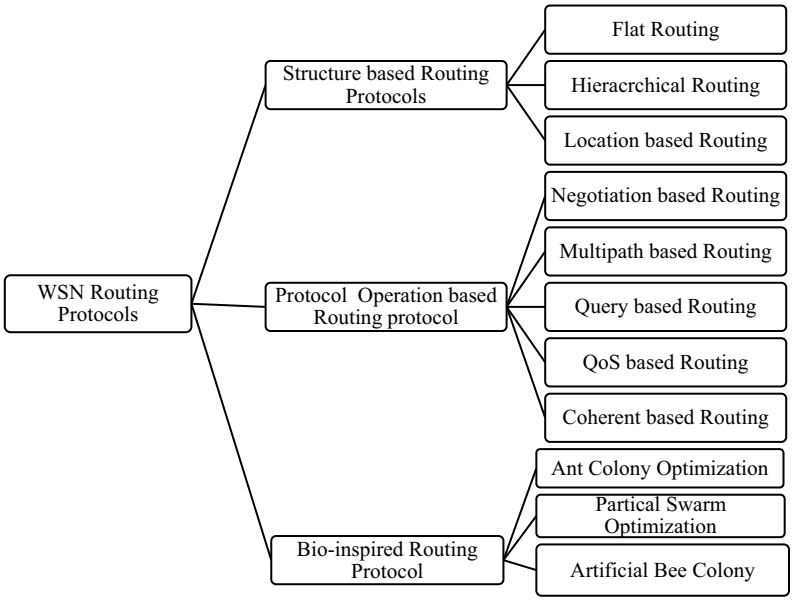



**Fig. 1** Classification of WSN routing techniques

of sensor nodules, it is impossible to assign a universal identifier to every node. Hussein [6] have presented Sensor Protocols for Information via Negotiation (SPIN) routing protocol in which each sensor node is considered as the BS for the data transmission. It is a multipath routing technique and can be used for mobility and data aggregation. It has limited scalability and more power consumption. Further, Yu et al. [7] anticipated directed diffusion (DD) that combines the data coming from all the sensor nodes and removes the data redundancy and minimizes transmissions overheads to save the energy for prolonged network lifespan.

Subsequently, Rumor routing [8] is a variety of direct diffusion and is principally planned for applications where geographic steering isn't doable. Further, the Minimum Cost Forwarding Algorithm (MCFA) [9] is used to maintain the routing table for storage of sensor ID to decide the direction of routing towards the BS. In MCFA data is transmitted to the nearest sensor node till the BS is reached. Further, in [10] have presented another kind of directed diffusion as the Gradient based routing (GBR) based on total hops and height of node. It forwards the packets to the node having a larger gradient. Later on Constrained anisotropic diffusion routing (CADR) been analyzed in [11]. In CADR, each node estimates cost objective and routes data using the local cost gradient and user requirement. In IDSQ, the query node can estimate that which node can transmit the needful data by balancing the energy cost function. Another flat routing algorithm such as COUGAR [12], Energy aware routing [13] [14] and Routing Protocols with Random Walks [15] have been applied to different routing problems in WSN. Flat algorithms are simple to implements and energy consumption is moderate which depends upon the traffic pattern. Flat

routing has several limitations like larger contentions, more collision overheads, less synchronization and more latency.

To get the better scalability and spatial information of network hierarchical routing techniques which are also known as cluster based routing techniques are used. The review of previous work presents on hierarchical techniques used for WSN routing. Salem et al. [16] has proposed the popular cluster based routing technique named as Low Energy Adaptive Clustering Hierarchy (LEACH). LEACH randomly selects the sensor node as the CHs and rotates it till the even distribution of energy in all the nodes. It uses TDMA/CDMA to get the inter and intra-cluster collision. It results in better network lifetime, lower energy consumption and optimized path. Power-Efficient Gathering in Sensor Information Systems (PEGASIS) is an improved version of LEACH [17]. It is an optimal chain based protocol. This sensor communicates with the nearest node only and forms the routing path towards the BS. It has local coordination between the sensor nodes for power optimization and longer network life. LEACH forms the clusters whereas PEGASIS avoids forming the clusters and uses a single node from the routing chain to transmit the information. Further, Threshold-sensitive Energy Efficient Sensor Network protocol (TEEN) [18] and Adaptive Periodic Threshold-sensitive Energy Efficient Sensor Network protocol (APTEEN) [19] protocols have been used for timing constraint applications. In TEEN sensor nodules continuously senses the channel, when the sensor is triggered by threshold it starts the communication. The major cons of TEEN technique are that if the threshold is not received then the node remains in idle mode. APTEEN is a modified version of TEEN which uses both proactive and reactive types of routing. APTEEN deals with attributes to be sensed, the threshold to be sensed, count time of transmission and schedule to assign the slot for communication to each sensor node. Subsequently, Small Minimum Energy Communication Network (MECN) [20] protocol which uses small and low power GPS for the sub-clustering network has been proposed. It finds the sub-network consists of limited sensor nodes and then transmits the data so that the energy required for the transmission will be lesser. It results in more overheads in the sensor network. Afterward, in [21] have presented the self-organizing tree based protocol (SOP) to support the heterogeneous mobile or stationary wireless sensor network. It consists of designed router nodes to which sensor nodes transmit the data. Router nodes then communicate the data with the BS. Energy efficient Virtual Grid Architecture routing protocol (VGA) [22] has been presented to maximize the network lifetime. It is GPS free approach that forms the smaller equal, adjacent and non-overlapping clusters. A hierarchical power-aware routing (HPAR) was presented in [23], which divides the sensor network in a small group of the sensor based on geographical proximity. It then transmits the data to the other nearest small zone to optimize the power by considering the path which has maximum remaining power which is also called as maximum power minimum distance path. Thus, cluster based or hierarchical routing techniques shown robust performance because of the absence of the collision reduced duty cycle and lower latency. It consists of a uniform but uncontrolled energy distribution. Routing is cluster based, simple but not optimal. Data aggregation is done by CHs in cluster based routing techniques.

In location based routing techniques, the location of the sensor node is considered for the routing of information obtained from a global positioning system (GPS). It considers the location of the node module for path optimization and energy management. The different location based routing techniques used in various types of WSN architectures are discussed below. An energy aware location based algorithm Geographic Adaptive Fidelity (GAF) [24] is proposed for the mobile ad-hoc sensor network. The network is divided into fixed sensor zones to form the grid. The sensor network uses GPS for its location. Later on, Geographic and Energy Aware Routing (GEAR) [25] have been proposed by Yu et al. which uses geographic data to respond to the queries. The data is transmitted to the destination node based on the nearest neighbor information obtained from GPS. Location based algorithms have been proved more efficient for the large mobile, heterogeneous and distributed network because of the ability to handle the location of sensor nodes. Received signal strength indicator (RSSI) algorithms are radio frequency based algorithm which finds the distance between the transmitter sensor node and receiver sensor node. RSSI algorithms are better suitable in WSN because of inexpensive, energy effective and simple hardware structure. In a large network, it needs larger collections of data samples [26]. RSSI has given better performance in the outdoor environment than the indoor environment. This increase in the anchor node reduces the localization error [27].

## *2.2 Protocol Operation Based Routing*

These types of routing protocols are based on the protocol operation that is being used for the routing. In a multipath routing protocol, multiple paths are used for data transmission rather than the single path for improvement in the network performance. It gives the alternate path when any path fails in data transmission. The switching cost is more in multipath routing algorithms [28]. Query based routing protocol is another protocol operation based routing protocol. Rumor routing protocol [8] is a query based routing protocol in which data is transmitted to the neighboring node present in the routing table depending upon the query by another node. SPIN algorithms [6] are the negotiation based routing protocol which eliminates the data redundancy based on negotiation. QoS based Sequential Assignment Routing (SAR) [29] has been used for improving the quality of network such as energy, delay, bandwidth, etc. In a single winner algorithm (SWE) and multiple winner algorithm (MWE) which are non-coherent and coherent routing protocols such as routing protocols generally used for the energy efficient routing but has more processing cost. All the above algorithms can be fused to develop the hybrid protocols to combine the advantages of different algorithms to tradeoff between various network parameters. The different routing techniques are summarized in the following Table 1 according to the different metrics.

**Table 1** Comparative analysis of different WSN routing techniques

| Method | Type | Mobility | Scalability | Energy usage | Data aggregation | Multi-path | Localization | Query based |
|---|---|---|---|---|---|---|---|---|
| SPIN [6] | Flat | Possible | Low | Limited | Yes | Yes | No | Yes |
| DD [7] | Flat | Limited | Low | Limited | Yes | Yes | Yes | Yes |
| Rumor [8] | Flat | Limited | Good | Limited | Yes | No | No | Yes |
| GBR [10] | Flat | Limited | Low | Limited | Yes | No | No | Yes |
| MCFA [9] | Flat | No | Low | Limited | No | No | No | No |
| CADR [11] | Flat | No | Low | Limited | Yes | No | No | No |
| COUGAR [12] | Flat | No | Low | Limited | Yes | No | No | Yes |
| EAR [14] | Flat | Limited | Low | Limited | Yes | No | No | Yes |
| LEACH [16] | Hierarchical | Yes | Good | Maximum | Yes | No | Yes | No |
| TEEN & APTEEN [18][18] | Hierarchical | Yes | Good | Maximum | Yes | No | Yes | No |
| PEGASIS [17] | Hierarchical | Yes | Good | Maximum | No | No | Yes | No |
| MECN [20] | Hierarchical | No | Low | Maximum | No | No | No | No |
| SOP [21] | Hierarchical | No | Low | Limited | Yes | No | No | No |
| HPAR [23] | Hierarchical | No | Low | Limited | Yes | No | No | No |
| VGA [22] | Hierarchical | No | Low | Limited | Yes | Yes | Yes | No |
| GAF[24] | Location | Limited | Low | Limited | No | No | No | No |
| GEAR [25] | Location | Limited | Low | Limited | No | No | No | No |
| RSSI [27] | Location | Yes | Yes | Moderate | Yes | No | Yes | No |

## 3 Bio-inspired Routing

Bio-inspired algorithms are based on the biological phenomenon which is implemented in the form of computational algorithms. These algorithms can deal with multi objective function problem solving, the capability to model and solve a complex problem [30]. This section focuses on work carried out on two popular bio-inspired swarm based algorithms such as ABC and ACO for the different wireless sensor network clustering and routing problems.

### *3.1 Ant Colony Optimization Algorithm*

ACO algorithm is a multi-objective optimization algorithm that is used for WSN routing, clustering and data aggregation. It is the biological population based algorithm based on the pheromone trail laying behavior of real ant colonies. Communication between various ants is based on a chemical called pheromone. Ants find the shortest path to the food source based on pheromone secreted by the ants. They follow the pheromone secreted path towards food. If any ant gets lost during the food searching, then it senses the pheromone and follows the path. Pheromone may get

evaporated due to environmental conditions; therefore, ants update the path periodically to increase the concentration of the pheromone level [31]. This algorithm is a popular algorithm for WSN routing problem. This section provides detailed information about previous work carried out on the ACO algorithm in the area of WSN. Supreet Kaur et al. [32] have used clustered based ACO to find the shortest path in the sensor node and sink node. It provided better network lifetime, stability period and throughput. Compressive sensing provided better compression of data to be transmitted. Jamal et al. [33] covered the generalized evolution criteria and performance analysis of ACO concerning to WSN. It considered time synchronization, data security and aggregation, sensor node localization, data packet routing in the network and load balancing. Along with numerous advantages of ACO algorithms described previously it has some major limitations also. The value of pheromone depends on the scale of the problem or area and type of network used. Due to a random decision in problem solving, the theoretical analysis of the algorithm is quite difficult. Source nodes spent most larger time for route searching and maintenance because of dynamic topology. The population of ants is an important factor of the algorithm. Larger population may result early convergence to poor solution. The pheromone decay rate needs to be carefully controlled as lower number of ants may result in cooperative behavior [34].

### *3.2 Artificial Bee Colony Algorithm*

ABC is a bio-inspired algorithm which is based upon the food searching phenomenon of honey bees. Bee colony has three types of bees such as employee bees, onlooker bees and scout bees. Employee bees searches food sources within the vicinity of the food source in their memory and they share about food information with onlooker bees and then the onlooker bees choose the best food sources. Onlooker bees chooses a food source within the vicinity of the food sources by themselves. Scout bees hunt a new food source arbitrarily when an employed bee has been deserted a food source. It is based on the divide and conquers strategy. It can handle complex function and multiple objective functions. ABC algorithm has been widely used for WSN clustering and routing problem solving. This section focuses on several previous works that have been carried out on ABC algorithms for solving WSN routing and clustering problems [35]. In recent year, ABC has been successfully presented for the various routing scenarios. Besides the number of advantages of ABC algorithms, it has several challenges like it does not consider the location in the routing process in WSN [36, 37]. If the structure of the network is not fixed, excessive energy consumption and convergence time for optimization are more.

## 4 Discussions

The performance of some of the routing techniques degrades in the heterogeneous and mobile WSNs. A major problem associated with the mobile WSN is power management as sensor nodes require extra power for mobility. Current routing algorithms are mainly developed for application specific nature of the network and limited capabilities of sensor nodes. Very less concentration has been given on the security of the data. Various methods have been studied on WSN clustering which majorly focused on energy optimization and cluster optimization. Clustering is performed based on the location of the sensor node and the power of sensor nodes. The performance of the clustering algorithm has been studied based on mobility, scalability, self-organization, distribution, resource awareness, data aggregation, homogeneity and randomized rotation nature. The tradeoff between different parameters is difficult while optimizing the network parameters. Some of the clustering techniques are facing challenges in scalability, energy consumption, mobility and network overheads. To handle the multiple objective functions and complex problem solving, bio-inspired algorithms are used extensively. Various swarm based algorithms such as ACO, ABC, etc. have been studied for the WSN routing problem solving. These techniques can solve the complex problem, the ability to optimize multiple functions at a time and enhance self-performance iteratively. These techniques resulted in better power and routing optimization. Swarm based algorithms are time consuming when the structure of the network is extremely larger because of its iterative nature. Swarm based routing algorithms focus on the one to one data transmission in the sensor network. These algorithms are dependent on the random decisions which are quite difficult to express theoretically or mathematically.

## 5 Conclusion

This paper summarizes the analysis of the various WSN routing techniques in detail. The performance of these routing techniques is measured based on mobility, power optimization, route optimization, location awareness, localization, scalability, network lifetime, multipath structure, network complexity, network overheads and security, etc. Forcing to optimize one parameter may disturb the optimization of other parameters. These algorithms are dependent upon the network structure, size of the network, type of sensor nodes, platform and environment of sensor deployment, mobility of network, homogeneous and heterogeneous structure of the network. These routing algorithms may fail to optimize more than one parameter at a time efficiently. In future, more focus can be given on the tradeoff between various network parameters to improve the network quality of service.

## References


1. Zagrouba R, Kardi A (2021) Comparative study of energy efficient routing techniques in wireless sensor networks. Information 12(1):42
2. Fahmy HMA (2021) WSN Applications. In: Concepts, Applications, Experimentation and Analysis of Wireless Sensor Networks, pp 67–232. Springer, Cham. https://doi.org/10.1007/978-3-030-58015-5_3
3. Devika G, Karegowda AG, Ramesh D (2020) Survey of WSN routing protocols. Int J Appl Evolut Comput (IJAEC) 11(1):34–51
4. Chan L, Chavez KG, Rudolph H, Hourani A (2020) Hierarchical routing protocols for wireless sensor network: a compressive survey. Wirel Netw 26(5):3291–3314
5. Shafiq M, Ashraf H, Ullah A, Tahira S (2020) Systematic literature review on energy efficient routing schemes in WSN–a survey. Mob Netw Appl 1–14
6. Hussein SA (2020) A new wireless sensor networks routing algorithm based on SPIN protocols and circumference technique 6(9):21–27
7. Yu X, Li F, Li T, et al (2022) Trust-based secure directed diffusion routing protocol in WSN. J Ambient Intell Human Comput 13:1405–1417. https://doi.org/10.1007/s12652-020-02638-z
8. Kheroua L, Moussaoui S, el houda Baroud N, Chaib O (2018) An optimized relative coordinates rumor routing (ORCRR) for wireless sensor networks. In: 2018 International Conference on Smart Communications in Network Technologies (SaCoNeT), pp 312–316, October 2018. IEEE
9. Ramos AR, Velez FJ, Gardašević G (2020) Performance evaluation of source routing minimum cost forwarding protocol over 6TiSCH applied to the OpenMote-B platform. In: José R, Van Laerhoven K, Rodrigues H (eds.) 3rd EAI International Conference on IoT in Urban Space, p 123. Urb-IoT 2018. EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-28925-6_11
10. Migabo ME, Olwal TO, Djouani K, Kurien AM (2017) Cooperative and adaptive network coding for gradient based routing in wireless sensor networks with multiple sinks. J Comput Netw Commun 2017:10 p. Article ID 5301462. https://doi.org/10.1155/2017/5301462
11. Moundounga ARA, Satori H, Satori K (2020) An overview of routing techniques in WSNs. In: 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), pp 1–7. IEEE
12. Behera TM, Samal UC, Mohapatra, SK (2019) Routing Protocols. In: Computational Intelligence in Sensor Networks, pp 79–99. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-319-78262-1_300574
13. Li Z, Shan D (2020) Research on 3D wireless sensor networks ISC-EAR routing algorithm. Adv Netw 8(2):16
14. Jurado-Lasso FF, Clarke K, Cadavid AN, Nirmalathas A (2021) Energy-aware routing for software-defined multihop wireless sensor networks. IEEE Sens J 21(8):10174–10182
15. Nguyen MT, Teague KA (2017) Compressive sensing based random walk routing in wireless sensor networks. Ad Hoc Netw 54:99–110 (2017)
16. Salem AOA, Shudifat N (2019) Enhanced LEACH protocol for increasing a lifetime of WSNs. Pers Ubiquit Comput 23(5):901–907
17. Khedr AM, Aziz A, Osamy W (2021) Successors of PEGASIS protocol: a comprehensive survey. Comput Sci Rev 39:100368
18. Asqui OP, Marrone LA, Chaw EE (2020) Evaluation of TEEN and APTEEN hybrid routing protocols for wireless sensor network using NS-3. In: Rocha Á, Ferrás C, Montenegro Marin C, Medina García V (eds.) Information Technology and Systems. ICITS 2020. AISC, vol 1137, pp 589–598. Springer, Cham. https://doi.org/10.1007/978-3-030-40690-5_56
19. Wang M, Wang S, Zhang B (2020) APTEEN routing protocol optimization in wireless sensor networks based on combination of genetic algorithms and fruit fly optimization algorithm. Ad Hoc Netw 102:102138
20. Almesaeed R, Jedidi A (2021) Dynamic directional routing for mobile wireless sensor networks. Ad Hoc Netw 110:102301 (2021)

21. He Q, Mou J, Lin B (2021) A robust self-organizing tree-based routing protocol for wireless sensor networks. Math Probl Eng 2021:13 p. Article ID 5932347. https://doi.org/10.1155/2021/5932347
22. Vikneshkumar D, Kaleeswaran D, Yuvaraj D (2020) Analysis on performance comparison of virtual grid-base dynamic route adjustment in wireless detector networks. ICTACT J Commun Technol 11(1):2138–2142
23. Li Q, Aslam J, Rus D (2001) Hierarchical power-aware routing in sensor networks. In: Proceedings of the of the DIMACS Workshop on Pervasive Networking
24. Devi RR, Sethukarasi T (2020) Energy harvesting in physical sensors using hybrid selection of active nodes in optimized GAF protocol for wireless sensor networks. Solid State Technol 63(6):23087–23097
25. Priyadarshi R, Soni SK, Sharma P (2019) An enhanced GEAR protocol for wireless sensor networks. In: Nanoelectronics, Circuits and Communication Systems, pp 289–297
26. Agarwal A, Dev A (2020) Extended RSSI based cluster head selection algorithm for wireless sensor networks. Int J Futur Gener Commun Netw 13(1):559–568
27. Grazia CA, Klapez M, Casoni M (2020) IRONMAN: infrastructured RSSI-based opportunistic routiNg in mobile adhoc networks. In: 2020 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob) (50308), pp 14–19. IEEE
28. Fu X, Fortino G, Pace P, Aloi G, Li W (2020) Environment-fusion multipath routing protocol for wireless sensor networks. Inf Fusion 53:4–19
29. Nakas C, Kandris D, Visvardis G (2020) Energy efficient routing in wireless sensor networks: a comprehensive survey. Algorithms 13(3):72 (2020)
30. Bhasgi SS, Terdal S (2021) A survey on bio-inspired routing algorithms in wireless sensor network. Int Trans J Eng Manag Appl Sci Technol 12(2):12A2Q-1
31. Chen X et al (2020) Artificial intelligence-empowered path selection: a survey of ant colony optimization for static and mobile sensor networks. IEEE Access 8:71497–71511
32. Kaur S, Mahajan R (2017) ACCGP: enhanced ant colony optimization, clustering and compressive sensing based energy efficient protocol. Adv Wirel Mob Commun 10(1):89–109
33. Nasir HJ, Ku-Mahamud KR, Kamioka E (2017) Ant colony optimization approaches in wireless sensor network: performance evaluation. J Comput Sci 13(6):153–164
34. Kakarash ZA, Karim SHT, Ahmed NF, Omar GA (2021) New topology control base on ant colony algorithm in optimization of wireless sensor network. Passer J 3(2):2
35. Muthumayil K, Buvana M, Jayasankar T (2021) Energy utilization using artificial bee colony algorithm for network life time enhancement of homogeneous WSNs. Int J Mod Agric 10(2):1649–1656
36. Wang Z, Ding H, Li B, Bao L, Yang Z (2020) An energy efficient routing protocol based on improved artificial bee colony algorithm for wireless sensor networks. IEEE Access 8:133577–133596
37. Wang H, Wang W, Xiao S, Cui Z, Xu M, Zhou X (2020) Improving artificial bee colony algorithm using a new neighborhood selection mechanism. Inf Sci 527:227–240

# Sign Language Interpreter Using Inception V2 and Faster R-CNN


**Ch. V. N. Koushik, Ch. Tarun, R. V. Neel Kamal, and T. Anuradha**

**Abstract** Sign language is used for communication by many people who have speech or hearing impairment. People with knowledge in particular sign language can easily understand this language but it is difficult for people who don't know that language. So, there is a need for sign language interpreter to make the communication easier between a person using sign language and a normal person who can't understand sign language. Most of existing research were able to identify only the static symbols of American Sign Language with considerable accuracy but failed to identify the motion symbols 'J' and 'Z'. The proposed system tries to do an accurate translation of American Sign Language alphabet symbols using deep learning techniques called inception V2 and Faster RCNN. The novelty of the proposed system is identification of all alphabet symbols including motion symbols with overall model accuracy of 98% by employing inception V2 combined with Faster RCNN technique. The model was tested by displaying signs continuously on webcam.

**Keywords** CNN · Faster RCNN · Inception V2 · ROI · RPN · Sign language

## 1 Introduction

Deaf and Dumb people have their own language known as sign language. There are different sign languages which are analogous to our speaking languages such as English, German, and Hindi etc. which show signs using single hand or both hands. To make the communication between a signer and non-signer, researchers have tried to develop automatic sign language translator for different languages using machine learning and deep learning techniques. But they faced problems like identification of only some of the symbols or identification of symbols with less accuracy. American Sign Language (ASL) [1] contains single hand gestures for showing all 26 alphabets of American English. These signs are called American manual language. It consists

Ch. V. N. Koushik (✉) · Ch. Tarun · R. V. Neel Kamal · T. Anuradha
Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, AP, India
e-mail: vnkoushik.ch@gmail.com



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_71

of 24 static symbols and 2 motion symbols to represent the alphabets 'J' and 'Z' which require movement of the hand.

Most of existing research on American Sign Language translation failed to identify the two motion symbols 'J' and 'Z'. And some research papers identified with less accuracy. By combining faster region based convolution neural network (Faster rcnn) with inception V2 as base network, the proposed research was able to identify motion based gestures with overall accuracy of 98% which is high compared to existing research. The proposed system tries to do a real time translation of hand gestures into equivalent English text, by taking hand gestures as input through video and translating it to text which could be easily understood by a non-signer.

## 2 Related Work

Research was done to convert alphabets of English sign language into text using machine leaning and deep learning techniques. In [2], authors experimented with Self-made American sign language dataset with 26 signs using HSV color space for segmentation and multiclass SVM for classification and obtained an accuracy of nearly 93%. PCA is used to extract the features and KNN model is used for classification in [3] with self-made American sign language data set having 24 signs excluding Z and J, obtained an accuracy of 0% for A, M, N, S and Accuracy of E is 70% for rest accuracy is 100%. In [4], authors used self-generated data set with alphabet excluding J and Z. each letter has 500 images and entire dataset consists of 60,000 images. They Used Deep Convolutional Neural Network with single Convolutional layer, Max pooling layer and obtained Precision of 82%, recall of 80%. In [5] also authors classified only static gestures using a self-made dataset with 26 English language dataset and convolutional neural network model using Inception v3. They obtained accuracies >90%. In [6] authors classified 24 alphabets and 0–9 numbers using 33,000 images dataset and obtained 94.67% accuracy. Authors experimented with self-made dataset with 24 letters excluding Z and J in [7] using CNN algorithm with orientation like center of mass, status of fingers, achieved an accuracy of 92%. In [8] authors used Convolutional Neural network with stochastic gradient descent as optimizer with image size of 200 $\times$ 200 for classifying alphabets. They experimented with NZASL dataset and their own dataset and obtained 82.5% validation set accuracy on NZ ASL dataset and 67% accuracy on letters of the alphabet on their own data set. In [9] authors used a self-made data set of American Sign Language alphabet and CNN algorithm for classification and obtained an accuracy of 90%.

## 3 Proposed Method

The experimental work is done using American Sign Language alphabet data set available in ROBO flow website for training the classifiers. The dataset contains

letters A to Z except J, Z in image format and J, Z in video format. Dataset is divided into 1512 images for training, 144 for validation and 72 for testing. The application detects the symbols continuously shown on the webcam. The implementation is done first with Convolution neural network (CNN) algorithm. Most of the researchers used CNN to classify the letters. The architecture for classification model building with CNN was shown in Fig. 1. Initially the images in the dataset are converted into HSV color space. Then binary thresholding is applied on the image using Otsu thresholding using the opencv library. After dataset pre-processing, training of the CNN model is done using stochastic gradient descent as optimizer, categorical cross entropy as loss function. At the time of real time deployment of the model when the user displays the sign in front of application then image gets converted to HSV color space and binary thresholding is applied and then sign present in the image is detected. Table 1 shows the CNN model summary used in the experimentation. In [15] and [16], authors experimented with Faster-RCNN to identify Arabic sign language and Bangladeshi sign language. Authors in [13] and [14] have discussed about various deep learning based object detection techniques including Faster-RCNN.

In the second phase, experimentation is done with Faster Region based Convolution Neural Network (Faster R-CNN) [10]. Faster Region Based Convolutional

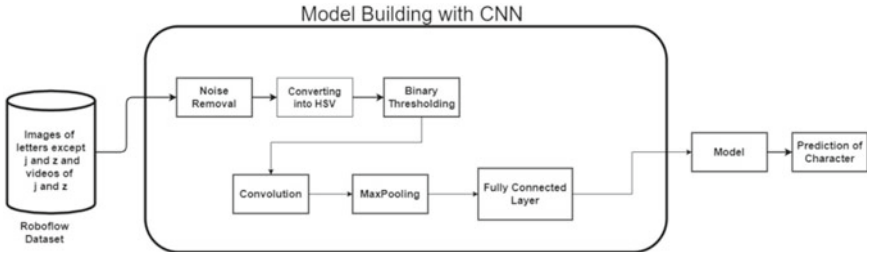



**Fig. 1** Architectural diagram of CNN classifier building

**Table 1** Summary of CNN

| Layer (type) | Output shape | #Param |
|---|---|---|
| conv2d_1 (Conv2D) | (None, 62, 62, 32 | 896 |
| max_pooling2d_1 (MaxPooling2 | (None, 31, 31, 32) | 0 |
| conv2d_2 (Conv2D) | (None, 29, 29, 32) | 9248 |
| max_pooling2d_2 (MaxPooling2) | (None, 14, 14, 32) | 0 |
| conv2d_3 (Conv2D) | (None, 12, 12, 64) | 18,496 |
| max_pooling2d_3 (MaxPooling2) | (None, 6, 6, 64) | 0 |
| flatten_1 (Flatten) | (None, 2304) | 0 |
| dense_1 (Dense) | (None, 256) | 590,080 |
| dropout_1 (Dropout) | (None, 256) | 0 |
| dense_2 (Dense) | (None, 26) | 6682 |

Neural Network (Faster R-CNN) algorithm is robust to outliers and noisy data. Faster R-CNN is embedded with Regional Proposal Network to detect objects in the image. A bounding box around the trained object is returned as output. It returns bounding box coordinates in the form of [ymin, xmin, ymax, xmax].

A dataset with letters A to Z are considered except j, z which are taken as videos. Dataset is downloaded in tf record format from ROBO flow website. Dataset is passed as input to Faster RCNN algorithm with inception V2 as base network. Inception V2 returns feature map and it acts as input for Region Proposal Network (RPN) algorithm. RPN algorithm working with anchor boxes is used to find regional proposals in the feature map generated through base network. The feature map derived from base network is passed through a rectangular sliding window of size nxn, where for each window K region proposals are generated. Each proposal is parameterized according to a reference box which is called an anchor box. Scales, Aspect Ratio are parameters for anchor boxes. Generally, there are 3 scales and 3 aspect ratios and thus there is a total of $K = 9$ anchor boxes. In this project we considered different scales like 0.2, 0.5, 1.0 and 2.0 and aspect ratios like 0.5, 1.0 and 2.0. In order to classify object proposal a metric named Intersection over union (IOU) with a threshold of 0.699999988079 is used. Anchor box is considered as background if the corresponding IOU is negative that is lesser than threshold and as foreground if the IOU is positive. Once the RPN is completed, the next step of Faster RCNN is ROI Pooling layer. ROI pooling layer is used to extract a fixed-length feature vector from each region proposal. Each Feature Vector is sent as input to fully connected layer using which we can detect the letter from the image. Softmax is used to return classification score whereas bounding box regression is used to set width and height of the bounding box.

*Feature Network*

Feature Network is usually a well-known pre-trained neural network such as Inception v2. This network is used to derive a feature map which is sent to RPN for object proposals. Figure 2 RPN is a network with 3 convolutional layers. There is one layer which inputs data to 2 layers i.e. classification and for bounding box regression.

*ROI Pooling*

In R-CNN [11] object detection is done by generating region proposals i.e., the bounding boxes through selective search. In Fast R-CNN [12], associated bounding

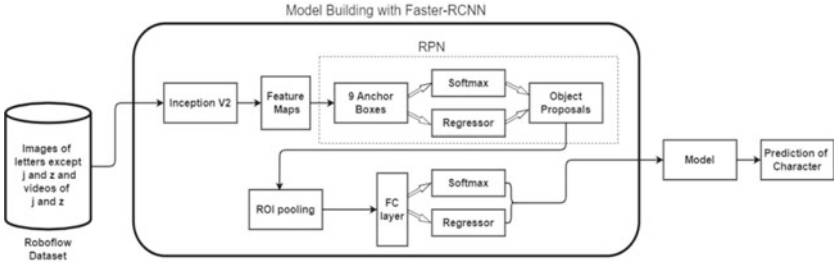



**Fig. 2.** Architecture of faster R-CNN

box position and object class is calculated based on multi task loss given in Eqs. 1–4.

$$L\left(p, u, t^u, v\right) = L_{\text{cls}}(p, u) + \lambda[u \geq 1]L_{\text{loc}}\left(t^u, v\right) \tag{1}$$

$$L_{\text{cls}}(p, u) = -\log p_u \tag{2}$$

$$L_{\text{loc}}(t^u, v) = \sum_{i \in \{x, y, w, \text{h}\}} \text{smooth}_{L_1}(t_i^u - v_i) \tag{3}$$

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \tag{4}$$

where L_cls and L_los is the log loss for true class u and bounding box. RPN returns object proposals on the feature map after classification and bounding box Regression. Equation 5 describes the formula of RPN. The returned Feature map is sent for ROI pooling. Classifier is used to determine probability of the proposal having the object and Regression is used to calculate regression coefficients to modify height and width of the object. Region of Interests (ROIs) are regions that have a high probability of containing any object. Faster RCNN Architecture diagram is shown in Fig. 2. ROI pooling takes each ROI from the input and takes the object proposal of input feature map that corresponds to ROI and converts that feature-map section into a fixed dimension map.

$$\text{L}\left(\{\text{p}_\text{i}\}, \{\text{t}_\text{i}\}\right) = \frac{1}{\text{Ncls}} \sum \text{L}_{\text{cls}}\left(\text{p}_\text{i}, \text{p}_\text{i}^*\right) + \lambda \frac{1}{\text{Nreg}} \sum \text{p}_\text{i}^* \text{L}_{\text{reg}}\left(\text{t}_\text{i}, \text{t}_\text{i}^*\right) \tag{5}$$

## 4 Experimental Results

Figure 3 shows the result obtained through CNN algorithm. It was observed that CNN gave better accuracy when the background of the image was plain but if the

**Fig. 3** CNN results

background was not plain, sometimes, the resultant output was wrong. And it was also observed that the sign need to be shown only in the fixed square box to get the correct result otherwise the sign cannot be recognized.

Figure 4 shows the change of accuracy Vs. epochs with epochs on x-axis and accuracy on y-axis for CNN algorithm. It was observed that training and testing accuracy was greater than 90% after 25 epochs. Figure 5 shows the change in loss with change in epochs for CNN. It was observed that loss became less than 0.2 for test data after 15 epochs and it was almost constant after that.

Figure 6 shows the result obtained through the Faster-RCNN algorithm. It was observed that even though the background of the image has mixed colors, the correct sign was recognized and also got the bounding box around the recognized sign to make it understandable to the users. When symbols were shown continuously, the output displayed the sequence text. Figure 7 shows output of faster R-CNN, when motion gestures 'J and Z' are given as inputs. It can be observed that the algorithm identified the symbol J with 96% and Z with 92% confidence. Figure 8 depicts the

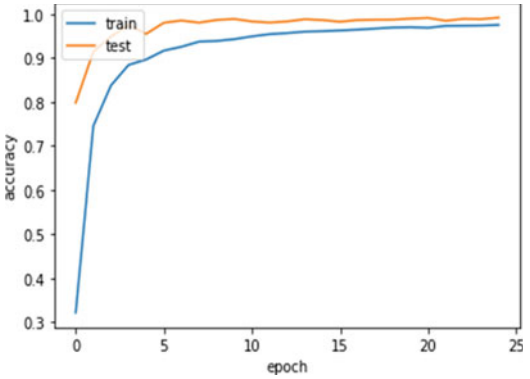



**Fig. 4** Epochs vs. accuracy in CNN

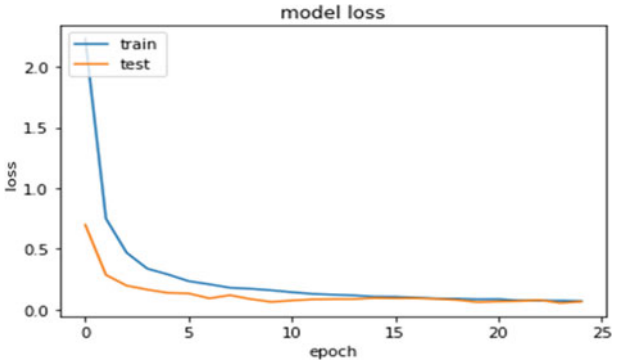



**Fig. 5** Epochs vs. loss in CNN

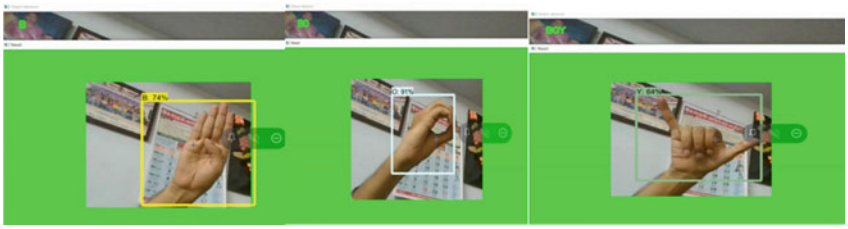


**Fig. 6** Result of faster R-CNN showing B, O, Y in sequence

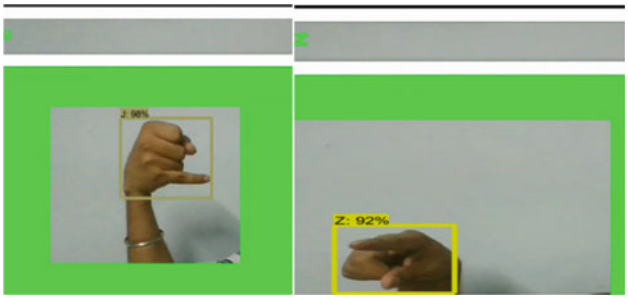



**Fig. 7** Output of faster R-CNN for signs J and Z

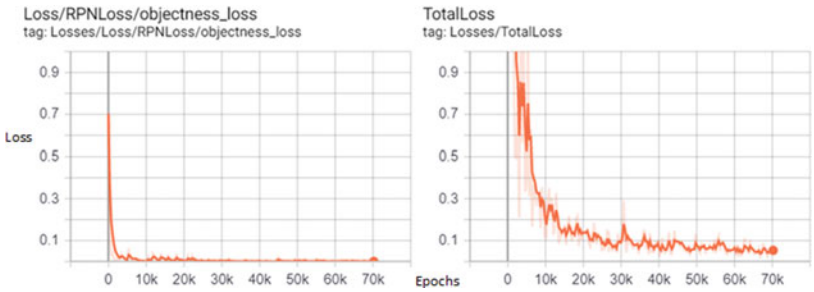



**Fig. 8** Loss of RPN and softmax layers of faster R-CNN

loss of Region Proposal Network and loss of softmax classifier in Faster R-CNN architecture. Here x-axis shows number of epochs and Y-axis shows the loss. Final loss after softmax layer reduced to 0.1 after 70 k steps. Loss value is constant almost from 40,000 steps in the classifier.

## 5 Analysis and Discussion

Among the two classifiers used in the proposed work, faster R-CNN model gave higher accuracy than the CNN model. And also, CNN model requires pre-processing techniques likely changing image into HSV color space and binary thresholding while Faster RCNN does not have this overhead. For CNN model background needs to be clear, a small disturbance in background caused decline in the accuracy. Figure 9 shows the outputs of Faster R-CNN and CNN models when the sign 'A' is shown with noisy background on webcam. Faster R-CNN gave correct output with 94% confidence, whereas CNN failed to show the correct output as the background was noisy. The confidence of Faster R-CNN model in recognizing the symbol can be seen above the bounding box.

Figure 10 describes performance of the model. It shows precision, recall and f1-score of the model for different letters and overall accuracy of the model when model is applied on testing data. Table 2 shows the comparison of accuracies for the literature surveyed and the proposed work using CNN and faster R-CNN. It can be observed that the proposed work using CNN algorithm with preprocessing techniques employed worked almost equal to or better than existing works which identified all 26 alphabets with >90% train and test accuracies after 25 epochs (Fig. 4). Proposed Faster R-CNN with inception V2 as base network outperformed all existing works with overall model accuracy of 98% (Fig. 10).

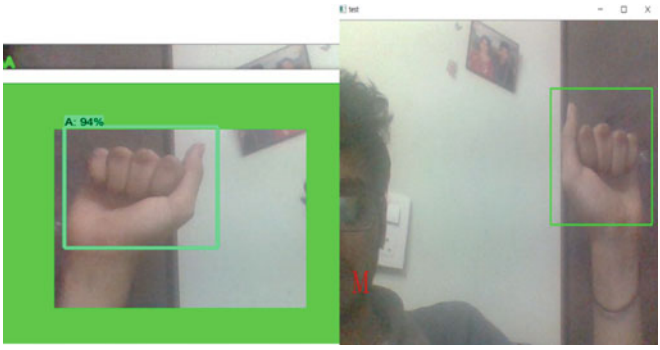
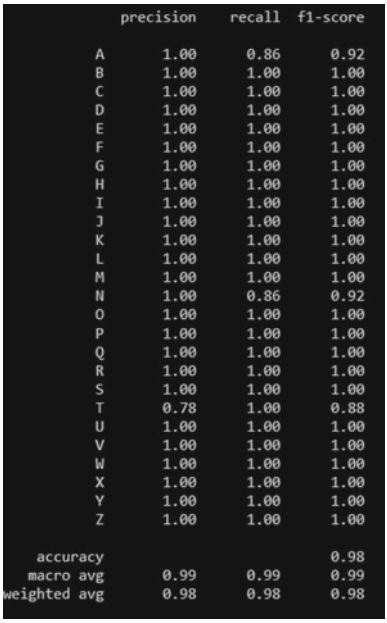



**Fig. 9** Output of sign 'A' with noisy background (1) Faster R-CNN and (2) CNN



| | precision | recall | f1-score |
|---|---|---|---|
| A | 1.00 | 0.86 | 0.92 |
| B | 1.00 | 1.00 | 1.00 |
| C | 1.00 | 1.00 | 1.00 |
| D | 1.00 | 1.00 | 1.00 |
| E | 1.00 | 1.00 | 1.00 |
| F | 1.00 | 1.00 | 1.00 |
| G | 1.00 | 1.00 | 1.00 |
| H | 1.00 | 1.00 | 1.00 |
| I | 1.00 | 1.00 | 1.00 |
| J | 1.00 | 1.00 | 1.00 |
| K | 1.00 | 1.00 | 1.00 |
| L | 1.00 | 1.00 | 1.00 |
| M | 1.00 | 1.00 | 1.00 |
| N | 1.00 | 0.86 | 0.92 |
| O | 1.00 | 1.00 | 1.00 |
| P | 1.00 | 1.00 | 1.00 |
| Q | 1.00 | 1.00 | 1.00 |
| R | 1.00 | 1.00 | 1.00 |
| S | 1.00 | 1.00 | 1.00 |
| T | 0.78 | 1.00 | 0.88 |
| U | 1.00 | 1.00 | 1.00 |
| V | 1.00 | 1.00 | 1.00 |
| W | 1.00 | 1.00 | 1.00 |
| X | 1.00 | 1.00 | 1.00 |
| Y | 1.00 | 1.00 | 1.00 |
| Z | 1.00 | 1.00 | 1.00 |
| accuracy | | | 0.98 |
| macro avg | 0.99 | 0.99 | 0.99 |
| weighted avg | 0.98 | 0.98 | 0.98 |

**Fig. 10** Describes accuracies of corresponding letters

**Table 2** Comparison with previous works

| Reference number | Classification technique | Alphabets identified | Accuracy |
|---|---|---|---|
| [2] | SVM | All 26 (A–Z) | 93% for static symbols |
| [3] | KNN | Only 24 static alphabets (excluding J and Z) | 0% for A, M, N, S, 70% for E and 100% for remaining |
| [4] | CNN | Only 24 static alphabets (excluding J and Z) | Precision 82% recall 80% |
| [5] | CNN | Only 24 static alphabets (excluding J and Z) | >90% |
| [6] | CNN | Only 24 static alphabets (excluding J and Z) | 94.6% |
| [7] | CNN | Only 24 static alphabets (excluding J and Z) | 92% |
| [8] | CNN | All 26 (A–Z) | 82.5% validation set accuracy on NZ ASL dataset and 67% on their own dataset |
| [9] | CNN | All 26 (A–Z) | 90% |
| Proposed work | CNN | All 26 (A–Z) | >90% for train test data |
| Proposed work | Faster RCNN | All 26 (A–Z) | 98% for all symbols |

## 6 Conclusion

Conversion of American Sign Language to text was implemented using two deep learning models namely CNN and Faster R-CNN with inception V2 as base model. Accuracies of CNN model varied in noisy backgrounds whereas Faster RCNN model worked well even with a noisy background as Faster RCNN was robust to outliers. The proposed Faster RCNN worked well even with motion symbols of sign language with overall accuracy of 98% so the model can be used for all symbols equally well which were a drawback of previous researches. The future work corresponds to conversion of generated text to speech to give non signers a real speech to speech communication feel.

## References

1. American Sign Language. https://en.wikipedia.org/wiki/American_Sign_Language. Accessed 20 Nov 2020
2. Kumar KT, Dominic MM (2016) Sign language recognition. In: 2016 3rd international conference on recent advances in information technology (RAIT), Dhanbad, India, pp 422–428. https://doi.org/10.1109/RAIT.2016.7507939
3. Upendran S, Thamizharasi A (2014) American Sign Language interpreter system for deaf and dumb individuals. In: 2014 international conference on control, instrumentation, communication and computational technologies (ICCICCT), Kanyakumari, India, pp 1477–1481. https://doi.org/10.1109/ICCICCT.2014.6993193
4. Ameen S, Vadera S (2017) A convolutional neural network to classify American Sign Language fingerspelling from depth and colour images. Expert Syst 34:e12197. https://doi.org/10.1111/exsy.12197
5. Das SG, Suratwala K, Kalbande D (2018) Sign language recognition using deep learning on custom processed static gesture images. In: 2018 international conference on smart city and emerging technology (ICSCET), Mumbai, pp 1–6. https://doi.org/10.1109/ICSCET.2018.8537248
6. Beena MV (2017) Automatic sign language finger spelling using convolution neural network: analysis
7. Bachani S, Dixit S, Chadha R (2020) Sign language recognition using neural network. Int Res J Eng Technol 7(4):583–586
8. Bheda V, Radpour D (2017) Using deep convolutional networks for gesture recognition in American sign language
9. Shahriar S et al (2018) Real-time American sign language recognition using skin segmentation and image category classification with convolutional neural network and deep learning. In: TENCON 2018 - 2018 IEEE region 10 conference, Jeju, Korea (South), pp 1168–1171. https://doi.org/10.1109/TENCON.2018.8650524
10. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. In: Proceedings of the advances in neural information processing systems, Montreal, Canada, pp 91–99
11. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE conference on computer vision and pattern recognition, pp 580–587. https://doi.org/10.1109/CVPR.2014.81
12. Girshick R (2015) Fast R-CNN. In: Proceedings of the IEEE international conference on computer vision (ICCV), Santiago, Chile, pp 1440–1448

13. Wang J, Zhang T, Cheng Y, Al-Nabhan N (2021) Deep learning for object detection: a survey. Comput Syst Sci Eng 38(2):165–182
14. Jiao L et al (2019) A survey of deep learning-based object detection. IEEE Access 7:128837–128868. https://doi.org/10.1109/ACCESS.2019.2939201
15. Alawwad RA, Bchir O, Ben Ismail MM (2021) Arabic sign language recognition using faster R-CNN. Int J Adv Comput Sci Appl (IJACSA) 12(3):2021. http://dx.doi.org/10.14569/IJACSA.2021.0120380
16. Hoque OB, Jubair MI, Islam MS, Akash A, Paulson AS (2018) Real time Bangladeshi sign language detection using faster R-CNN. In: 2018 international conference on innovation in engineering and technology (ICIET), pp 1–6. https://doi.org/10.1109/CIET.2018.8660780

# Simultaneous Sparse Representations with Partially Varying Support


**Lakshmi Madhuri Sathi, Varsha Juluri, Santhoshini Tangudu, Swathy Sreeram, Kavya Kuzhithara Sajan, and Sandeep Palakkattillam**

**Abstract** The idea of sparse representations approximates a signal as a linear combination of a few atoms from a redundant over complete dictionary. Orthogonal Matching Pursuit (OMP) is a greedy algorithm used for the computation of sparse representations. The idea of simultaneous sparse representations is to jointly compute the sparse representations of a group of signals with a common support for their corresponding sparse representations. The OMP algorithm was later extended to Simultaneous-OMP (SOMP) for computing simultaneous sparse representations of a group of signals. The strict constraint on the support of non-zero coefficients makes SOMP unusable in many situations. In this work, an extension of the SOMP algorithm for computing simultaneous sparse representations with a partially varying support is proposed. The experiments demonstrate that the proposed algorithm achieves superior performance over SOMP, when the support of the non-zero sparse representation coefficients is not exactly same for all the sparse representations.

**Keywords** Dictionary · OMP · S-OMP · Sparse representation · Support

## 1 Introduction

The rapid advancement of technology in our day-to-day life, provides us access to various sources of information like videos, medical data etc. Each of these signals have some internal structure which is the fundamental characteristic of that signal. It is crucial to identify these structures for achieving a good performance in various signal processing techniques used for processing such signals. In order to characterize and to mathematically describe these structures, there is a dependence on various models [1]. The performance of various signal processing techniques relies on how good these models characterize the signals of interest. One among them is the sparse-land model [2], which is a recently popularized and widely studied model. It assumes that a given signal $y \in \mathbb{R}^n$ can be represented as a linear combination of a few prototype

L. M. Sathi · V. Juluri · S. Tangudu · S. Sreeram · K. Kuzhithara Sajan · S. Palakkattillam (✉)
Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_72

signals called atoms, selected from a collection of such atoms called the dictionary, which can be represented as Eq. (1),

$$y = D\alpha \tag{1}$$

where, $\alpha \in \mathbb{R}^m$ is the sparse representation vector with only a few non-zero coefficients, and $D \in \mathbb{R}^{n \times m}$ is the dictionary matrix ($m \gg n$) [3, 4]. The dictionary $D$ is selected in such a way that the columns of $D$ are highly correlated to the signal [5, 6], which makes the sparse representation $\alpha$ more sparse. The previous works have shown that the sparse-land model produces promising results in different signal and image processing problems, such as, image denoising [7], single image resolution [8], and various machine learning problems [9, 10]. In the literature this problem is commonly known as the 'sparse coding' problem [2]. In the case of practical signals, the sparse coding problem can be relaxed as the approximation $y \approx D\alpha$. To compute sparse representation, two different approaches can be used [2], namely, convex relaxation and greedy methods. In this project, the focus is on the computationally simpler greedy algorithms. Examples of greedy algorithms include Matching Pursuit (MP) and Orthogonal Matching Pursuit (OMP) [11]. In several situations, computation of sparse representations of a group of $P$ signals with a common support maybe required. For example, sparse representations of multiple observations of a signal corrupted with random noise [12]. It has been proven [13, 14] that the simultaneous computation of sparse representations of a group of signals is advantageous over the independent computations using one of the above methods for finding the sparse representation of a single signal.

The simultaneous sparse coding problem is based on a strict assumption that the sparse representations represented in a matrix form $\Gamma = [\alpha_1, \alpha_2, ..., \alpha_p]$, where $\alpha_1, \alpha_2, ..., \alpha_p$ are the columns of the matrix $\Gamma$, and sparse representations of a group of $P$ signals where all of them have same support for their non-zero coefficients. In many practical situations, this condition may not be exactly satisfied. In several situations, a partially common support may be needed. An example situation where simultaneous sparse representation with a relaxed constraint needed is the sparse representation of color image patches. If there was a tight constraint that the sparse representation of all the patches should have the same support, then such a representation restricts the variability in colors. The SOMP algorithm cannot be used in this case because of the strict assumption of common support. In this paper, an algorithm is proposed for computing simultaneous sparse representations of a group of signals with a partially varying (relaxed) support for their non-zero values. Let, $y_1$, $y_2$, …, $y_p$, be $P$ signal vectors and $\alpha_1$, $\alpha_2$, …, $\alpha_p$ are their corresponding sparse representations. The assumption is that there are at most $T$ non-zero values in each of $\alpha_1$, $\alpha_2$, …, $\alpha_p$. Let $T_1$ coefficients out of $T$ have the same support among $\alpha_1$, $\alpha_2$, …, $\alpha_p$, and $T_2$ coefficients have varying support, where $T_1 + T_2 = T$. This work focuses on a greedy method for computing the simultaneous sparse representations under this condition. This paper shows that the proposed algorithm achieves better performance over SOMP using simulations on synthetic data.

## 2 Related Works

A wide variety of algorithms for computing the sparse representations have been previously proposed in the literature by utilizing the convex relaxation approach and the greedy approach. In the work presented in [7], the authors have proposed an algorithm for denoising natural images. Sparse Representation algorithms show favorable results while dealing with image fusion techniques used in the medical field [15]. In [16], the authors have proposed a model using a dictionary learning algorithm, and chosen the dictionary columns that are similar to the input signal. In [17], the authors review the OMP algorithm which is closely related to the proposed work. In the research paper [18], the authors modeled the image denoising technique by predicting on the correlated sparse model. Orthogonal Matching Pursuit (OMP) [11] is a greedy algorithm that computes an approximate solution to the sparse coding problem. The OMP assumes that a signal vector $y \in \mathbb{R}^n$, dictionary $D$, and the stopping criteria are given. This signal vector $y$ is assumed to be containing additive white Gaussian noise of variance $\sigma^2$ as, $y = y_t + n$ where, $y_t$ is the original signal and $n$ is the noise vector. The OMP algorithm iteratively selects the support of non-zero coefficients by selecting the best column of the matrix $D$ that is matched to the residual left over from the previous iterations. The algorithm then computes the non-zero coefficients by solving least square minimization problems. In every iteration, the algorithm partially updates the sparse representation vector by adding one non-zero coefficient, and the recovered signal obtained in that iteration can be computed by multiplying the dictionary with the presently estimated sparse representation vector. The iterations can be stopped when the number of non-zero values reach some fixed threshold or when the residual energy is approximately equal to the noise variance $\sigma^2$. Similar to the algorithms proposed for computing the sparse representation of a single signal vector, various algorithms have been also proposed for computing sparse representations of a group of signals $Y = [y_1, y_2, \ldots, y_p]$ by utilizing convex relaxation [13] and greedy approaches [14]. In this paper, the authors have proposed Simultaneous Orthogonal Matching Pursuit (SOMP), which can be seen as an extension of the OMP algorithm into the case of simultaneous sparse representations. Each of the signals is assumed to be corrupted by Additive White Gaussian noise of variance $\sigma^2$ as $Y = Y_t + N$ where, $Y_t$ is the original signal whereas, the available signal is $Y_t + N$ where $N$ is the noise. The SOMP algorithm also selects the supports of non-zero values by choosing the best column of the matrix $D$ that is matched to all the residuals that is left over from the previous iterations. The algorithm can be terminated in a similar fashion to that of the OMP algorithm.

## 3 Methods

In this section, the proposed algorithm is described for computing simultaneous sparse representations of a group of signals with a relaxed constraint on the support of non-zero sparse representation coefficients. Let us consider a set of $P$ true signals, $y_{01}, y_{02}, \ldots, y_{0P}$ such that $y_{0j} \approx D\,\alpha_{0j}$ with $||\alpha_{0j}||0 \leq T \;\; for \;\; j = 1, 2..., P$. As the true signals cannot be measured exactly in many cases, the noisy versions corrupted by AWGN of variance $\sigma^2$ were considered, i.e., $y_j = y_{0j} + n_j$ where, $n_j \in \mathbb{R}^n$ is the noise vector. The proposed algorithm computes the sparse representations of this group of input signals $y_1, y_2, \ldots, y_p$ by operating in two modes, i.e., SOMP followed by OMP mode. The residual and dictionary columns selected by the SOMP iterations are saved and carried to the OMP mode. Depending on the method for switching from the SOMP mode to OMP mode 3 different cases of the proposed algorithm are discussed.

### *3.1 Case 1*

Case 1 assumes a noiseless scenario, where $T_1$ and $T_2$ are known. The SOMP mode undergoes exactly $T_1$ iterations and saves the selected dictionary columns and the residual signals obtained in the last iteration i.e., $T_1$ $^{\text{th}}$ iteration. The algorithm then proceeds with the OMP mode, which it starts with the residuals and the dictionary columns as obtained from the SOMP mode. The proposed algorithm continues in the OMP mode for $T_2$ iterations and the sparse representation vectors obtained from OMP mode are chosen as the final estimates of the sparse representations- $\alpha_{01}, \alpha_{02}, \ldots \alpha_{0p}$.

**Algorithm 1:** Algorithm of Case 1

Input:( $y_1$, $y_2$, ... $y_p$), $T_1$, $T_2$, $D$, Initialization: $R = Y$, $t = 0$, $d_{somp} = [null]$;
while $t <= T_1$ do
  Choose the best column $d_j$ from $D$ using, $max(abs(D^T R))$;
Augment $d_j$ to the partial dictionary matrix $d_{somp}$;
  Compute non-zero values, $\alpha' = (d_{somp}^T d_{somp})^{-1} d_{somp}^T Y$;
  Update the residual as $Y - (d_{somp}\alpha')$;
  Save residual to $R_{somp}$ and increment $t$;
end
for $i$ from 1 $to$ $P$ do
  $r = R_{somp}(:, i)$ $and$ $t_1 = 0$;
while $t_1 <= T_2$ do
    Choose the best column $d_k$ from $D$ using, $max(abs(D^T r))$;
    Append the index of the best column that is chosen to the in $d_x$ and $d_k$ to the matrix $d_{somp}$, which are obtained from SOMP;
    Compute non-zero values, $\Gamma' = (d_{somp}^T d_{somp})^{-1} d_{somp}^T Y$;
    Update the residual as $Y - d_{somp}\Gamma'$ and increment $t_1$;
end
  Compute $\Gamma$ from $\Gamma'$ at the indices that are saved;
  Compute the estimate of $y_{01}, y_{02}, \ldots y_{0P}$ using $Y = D\Gamma$;
end

Case 1 assumes that the number of non-zero values of the common and varying supports are known exactly. However, in most of the practical situations, the values of $T_1$ and $T_2$ may not be known.

## 3.2 Case 2

The sparse representations with partially varying support can be written as, $\alpha = \alpha_{T1} + \alpha_{T2}$ where, $\alpha_{T1,T2} \in \mathbb{R}^m$ are the sparse representation vectors with non-zero values at common and varying supports respectively. Then we can write,

$$y_{0j} = D(\alpha_{T1} + \alpha_{T2}) = {y_{0j}}^{T1} + {y_{0j}}^{T2} \tag{2}$$

where ${y_{0j}}^{T1}$ and ${y_{0j}}^{T2}$ in Eq. (2) denote the signal component produced by the non-zero coefficients with common and varying supports respectively. In case 2, the assumption is that the average energy of the signal ${y_{0j}}^{T1}$ i.e., $\|{y_{0j}}^{T1}\|_2^2$ is assumed to be known and is denoted as $\beta$. Here, the observed signals $y_1$, $y_2$, …, $y_P$ are noisy. In case 1 the proposed algorithm continues to be in the SOMP mode until the average energy of the restored signal becomes $\beta$ i.e., whenever the signal components produced by the non-zero coefficients of common supports are retrieved. The SOMP

iterations will be stopped when the average energy of the recovered signal is greater than $\beta$ and the residual and the selected dictionary columns are saved. The proposed algorithm then runs in OMP mode till the stopping criteria is reached, with the initial residual and selected dictionary columns obtained from SOMP mode.

**Algorithm 2**: Algorithm of Case 2

Input: $(y_1, y_2, \ldots y_P)$, $\beta$, $\sigma^2$, $D$, Initialization: $R = Y, d_{somp} = 0, \hat{Y}_s = 0$;
while(average energy of $\hat{Y}_s <= \beta$) do
  Choose the best column $d_j$ from $D$ using, $max\left(abs(D^T R)\right)$;
Augment $d_j$ to the new dictionary $matrix\ d_{somp}$;
  Compute non-zero values, $\alpha' = \left(d_{somp}^T d_{somp}\right)^{-1} d_{somp}^T Y$;
  Update the residual as $Y - \left(d_{somp}\alpha'\right)$;
  Save residual to $R_{somp}$ and increment $t$;
  Compute $\hat{Y}_s = d_{somp}\alpha'$, where $\hat{Y}_s$ is the partially recovered signal obtained in individual iteration;
end
for $i$ from 1 $to$ $P$ do
$r = R_{somp}(:, i)$;
while (energy of residual$(r) > \sigma^2$) do
    Choose the best column $d_k$ from $D$ using, $max\left(abs(D^T R)\right)$;
    Append the index of the best column that is chosen to the in $d_x$ and $d_k$ to the matrix $d_{somp}$ , which are obtained from SOMP;
    Compute non-zero values, $\Gamma' = \left(d_{somp}^T d_{somp}\right)^{-1} d_{somp}^T Y$;
    Update the residual as $y - \left(d_{somp}\Gamma'\right)$;
end
  Compute $\Gamma$ from $\Gamma'$ at the indices that are saved
  Compute the estimate of $y_{01}, y_{02}, \ldots y_{0P}$ using $Y = D\Gamma$ ;
end

In case 2, it has been assumed that the average energy of the common component of the signal is known. However, in many of the practical situations this average energy may not be known exactly.

### *3.3 Case 3*

In case 3, the SOMP mode of the proposed algorithm continues for a fixed number of iterations regardless of the values of $T_1$ and $T_2$. The exact values of $T_1$ and $T_2$ are assumed to be unknown. The proposed algorithm begins in the SOMP mode and continues for a fixed $l_1$ iterations and then switches over to the OMP mode. The iterations of the OMP mode are continued till the residual energy becomes less

than or equal to $\sigma^2$. The residual and the selected dictionary columns obtained from the last iteration of the SOMP mode are saved and are carried to OMP. In practice, an appropriate value of $T_1$ has to be calculated experimentally for a given class of signals. An example database of signals can be taken from the given class of signals and the value of $l_1$ for which the sparse representation error is minimum can be computed by running the proposed algorithm for different values of $l_1$.

**Algorithm 3**: Algorithm of Case 3

Input: $(y_1, y_2, \ldots y_P)$, $l_1$, $\sigma^2$, $D$, Initialization: $R = Y, t = 0, d_{somp} = 0$;
while $t <= l_1$ do
  Choose the best column $d_j$ from $D$ using, $max\left(abs(D^T R)\right)$;
Save $d_j$ to the new dictionary matrix $d_{somp}$;
  Compute non-zero values $\alpha' = \left(d_{somp}^T d_{somp}\right)^{-1} d_{somp} Y$ ;
  Update the residual as $Y - \left(d_{somp}\alpha'\right)$;
  Save residual to $R_{somp}$ and increment $t$;
end
for $I$ from 1 $to$ $P$ do
$r = R_{somp}(:, \mathrm{i})$;
while(energy of residual$(r) > \sigma^2$) do
    Choose the best column $d_k$ from $D$ using, $max\left(abs(D^T R)\right)$;
    Append the index of the best column that is chosen to the in $d_x$ and $d_k$ to the matrix $d_{somp}$ , which are obtained from SOMP;
    Compute non-zero values, $\Gamma' = \left(d_{somp}^T d_{somp}\right)^{-1} d_{somp}^T Y$,
    Update the residual as $y - \left(d_{somp}\Gamma'\right)$ ;
end
  Compute $\Gamma$ from $\Gamma'$ at the indices that are saved;
  Compute the estimate of $y_{01}, y_{02}, \ldots y_{0P}$ using $Y = D\Gamma$;
end

## 4 Experimental Results

In contrast with the conventional sparse coding, the algorithms for computing simultaneous sparse representations are really limited. SOMP has been consistently used for simultaneous sparse representations over several years. Hence, the proposed algorithm is evaluated by comparing its performance with SOMP, which is probably the best closest candidate for a comparison. This section describes the experimental setup that is used for testing the proposed algorithm and comparing its performance with SOMP. The performance comparison of OMP and SOMP is also carried out to verify the effectiveness of SOMP over OMP in the case of simultaneous sparse

representation of a group of signals with common support. All the tests are carried out on synthetic data generated using MATLAB. An over-complete dictionary, of size $20 \times 50$, filled with random Gaussian elements of zero mean and unit variance is generated. Each of the columns in the dictionary is normalized, so that the normalized to have unit $l_2$ norm. The performance evaluation is carried out by computing the Mean Square Error between the signal computed by the proposed algorithm/SOMP/OMP and the original signals $y_0 = [y_{01}, y_{02}, \ldots, y_{0P}]$.

In Fig. 1(a), the simulation takes different SNR values by adding noise of different variance and uses OMP and SOMP for recovering the signal. For large SNR values, the performance of OMP and SOMP have comparable performance. In Fig. 1(b), $snr = 2$ and $T = 10$. $MSE$ increases as $T$ increases, since it is deviating further from the sparsity assumption. The Fig. 2 shows the comparison between case 1 of the proposed algorithm and SOMP. Here, SNR is varied from 1 to 10 and $T_1$ is taken as 1 and 3, the results of each are shown in Fig. 2(a) and 2(b) respectively. The Fig. 3 correspondingly shows the comparison between the proposed algorithm of case 2 and SOMP. The SNR is varied from 1 to 10 and $T_1$ is taken as 1 and 3 with constant $\beta$ and the results of each are shown in Fig. 3(a) and 3(b) respectively. In this experiment it is shown that SOMP is advantageous over OMP in simultaneous computation of

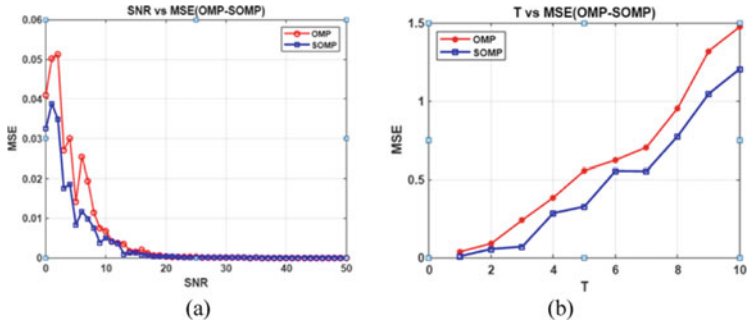



**Fig. 1** Graphs of OMP and SOMP

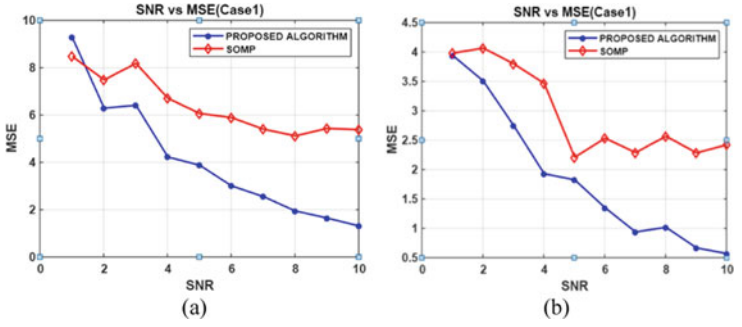



**Fig. 2** Graphs of SNR vs MSE for proposed algorithm and SOMP

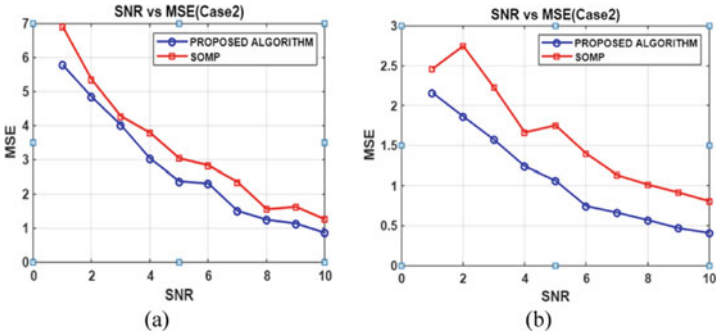



**Fig. 3** Graphs of SNR vs MSE for proposed algorithm and SOMP

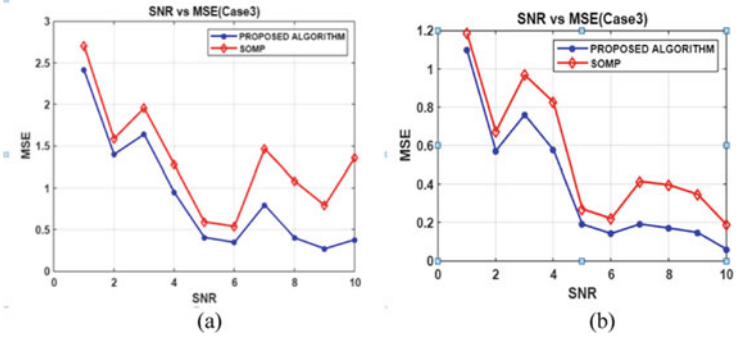



**Fig. 4** Graphs of SNR vs MSE for proposed algorithm and SOMP

sparse representations of groups of signals with a common support in a noise-less scenario.

In Fig. 4 the performance of case 3 of the proposed algorithm is compared against SOMP. The actual value of $l_1$ in case 3 of the proposed algorithm is identified experimentally for a set of signals. The MSE increases with the increase in the probability of getting actual $l_1$ which is shown in Fig. 4(a) and 4(b) respectively.

## 5 Discussion

In Fig. 1(a), Fig. 1(b) SOMP shows a superior performance when compared to OMP. In Fig. 2 the proposed algorithm performs better than SOMP significantly with higher SNR values. Similarly, in Fig. 3 and Fig. 4, the proposed algorithm achieves better performance when compared to SOMP in the respective cases. MSE decreases with increase in SNR, since signal strength increases with SNR. With increase in common supports i.e., $T_1$, MSE decreases in each case. The proposed algorithm gives better

results in case 3 when compared with case 1, case 2 and SOMP. The performance comparison between SOMP and the proposed algorithm has been evaluated and differentiated in the experiments.

## 6 Conclusions

The proposed algorithm achieves better performance over SOMP as SNR increases, whereas the performance is comparable for smaller SNR values. In many practical scenarios, the correct values of the parameters assumed in case 1 and case 2 may not be known. In order to alleviate this issue, the algorithm proposed as in case 3 is implemented where the SOMP iterations are fixed to a constant value which can be tuned to different cases by using example signals. The focus of the future work is to make the proposed algorithm parameter free, so that switching from SOMP to OMP mode can be done automatically. As such, the future research also targets to utilize the proposed algorithm for color image restoration problems, in which, simultaneous sparse representation of color image patches with partially varying support is expected to be beneficial.

## References

1. Elad M (2012) Sparse and redundant representation modelling–what next. IEEE Signal Process Lett 19(12):922–928
2. Elad M (2010) Sparse and Redundant Representations–From Theory to Applications in Signal and Image Processing. Springer, New York. https://doi.org/10.1007/978-1-4419-7011-4
3. Baraniuk RG, Candes E, Elad M, Ma Y (2010) Applications of sparse representation and compressive sensing [Scanning the Issue]. Proc IEEE 98(6):906–909
4. Zhang Z, Xu Y, Yang J, Li X, Zhang D (2005) A survey of sparse representation: algorithms and applications. IEEE Access 3:490–530
5. Elad M, Aharon M (2006) Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans Image Process 15:3736–3745
6. Sadeghi M, Babaie-Zadeh M, Jutten C (2013) Dictionary learning for sparse representation: a novel approach. IEEE Signal Process Lett 20(12):1195–1198
7. Thangavel SK, Rudra S (2020) A sparse representation based learning algorithm for denoising in images. https://doi.org/10.1007/978-3-030-37218-7-89
8. Ravishankar S, Reddy CN, Tripathi S, Murthy KVV (2011) Image super resolution using sparse image and singular values as priors. In: Real P, Diaz-Pernil D, Molina-Abril H, Berciano A, Kropatsch W (eds.) Computer Analysis of Images and Patterns. CAIP 2011. LNCS, vol 6855. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23678-5_45
9. AswathiVarsha KTK, Lalitha S (2017) Stress recognition using sparse representation of speech signal for deception detection applications in Indian context. In: IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp 1–7
10. Menon R, Nair SS, Srindhya K, Kaimal MR (2013) Sparsity-based representation for categorical data. In: IEEE Recent Advances in Intelligent Computational Systems (RAICS), pp 74–79

11. Pati YC, Rezaiifar R, Krishnaprasad PS (2005) Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, vol 1, pp 40–44. Pacific Grove, CA, USA
12. Eldar YC, Mishali M (2009) Robust recovery of signals from a structured union of subspaces. IEEE Trans Inf Theory 55(11):5302–5316
13. Tropp JA, Gilbert AC, Strauss MJ (2003) Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. Signal Process 86(3)
14. Tropp JA, Gilbert AC, Strauss MJ (2003) Algorithms for simultaneous sparse approximation. Part II: Greedy pursuit. Signal Process 86(3)
15. Li Q, Wang W, Chen G, Zhao D (2021) Medical image fusion using segment graph filter and sparse representation. Comput Biol Med 131
16. Singh S, Anand RS (2020) Multimodal medical image sensor fusion model using sparse K-SVD dictionary learning in nonsubsampled shearlet domain. IEEE Trans Instrum Meas 69(2):593–607
17. Gao X, Wang X, Zhou J (2020) A robust orthogonal matching pursuit based on L1 norm. In: Chinese Control And Decision Conference (CCDC), pp 3735–3740
18. Liu H, Zhang J, Mou C (2021) Image denoising based on correlation adaptive sparse modeling. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

# Solar Power Based Agriculture Robot for Pesticide Spraying, Grass Cutting and Seed Sowing


**Bysani Sai Yaswanth, N. Pruthvi Raj, B. P. Rahul, Venktesh M. Moger, and B. T. Venkatesh Murthy**

**Abstract** In India, most people are dependent on agriculture for their livelihood. Seed sowing, grass cutting and pesticide spraying are the various works that are carried out in the agriculture fields. Though there are many types of equipment available in the market, all are not used for multiple purposes. By developing a multipurpose agriculture robot input costs can be reduced. The proposed agricultural robot is an Internet of Things (IoT) based system. Through which users can monitor the condition of the crop and also perform some specific operation using an android app. For a non-technical farmer, it also helps to find a suitable crop based on some parameters related to crop using Machine learning. This paper intends to configure, create and manufacture the robot which carries out the sowing of seeds, cutting the grass and spraying of pesticides. This ensures the carrying out of agriculture operations in an efficient way.

**Keywords** Automation · Internet of Things · Machine learning · Microcontroller · Sensors · Solar panel

## 1 Introduction

Agriculture is the foundation of Indian culture and civilization. India positions second in the world in terms of farm output. A portion of the serious issues in the Indian

B. Sai Yaswanth · N. P. Raj · B. P. Rahul · V. M. Moger · B. T. V. Murthy (✉)
Department of ECE, Siddaganga Institute of Technology, Tumakuru 572103, Karnataka, India
e-mail: btv_murthy@sit.ac.in

B. Sai Yaswanth
e-mail: 1si17ec017@sit.ac.in

N. P. Raj
e-mail: 1si17ec075@sit.ac.in

B. P. Rahul
e-mail: 1si17ec076@sit.ac.in

V. M. Moger
e-mail: 1si17ec103@sit.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_73

farming are rising input costs, accessibility of skilled workers, availability of water. To avoid these issues, automation technologies are used. It helps farmers to reduce the human efforts and investments [1]. Developing a unique vehicle which can perform all these operations are more helpful to the farmers. In this paper a robot is developed to carryout operations automatically. The major idea of this paper is to propose an agriculture robot which performs sowing of seeds, cutting the grass, sprinkling of water, spraying of pesticides and prediction of the suitable crop based on climatic conditions.

Nowadays, farmers are spending more amount on machinery to reduce the labour. There is various equipment available in the market for carrying out agricultural operations. But they have to be operated manually to complete the task. The final product after using these equipments are less when compared to the invested money. This is good opportunity to address these issues by using the recent technologies.

There are so many Agribots that are manufactured. But all of them can not be used in a multipurpose way. In this paper, Agribot is developed using new technologies such as IoT and Machine Learning. These features make this robot unique and novel.

## 2 Related Works

In [1], the authors explained the process of recharging the battery, utilization of solar panel and controlling of robot using android app and Bluetooth module. It helps to overcome the problems encountered in conventional methods and carries out agriculture processes in an efficient way. Long range communication between the robot and user is not possible in this scenario. In [2], the authors discussed about the monitoring and control of field with the help of IoT. With this, condition of the field can be monitored from any part of world using Internet. They have developed two control systems. One is field control and the other is robot control. In field control, sensor data is shared to the farmer through android app. In robot control, robot sprays the pesticides to control the attack of pests. In this scenario, sensors are placed in fixed location. So, that data is inaccurate to estimate the condition of the field. In [3], the authors demonstrated the design of the seed sowing robot. Robot is able to sow the seeds of different sizes with high precision. It also ensures the precise spacing between the seeds. Wi-Fi acts as receiver to establish communication between the user and robot. Robot turns off automatically when there is obstacle in-front of it.

In [4], the authors clarify the distinctive complex strategies like planting of seeds, splashing pesticides alongside grass cutting. This arrangement of robot is fueled by solar energy where the solar panel gives the energy to the robot. By utilizing Bluetooth and android user gives the instructions to the robot to play out the necessary functions and the deployment of the robot. In this scenario ploughing operation is not done and robot works for short range.

In [5], the authors examined on planning an application that will permit farmers to foresee the local crops of explicit yields relying upon actual boundaries like precipitation and temperature. Utilizing crop data set of different harvests from different areas

of India, precipitation and temperature data set for same region, they have proposed a model which is utilized to foresee crop yield and gives farmers the suggestion set to upgrade their harvest determination dependent on factors like area, field size, temperature, precipitation, and different harvest data set [6]. The principle limitation of this model is that it is restricted to three crops.

After the thorough literature survey, the main objectives of the paper are:

1. To perform agricultural operations like seed planting, pesticide spraying and grass cutting.
2. To develop a system for Environment Monitoring using IoT.
3. To predict the suitable crop for cultivation, based on weather conditions using Machine Learning.
4. To develop an android application for robot control and mechanism using the Internet.
5. To share live video of the robot to monitor the field.

## 3 Methods

The block diagram of the proposed system is shown in the Fig. 1. It consists of Arduino mega, Solar Panel, Wi-Fi module, rain sensor, Soil moisture sensor, temperature sensor and pH sensor. The system is powered by using the battery, which is charged from the solar panel by using the solar power.

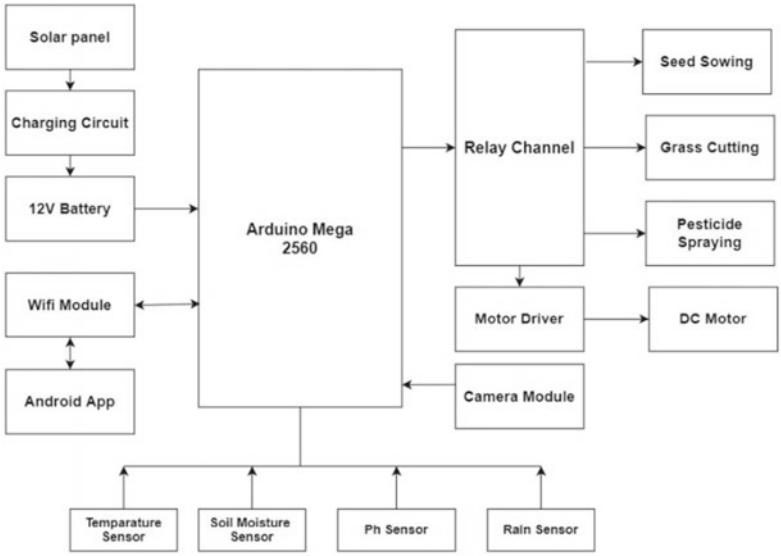



**Fig. 1** Block diagram of the proposed system

### 3.1 Power Supply

The Robot is powered by a 6 A h Lithium battery, which indeed is charged using the Solar Panel. A solar panel rated 20 W of 36 cells, can charge the 12 V battery in 4 to 5 h. Solar Panel provides a renewable and Ecofriendly source of energy. It is made up of Photovoltaic (PV) cells and works using the Photoelectric effect. The solar panel absorbs sunlight and converts this solar energy into electrical energy. This Electrical energy is used to charge the Lithium battery through a charging circuit [7–9].

### 3.2 Robot Structure

The agriculture robot is designed using Chlorinated Polyvinyl Chloride (CPVC) pipes of diameter 1.5 in. and a play-board to build a strong and lightweight body frame. The width of the frame is 30 cm and the length is 65 cm approximately. The funnels, solar panels, battery and other modules are deployed on that frame [10, 11].

### 3.3 Seed Sowing

Seed sowing is performed with the help of the two funnels. One funnel is used for the large size seeds and another funnel is used for small size seeds. These funnels are connected to a frame. The whole frame is controlled by the servo motor. When servo motor is turned on, the frames moves in the to and fro motion and seed gets dispensed [12].

### 3.4 Pesticide Spraying

Pesticide Spraying is used to protect the crop from harmful pests. It is performed with a 750 mL container and submersible pump. When the pump is turned on, another servo motor is activated and pesticide is sprayed in the field at an angle of 60° to 180° [13].

### 3.5 Grass Cutting

Grass Cutting is used to remove the unwanted grass in the field. This operation is performed with the help of two 5 V DC motors and blades. Time required is more

when the grass is thick. From above three sections agriculture operations can be performed in an efficient way.

### *3.6 Sensors*

Sensors records the physical properties of the field. So that user can monitor the field and take the immediate actions whenever it is required. Sensors involved in this paper are Temperature sensor, Soil Moisture Sensor, pH sensor and Rain sensor. When the robot is turned on sensor data is continuously shared to the cloud. To share the data to the cloud ESP32 WiFi module is module is used [14, 15].

### *3.7 ESP32-CAM Module*

ESP32-CAM is a low-cost development board with a WiFi camera. It allows creating an Internet Protocol (IP) camera for video streaming with different resolutions. It has a built-in PCB antenna. With the help of this module, the live broadcasting of the video is done. The users can monitor the fields and crop using an Android application.

### *3.8 Crop Prediction Using Machine Learning Algorithm*

The majority of the farmers think about which crops to be planted in a particular region dependent on climatic conditions. But for a non-specialized farmer, it is hard to recognize the appropriate yields for the particular region dependent on natural conditions. In this way, crop prediction assists the client with thinking about the harvests which can be planted in the specific area [16–19]. Crop prediction is performed utilizing the continuous sensor information in real-time and machine learning techniques. The decision tree algorithm is implemented in the classifier approach. In this application, available data need to be classified based on the parameters such as humidity, moisture, pH and rainfall to predict the output. Android Application is designed in such a way, whenever the user enters the parameters machine algorithm executes in the server and returns the predicted crop to the android application.

### *3.9 Android Application*

The Android application is developed using an MIT App Inventor, which is a web application development environment. The App is user friendly and developed mainly for the following applications:

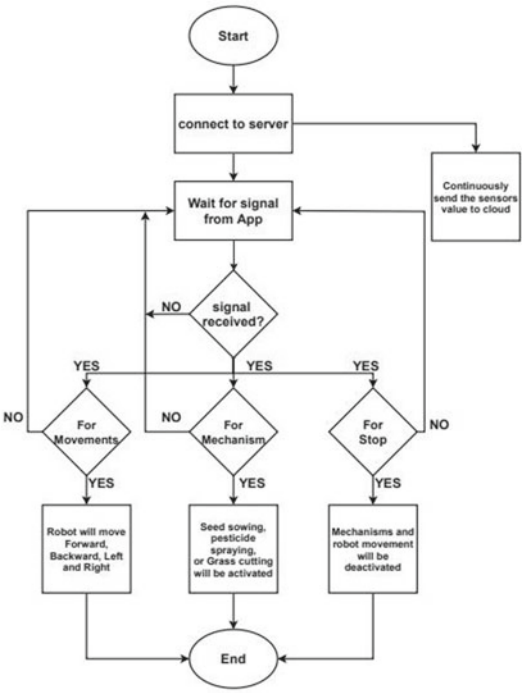



**Fig. 2** Flow chart of the proposed system

1. Controlling robot movements and operations.
2. Live video display.
3. To display the output of the sensors.
4. To view the predicted crop.

### 3.9.1 Working Principle

The following steps describe the work flow of the robot.

1. Making the setup as required.
2. Switching on the robot with the help of power supply.
3. Connect the robot and mobile application to the internet to establish the communication between mobile app and robot.
4. The sensor's values are continuously sent to the cloud database.
5. The robot should stand by until it gets a message from the application.
6. Robot will move left, right, forward and backward according to the signal received for movements thereby turning on the camera module.
7. If the message is received for any of the operations, then the respective operations part will be activated.
8. If the message is not received go back to step 4.

9. If user is required to stop the robot, then STOP signal is used to inactivate the movements.

The first step is to connect the app to the robot. The real time sensor values are sent continuously to the database in cloud. The robot waits for the message from the user through the app. When the message is received, the robot will check for the type of the signal. If the message is for the movements like left, right, front and back, then the robot will move accordingly and if the message is for activation of seed sowing, the robot will start sowing the seeds on the path. If the signal is for grass cutting, the robot will start to cut the grass, similarly for pesticide spraying and seed sowing. Proposed flowchart is shown in the Fig. 2.

## 4 Results

The Agribot is assembled as shown in the Fig. 3 and the various operations are performed in the field and results are tabulated as shown in Table 1.

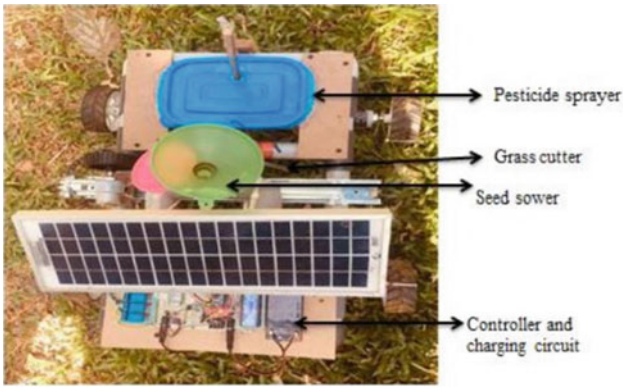



**Fig. 3** Agriculture robot

**Table 1** Case study on various operations w.r.t time

| Operation | Time taken for covering 10 m |
|---|---|
| Low grass cutting | 1 min 20 s |
| Medium grass cutting | 1 min 30 s |
| High grass cutting | 1 min 40 s |
| Loose soil ploughing | 1 min 40 s |
| Medium soil ploughing | 1 min 50 s |
| Hard soil ploughing | 2 min 00 s |
| Pesticide spraying | 1 min 20 s |

**Table 2** Power consumption of various operations

| Operation | Power Consumption for 1 h (in Watts) |
|---|---|
| Grass cutting | 6 |
| Ploughing | 1 |
| Pesticide spraying | 1.6 |
| Seed dispensing | 1.25 |
| Robot movement | 45–60 |

The grass cutting operation is performed on various type of grass. The approximate time required for performing grass cutting is around 1 min 30 s for 10 m using two blades. The Pesticide Spraying operation is performed at various angles such as 90, 60 and 180°. The approximate time required for performing pesticide spraying is around 1 min 20 s for 10 m using a submersible pump. The ploughing operation is performed on various type of soils. The approximate time required for performing the ploughing operation is around 1 min 20 s for 10 m using a ploughing rod. Designed robot is shown in the Fig. 3.

Grass cutting operation consumes around 6 W of power. 2.5 W from 5 V DC motor and there are 2 such motors adds up to 5 W. 1 W from stepper motor just to move up and down. Ploughing uses 1 W for the stepper motor to move the ploughing rods down. 1.6 W is consumed for pesticide sprayer which uses 5 V submerging pump. Seed dispensing consumes 1.25 W using SG90 servo motor. For the movements of robot it consumes 45–60 W depending on the friction of the surface. If obstacles is more (more friction of uneven road) the motor consumes more current simultaneously hence power consumption varies. The results are tabulated in Table 2.

The sowing mechanism is tested on the field using test seeds with different size of seeds. Approximate size of the seed is 1 cm in diameter. Seeds are dropped at a distance of 30 cm and with the help of the thick sheet soil is covered on the seed and robot move further. When all operations are activated a maximum of 10 A current is drawn. The batteries once charged powers the Agribot for about 1 h approximately depending on the soil conditions. ESP32-CAM which is mounted on the robot used to stream the video of the agriculture field using inbuilt WiFi. The output of the camera module is displayed in the android app as shown in Fig. 4. Sensors data is shown in the Fig. 5. If user wants to predict the suitable crop based on the different parameters, they can use the input textbox to provide the input to the algorithm, then algorithm predicts the suitable crop based on the Machine Learning model.

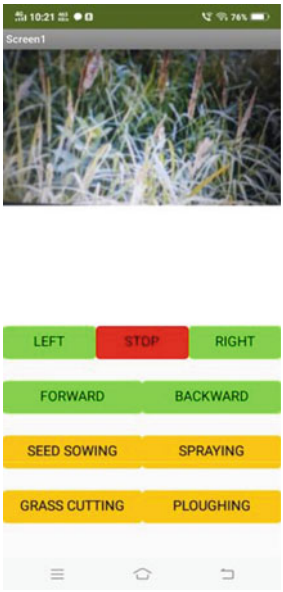



**Fig. 4** Camera module output


**Fig. 5** Crop prediction

## 5 Discussion and Conclusion

An autonomous multi-purpose agricultural robot is designed to play out the intricate cultivating undertakings like seed planting, grass cutting and pesticide spraying. The benefits of robot are reduced human intervention and efficient resources utilization such as the solar energy and also decreased human mediation and proficient assets use. Directions are passed to the system utilizing WiFi module, which guarantees remote controlling of robot with less intercession of human. The activities are performed utilizing android application. Creative seed planting, pesticide sprayer and grass cutting hardware has huge impact in agribusiness. For a non-technical farmer, Machine learning model helps to find the suitable crop for their fields. Camera module helps to monitor the field remotely. With the help of sensor data and IoT, farmer can monitor the environmental conditions of the field. By using the advanced technologies and automation in agriculture farmers can save more time and human efforts. The computational cost of the robot is around INR 12,000 and the working cost is INR 15,000.

## References

1. Ranjitha B, et al (2019) Solar Powered autonomous multipurpose agricultural robot using bluetooth/android app. In: 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), pp 872–877. IEEE
2. Jerosheja BR, Mythili C (2020) Solar Powered automated multi-tasking agricultural robot. In: 2020 International Conference on Innovative Trends in Information Technology (ICITIIT), Kottayam, India, pp 1–5. https://doi.org/10.1109/ICITIIT49094.2020.9071542
3. Ayaz M, Ammad-Uddin M, Sharif Z, Mansour A, Aggoune EH (2019) Internet-of-Things (IoT)-based smart agriculture: toward making the fields talk. IEEE Access 7:129551–129583. https://doi.org/10.1109/ACCESS.2019.2932609
4. Umarkar S, Karwankar A (2016) Automated seed sowing AGRIBOT using Arduino. In: 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, pp 1379–1383. https://doi.org/10.1109/ICCSP.2016.7754380
5. Rafath F, Rana S, Ahmed SZ, Begum JR, et al (2020) Obstacle detecting multifunctional AGRIBOT driven by solar power. In: 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI) (48184), Tirunelveli, India, pp 196–201. https://doi.org/10.1109/ICOEI48184.2020.9142983
6. Al-Mamun GMSU, et al (2019) Performance analysis of multi-purpose AGROBOT. In: 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), pp 1–4. https://doi.org/10.1109/WIECON-ECE48653.2019.9019906
7. Madiwalar S, Patil S, Meti S, Domanal N, Ugare K (2020) A survey on solar powered auto nomous multipurpose agricultural robot. In: 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 2020, pp 184–189. https://doi.org/10.1109/ICIMIA48430.2020.9074943
8. Mathad GP, Reddy S, Punith CM, Murthy BV (2018) IoT based mobile charging with solar energy by coin insertion. In: Proceedings of ISSRD International Conference, 24th June 2018, Bengaluru, India
9. Karthik S, Devaraj BN, Kashyap AK, Avinash M, Murthy BV (2019) Dynamic load shedding management for optimizing power distribution. In: 2019 International Conference on Communication and Electronics Systems (IC-CES), Coimbatore, India, 2019, pp 1719–1723

10. Pavan KU, Sahul MPV, Murthy BTV (2017) Implementation of stereo visual odometry estimation for ground vehicles. In: 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Tech-nology (RTEICT), Bangalore, 2017, pp 1173–1177
11. Coronado E, et al (2021) Towards a modular and distributed end-user development framework for human-robot interaction. IEEE Access 9:12675–12692. https://doi.org/10.1109/ACCESS.2021.3051605
12. Li S, Li S, Jin L (2020) The design and physical implementation of seeding robots in deserts. In: 2020 39th Chinese Control Conference (CCC), 2020, pp 3892–3897. https://doi.org/10.23919/CCC50068.2020.9189220
13. Tai ND, Thinh NT (2019) Design of cable measuring system of a robot spraying pesticides in agricultural farm. In: 2019 International Conference on System Science and Engineering (ICSSE), pp 577–580. https://doi.org/10.1109/ICSSE.2019.8823505
14. Ragavi B, Pavithra L, Sandhiyadevi P, Mohanapriya GK, Harikirubha S (2020) Smart agriculture with AI sensor by using AGROBOT. In: 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), pp 1–4. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00078
15. Maheswari R, Azath H, Sharmila P, Gnanamalar SS (2019) Smart village: solar based smart agriculture with IoT enabled for climatic change and fertilization of Soil. In: 2019 IEEE 5th International Conference on Mechatronics System and Robots (ICMSR), pp 102–105. https://doi.org/10.1109/ICMSR.2019.8835454
16. Yaswanth BS, Darshan RS, Pavan H, Srinivasa DB, Murthy BTV (2020) Smart safety and security solution for women using KNN Algorithm and IoT. In: 2020 Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT), pp 87–92
17. Taunk K, De S, Verma S, Swetapadma A (2019) A brief review of nearest neighbor algorithm for learning and classification. In: 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, pp 1255–1260
18. Meeradevi H, Salpekar H (2019) Design and implementation of mobile application for crop yield prediction using machine learning. In: 2019 Global Conference for Advancement in Technology (GCAT), BANGALURU, India, 2019, pp 1–6.https://doi.org/10.1109/GCAT47503.2019.8978315
19. Hasan M, Uddin KN, Sayeed A, Tasneem T (2021) Smart agriculture robotic system based on internet of things to boost crop production. In: 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pp 157–162. https://doi.org/10.1109/ICREST51555.2021.9331091

# Super Compact FR-4 Compatible 28 GHz Antenna for 5G Handheld Devices


**K. Jayanthi, D. Kumutha, and M. Jeyabharathi**

**Abstract** 5G is a very fast-growing technology with no limitation in real-time application. Millimeter-wave (mm-wave) band is a perfect choice to accomplish the bandwidth requirements and throughput for future 5G mobile applications. The frequency band deliberated are 28, 38, 60, and 73 GHz in 5G bands. A Super compact planar antenna for future 5G mobile communication is proposed which operates at 28 GHz. The prototype uses FR-4 substrate which has a dielectric constant of 4.4 and a loss tangent of 0.02. The antenna has a compact size of 10.87 × 12.364 $mm^2$ with a dielectric height of 1. 6 mm. The dimensions have been calculated using design equations and it is designed in HFSS 15.0 software. The patch is loaded with a U-shaped slot attached with the stub while the Ground of the antenna is not disturbed. Implementation of full ground is the perk of this project. The antenna resonates at 28 GHz, ranging between 26.32 to 30.26 GHz. The impedance bandwidth is obtained to be 3.94 GHz. The antenna parameters such as Return loss, Gain, Efficiency, and VSWR are measured and observed to be good. The antenna exhibits a very good return loss for 28 GHz at –32 dB and VSWR less than 2. The design complexity of proposed antenna is low since it consists of single layer. Generally defective ground structures are used to enhance the bandwidth and gain but without affecting ground structure the same can be achieved in the proposed design.

**Keywords** 5G communications · FR-4 · Millimeter wave (mmW) · MIMO · VSWR

## 1 Introduction

Nowadays, the Millimeter wave (MMWs) is predicted significantly for developing in the Fifth Generation (5G). Mostly, 5G is carried out with the many features that

K. Jayanthi · M. Jeyabharathi
Government College of Engineering, Salem, Tamil Nadu, India

D. Kumutha (✉)
SJB Institute of Technology, Kengeri, Bangalore, India
e-mail: kumuthad@sjbit.edu.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_74

are adopted with better network capacity, higher throughput and the speed of the data rates is efficient. On the other side, Millimeter Wave (mmW) is incorporating very efficiently in the real-time 5G application. During mobile communication, the MIMO antenna is inbuilt to increases the channel capacity for the Millimeter waves of 28 GHz for better performance. It is more suitable for the mm waves to analyze the antenna system design with 28 GHz for any application.

## 2 Related Work

A MIMO Antenna array with 4 number of 2-element arrays located at the center of the substrate, which is excited using power divider/combiner concept, radiates from 25.5 to 29.6 GHz. The DGS structure carved with rectangular, circular, zig-zag shaped slots produce better impedance bandwidth and increases peak gain and efficiency which is observed to be as 8.45 dB and 82% respectively [1–3]. Two circles with opposite directions having radius and length as 6.8, 8.9 (inner circle), and 3.6, 11.1 mm (outer circle) respectively are constituting to 32.5 and 26 GHz. The ground structure being in circular shape reduces the inductive and capacitive effect causes the antenna to operate from 24 to 39 GHz for mm-wave [4]. A single bowtie-shaped radiator converted into a 2-element antenna connected to the patch with ground having a slit introduced between the two bowties-shaped slot helps in achieving maximum isolation and causes resonance from 27 to 29.02 GHz with gain as 12 dB [5]. The radiation patch and a ground plane are integrated using AgHT-8 achieve the frequency range of 23.92 to 43.8 GHz at 58.71% of impedance bandwidth in optically transparent antenna [6].A rectangular-shaped slotted patch antenna is a 2-element array and it is based on the feed network for the T-junction power combiner and divider for the required array element, the ground plane structure of rectangular, circular, and Zigzag-shaped slot. The prototype operates between 25.5 to 29.6 GHz with an 8.3 dB gain [7]. A compact tree-shaped quad element MIMO is proposed with 4 different arcs to achieve a wide band from 23 to 40 GHz with resonance at 28, 33, and 38 GHz. The isolation is greater than 20db with the total gain being 8.87 dB and efficiency is above 70% [8]. A Dual-band antenna operating at 28 and 38 GHz is presented which yields 7.7 and 7.73 dB gain for 28 and 38 GHz respectively. Further $2 \times 2$ and $1 \times 4$ antenna arrays are designed for which the gain is observed to be 12.5 and 12.1 dB for the dual-band [9]. The design of the Multibaryonic H-Shaped coupled microstrip antenna is illustrated and proposed for 5G applications [10]. A shape of the Dual-band patch is in an elliptical slot resonating at 28 and 45 GHz having a good return loss at –40 and 7.6db gain for 28 GHz. Then the maximum gain occurs at 7.21 dB for the return loss of $-14$ dB at 45 GHz. The center series fed with quarter-wave feed techniques are achieved for the array configuration with a 12 dB gain criterion. Following, 28 GHz is designed for the array antenna for yielding 13.5 dB gain and efficiency of 98.75% is 1.19 W/kg, whereas the array antenna for 34 GHz is 1.16 W/kg and same for 45 GHz is 1.2 W/kg [11, 12]. For secure and high-speed transmission an ultra-wideband antenna covering 2.3 to 13.5 GHz is

proposed which meets the 10 db criterion for return loss and exhibits a maximum gain of 11 db [13]. Various Mutual coupling techniques, and MIMO parameters have been incurred in [11–16].

Compared with all the above-mentioned prototypes the proposed system differs from the material used, implementation of full ground and the less complexity of the design. In this paper, a design of super compact radiating system which adapts to FR-4 substrate even under higher frequencies is proposed for 5G mobile application. The proposed prototype is a simple rectangular patch antenna having a stub and U-Shaped slot in it. The radiator has enhanced gain with unidirectional radiation pattern since the ground plane is un-defected. The antenna exhibits good voltage standing wave ratio (VSWR), high gain, uniform radiation pattern and almost constant radiation efficiency over a wide range of frequencies. Section 3 of this paper presents the system design of the antenna which has the design and its evolution in stages. The simulated results are reported in Sect. 4 followed by the conclusion in Sect. 5.

## 3 System Design

A Compact rectangular planar antenna is presented, whose substrate size is 10.87 × 12.364 $mm^2$ and the patch size is 1.270 × 2.764 $mm^2$. HFSS software deployed for designing the antenna. FR-4 Substrate is used whose height is 0.8 mm, dielectric constant and loss tangent of $\varepsilon r = 4.4$ & $\tan\delta = 0.02$. The proposed radiator is designed and implemented using Ansoft HFSS15.0. The feed length is obtained as 1.81 mm. The dimension of the ground and substrate is same because the ground is not disturbed which is the perk of this prototype. The size of the antenna is much smaller than all other radiators stated in the previous literature. These dimensions have been incurred from the design equation which is as follows,

The wavelength ($\lambda$) is designed for resonant frequency (f) = 28 GHz and velocity of light C = (3 × $10^8$ m/s) to obtain 10.7 mm, formula is given in the Eq. (1).

$$\lambda = \frac{c}{f} = \frac{3 \times 10^8}{28 \times 10^9} = 10.7\text{mm} \tag{1}$$

Width of the Patch is designed using Eq. (2) for $\varepsilon_r = 4.4$,

$$W = \frac{c}{2f}\sqrt{\frac{2}{\varepsilon_r + 1}} = \frac{3 \times 10^8}{2 \times 28 \times 10^9} \times \sqrt{\frac{2}{4.4 + 1}} = 2.764\,\text{mm} \tag{2}$$

where C = Speed of Light = = (3 × $10^{11}$) mm, f = 28 GHz, $\varepsilon_r$ = Dielectric constant of the substrate = 4.4.

Now, length of the patch (L) is calculated using Eq. (3) and the effective length ($L_{eff}$) is calculated using Eq. (4) which is given below

$$L = L_{eff} - 2\Delta L = 1.270\,\text{mm}\,. \tag{3}$$

$$L_{eff} = \frac{c}{2f\sqrt{\varepsilon_{\text{reff}}}} = \frac{3 \times 10^8}{2 \times 28 \times 10^8\sqrt{1.565}} \tag{4}$$

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{\text{h}}{w}\right]^{\frac{-1}{2}} = 1.54$$

where, $\varepsilon_{reff}$ = Effective dielectric constant of substrate,

Then the Substrate length (L) and substrate Width (W) are calculated using the below formulas in the Eq. (5) and (6),

$$L_g = \text{L} + 6\,\text{h} = 1.270 \times 10^{-3} + 6 \times 1.6 \times 10^{-3} = 10.87\,\text{mm} \tag{5}$$

$$W_g = \text{W} + 6\text{h} = 12.36\ \text{mm} \tag{6}$$

where $\text{L}_g$ and $\text{W}_g$ are length and width of substrate, and 'h' value is given 1.6 mm.

Feed line length is calculated using the below Eq. (7).

$$\text{Feed length (Fl)} = \lambda_{\text{g}}/4 = 7.2672/4 = 1.81\,\text{mm} \tag{7}$$

where $\lambda_{\text{g}}$ is guided wavelength and it is given by, $\lambda\text{g} = \lambda/\sqrt{\varepsilon_{reff}} = 7.2672\,\text{mm}$ (Table 1).

The front view and rear view of the antenna dimension are shown in Figs. (1) and (2). Table [1] give the designed value of the Compact rectangular planar antenna to manipulate at the 28 GHz. Initially, a simple microstrip antenna was designed

**Table 1** Dimension of the antenna

| Key word | Size [mm] | Key word | Size [mm] |
|---|---|---|---|
| Ws | 12.36 | Pw | 2.764 |
| Ls | 10.87 | Pl | 1.27 |
| Fl | 1.81 | Lg | 10.87 |
| Fw | 0.5 | Wg | 12.36 |

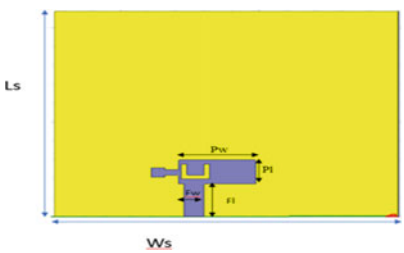



**Fig. 1** Front view of the antenna

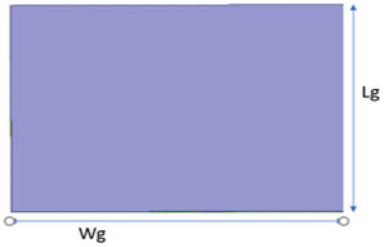



**Fig. 2** Rear view of the antenna

using the above-mentioned dimensions with feed at the center of the patch. The antenna exhibited very poor return loss. As an initial change, the feed was moved to the left at the distance of $\lambda/6$ from the center for which the antenna resonated at 26.8 GHz covering a frequency band of 25.6 to 28.73 GHz thus showing an impedance bandwidth of 2.33 GHz. In order to obtain the resonance at 28 GHz, further changes have been made. A U- shaped slot has been etched from the patch which caused a change in return loss which increased the bandwidth. The operating frequency was observed to be between 25.1 to 29.5 GHz below –10db. The resonance occurred at 26.5 GHz at –15db. In order to improve the return loss and to achieve resonance at 28 GHz, a stub was included at the left side of the patch. This inclusion caused a change causing the radiator to operate between 26.3 to 30.27 GHz and resonance at 28 GHz at –31 dB. The inclusion of stub finally achieved the resonance at 28 GHz with a good return loss which is the ultimate aim of the project. VSWR is observed to be less than 2 in the entire operating band and other antenna parameters like gain, efficiency is observed to be good. In all the reference papers the result is obtained only after disturbing the ground which causes backward radiation which ultimately produce undesirable radiation on mobile phones. Thus, this prototype exhibits a better result with full ground which makes this prototype ideal among all the proposed systems in the reference.

Some of the salient features of the proposed single element and modified corporate feed array structures are enumerated below: (i) Generally, defected ground plane and multilayered configurations are being used to enhance the bandwidth and gain but in the proposed configuration, a simple single layered with un-defected ground structure is employed to achieve the same. (ii) The proposed antenna has low design complexity since it consists of single layer and with a minimum number of slots. (iii) Generally, a substrate which has low dielectric constant is used for higher frequencies. But, the protype is designed using FR-4 substrate whose dielectric constant is 4.4 which is higher than the dielectric constant of the substrates used in the literature [1–16]. (iv). The presented antenna offers good gain over an entire frequency spectrum 5.77 and 15.5 dB are the single and array configurations average gain compared with all other radiators operating in mm-wave spectrum. (v). Moreover, the enhanced frequency response is achieved using a stub unlike DGS [1–5], power divider/combiner [7], quarter-wave feed [11, 12], Complicated structures [16] are discussed in literature.

## 4 Results and Discussion

The return loss, VSWR, realized gain, frequency vs efficiency results of proposed antenna are discussed as follows.

### *4.1 Return Loss*

The ratio of incident power of an antenna to its reflected power is nothing but return loss. It is measured in terms of dB and represented by S11. The return loss below –10 dB is taken as the operating band of the antenna. The return loss (S11) of the proposed antenna is as shown in Fig. 3. This antenna covers the band from 26.32 to 30.26 GHz with the impedance bandwidth of 3.94 GHz. The proposed antenna has a return loss of –31.6 dB at 28 GHz.

### *4.2 Voltage Standing Wave Ratio*

Voltage standing wave ratio is defined as the ratio between maximum amplitude of standing wave to its minimum amplitude. It is the measure of how much power is coupled to the load from source travels through transmission line. For better performance the VSWR should be less than or equals to two. From Fig. 4 it is noted that

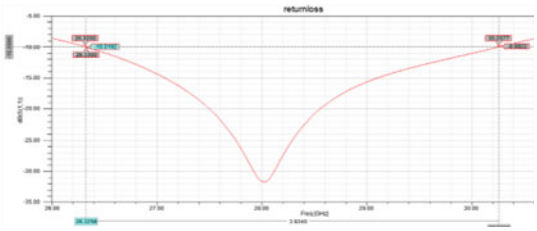



**Fig. 3** Simulated result of the return loss

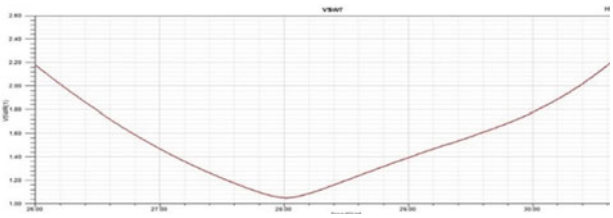



**Fig. 4** Simulated results of VSWR

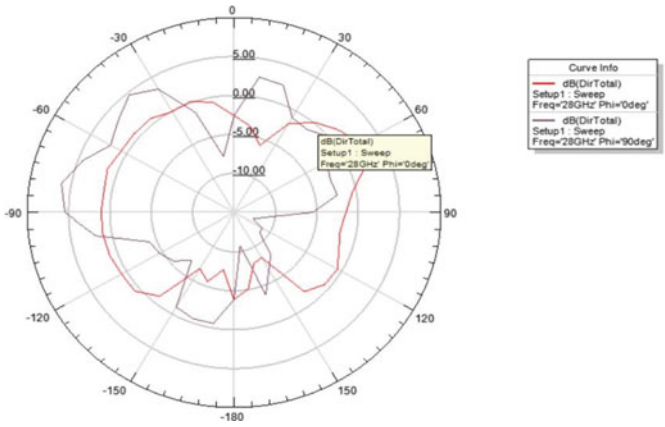



**Fig. 5** Radiation pattern of the antenna

the VSWR is less than two for the entire operating band from 26.32 to 30.26 GHz and it takes the value of 1.05 at 28 GHz.

## *4.3 Radiation Pattern*

The radiation pattern at 28 GHz is shown in Fig. 5. The red color line represents E-plane (XZ, phi = 0) and blue line represents H plane (YZ, phi = 90).

## *4.4 Radiation Efficiency*

The ratio between radiated power of the antenna to the incident power is defined as antenna radiation efficiency. In mathematical form,

$$\text{Efficiency}(\eta) = \text{Radiated Power (Prad)}/\text{Input power}(P \text{ in }/\text{out}) \quad (8)$$

In the designed prototype the efficiency is observed to be more than 75% at 28 GHz which presented in Fig. 6.

### *4.5 Realized Gain*

The total efficiency of the antenna, along with its directivity is used to calculate the realized gain of an antenna. The realized gain for the designed prototype is observed to be 5 db at 28 GHz as given in Fig. 7.

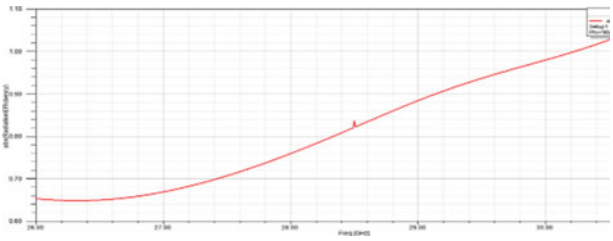


**Fig. 6** Radiation efficiency

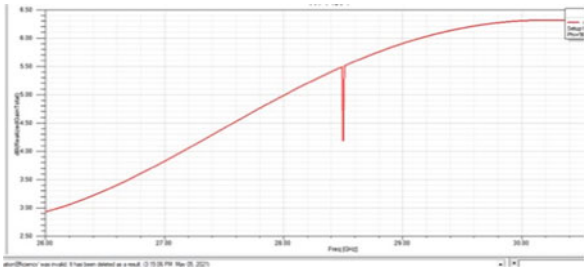


**Fig. 7** Realized gain

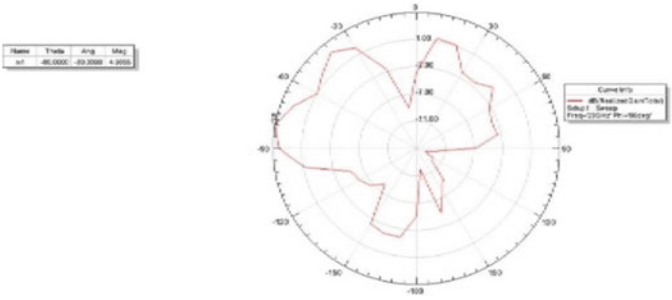


**Fig. 8** Gain

### *4.6 Gain-2D and 3D Pattern*

The ratio between radiation intensity of the designed antenna in desired direction to the isotropic antenna radiation intensity is called antenna gain. By definition the directivity and gain of the antenna are closely related but there may be a simple difference. In mathematical form it is given by Eq. (8).

$$\text{Gain(G)} = (\text{Efficiency}(\eta)) * (\text{Directivity(D)}) \tag{9}$$

Gain-2D and Gain-3D of the proposed super compact planar antenna is shown in Fig. (8) and (9). The 3D gain pattern of the proposed antenna system obtained at 28 GHz as in Fig. 9. The antenna achieves a maximum gain of 4.99 dB at the operating band. Thus, the prototype is designed in such a way that it stands out uniquely as it is implemented with full ground. Also, the antenna uses FR-4 substrate which is cost effective and has a dielectric constant 4.4. The works that are given in the literature have used substrate materials that have high dielectric constant which helps in obtaining good performance of the antenna at higher frequencies. But this paper has presented an antenna which exhibits a good return loss less than –30 db, VSWR less than 2 over the entire band and gain of 4.99 dB (Table 2).

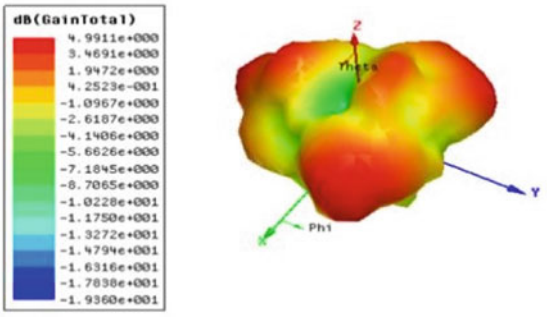



**Fig. 9.** 3D gain pattern

**Table 2** Comparison table

| Ref. no | Substrate | Operating frequency (GHZ) | Band-width (GHZ) | No of elements | Gain [dBi] | Efficiency (%) | Double sided radiation |
|---|---|---|---|---|---|---|---|
| [13] | AgHT-8 | 23.92–43.8 | 19.8 | 2 | 1.4 | 58.71 | Yes |
| [16] | Rogers-5880 | 23 to 40 | 17 | 4 | 8.87 | 70 | No |
| [7] | RT/Duroid 5880 | 25.5 to 29.6 GHz | 4.1 | 2 | 8.3 | 71 | Yes |
| [Proposed] | FR-4 | 28.32 to 30.26 | 3.94 | 1 | 5 | 75 | No |

The comparison table exhibits the novelty of the design comparing with papers that has been explained in literature. With only single element being implemented the prototype exhibits good results.

## 5 Conclusion

Thus, a Super compact planar antenna for future 5G mobile communication is designed which operates at 28 GHz. The prototype uses FR-4 substrate which has a dielectric constant of 4.4 and loss tangent of 0.02 with size of 10.87 × 12.364 $mm^2$ and dielectric height as 1.6 mm. The patch is loaded with a U-shaped slot attached with the stub while the Ground of the antenna is not disturbed. The antenna resonates at 28 GHz, ranging between 26.32 to 30.26 GHz. The impedance bandwidth is obtained to be 3.94 GHz. The antenna exhibits very good return loss for 28 GHz at –32 dB and VSWR is 1.05. The proposed antenna achieves a maximum gain and efficiency of 4.99 dB, 75% respectively. The proposed antenna is suitable as Array/MIMO configuration for base station and as well as for 5G enabled handheld devices.

## References

1. Hussain R et al (2017) Compact 4G MIMO antenna integrated with a 5G array for current and future mobile handsets. IET Microw Antennas Propag 11(2):271–279
2. Jilani SF, Alomainy A (2018) Millimeter-wave T-shaped MIMO antenna with defected ground structures for 5G cellular networks. IET Microw Antennas Propag 12(5):672–677
3. Kumutha D, Arulmozhi S, Meena K, Kumar TS, Madhumitha K, Aswini K (2019) A novel broadband and high-isolation dual polarized microstrip antenna for 5G application. In: 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), pp 1–6. IEEE
4. Khalid M et al (2020) 4-Port MIMO antenna with defected ground structure for 5G millimeter wave applications. Electronics 9(1):71
5. 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
6. Sharma S, Kanaujia BK, Khandelwal MK (2020) Analysis and design of single and dual element bowtie microstrip antenna embedded with planar long wire for 5G wireless applications. Microw Opt Technol Lett 62(3):1281–1290
7. Desai A, Upadhyaya T, Pate R (2018) Compact wideband transparent antenna for 5G communication systems. Microw Opt Technol Lett 2018:1–6
8. Khalid M, Iffat Naqvi S (2020) 4-Port MIMO antenna with defected ground structure for 5G Millimeter wave applications. Electronics 9:71
9. Sehrai DA, Abdullah M (2020) A novel high gain wideband MIMO antenna for 5G Millimeter wave applications. Electronics 9:1031–2020
10. Khan J, Sehrai DA, Ali U (2019) Design of dual band 5G antenna array with SAR analysis for future mobile handsets. J Electr Eng Technol
11. Khattak MI, Sohail A (2019) Elliptical slot circular patch antenna array with dual band behaviour for future 5g mobile communication. Prog Electromagn Res C 89:133–147

12. Jeyabharathi M, Jayanthi K (2020) Survey on planar antennas For 5G application. Int Res J Eng Technol 7(12):319–325
13. Kumutha D, Prabha NA (2021) Design of multibaryonic H-shaped coupled microstrip antenna for Wi-Fi, WiMAX and 5G applications. In: Komanapalli VLN, Sivakumaran N, Hampannavar S (eds.) Advances in Automation, Signal Processing, Instrumentation, and Control. LNEE, vol 700, pp 1341–1351. Springer, Singapore. https://doi.org/10.1007/978-981-15-8221-9_125
14. Ghazaoui Y, El Alami A, El Ghzaoui M, Das S, Barad D, Mohapatra S (2020) Millimeter wave antenna with enhanced bandwidth for 5G wireless application. J Instrum 15(01):T01003
15. Naqvi SI et al (2019) An integrated antenna system for 4G and millimeter-wave 5G future handheld devices. IEEE Access 7:116555–116566
16. Marzouk HM, Ahmed MI, Shaalan AHAM (2019) Novel dual-band 28/38 GHz MIMO antennas for 5G mobile applications. Prog Electromagn Res 93:103–117

# Survey on Software Solution for High Performance Packet Processing


**Nanda Kishore, S. Rajarajeswari, Pramod Sunagar, and Anita Kanavalli**

**Abstract** With the evolution of the internet, we see a huge spike in the number of network users. The range of generated internet traffic in most of the various locations is increasing rapidly. The latest networking environment should deliver high throughput, high speed, high bandwidth, and fewer delay properties. Else it leads to data loss which is costly in packet processing. This survey contributes to finding a faster software solution to packet processing framework to overcome the challenges faced by traditional networks. This paper aims at packet acquisition and distribution methods based on packet processing models such as Data Plane Development Kit (DPDK), Netmap, Netslice, and PF-RING. And come up with a model which can productively decrease packet loss, successively enhance the performance rate and reduce resource waste. Comparison between the packet processing frameworks depicts that DPDK based data processing has dominance over Netmap, Netslice, and PF-RING in packet processing.

**Keywords** Packet processing · DPDK · Netslice · Netmap · Pf_ring

## 1 Introduction

Packet processing is a collection of steps and algorithms which are performed on an incoming or outgoing packet of information or data. Traditional network traffic packet processing is facing acute efficiency and performance fall due to platform hardware

N. Kishore (✉) · S. Rajarajeswari · P. Sunagar · A. Kanavalli
Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore 560054, India
e-mail: nandakishore.jnv@gmail.com

S. Rajarajeswari
e-mail: raji@msrit.edu

P. Sunagar
e-mail: pramods@msrit.edu

A. Kanavalli
e-mail: anithak@msrit.edu



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_75

limitations. One more cause of performance bottleneck is the kernel network protocol stack performance.

Network communication methods have increased their performance. However, it does not guarantee high packet processing at a high packet arrival rate [1]. There are restrictions on the construction of network stacks that must be used with standard hard- ware. Package performance can work well by switching network stacks and without the need for very high-performance hardware. This led to research on how to integrate hardware and high-paced communication systems towards package performance. The rest of the paper is displayed as, Sect. 2 related work. Section 3 explains overview. Section 4 surveys about software solutions to packet processing. Section 5 future work. Section 6 compares techniques surveyed and ends with the conclusion and future work.

## 2 Related Work

This study put forward an I/O stack that has both the plus points of zero-copy and the use of the standard SSD page cache. In the I/O stack, the page archive applies a pre-program read. When you miss, the device will send the data directly to the user [2]. This paper unveils an examination of performance on network interface cards and package processing techniques. The assessment targets to assist network managers and experts determine which network card and technique must be adapted based on the network needs [3]. This paper suggests how to receive data based on zero-copy. When zero-copy pool resources are depleted during packet arrival at NIC, a standing resource is utilized to ensure normal data acceptance and to lessen the rate of packet loss [4].

This work put forward a design in which everything is decentralized, and operations are distributed accordingly. The packet processing functionality is taken care of with micro service control. This framework has a user-space network driver called ixy [5]. Kim has implemented a technique in which it is used in one-way communication using a modified commercial-off-the-shelf network card driver. It observed reliability and performance using PF-RING with zero-copy [6]. This paper offers two operations to make good use of the bus between the FPGA and the CPU. They evaluated the range of real-world CPU + FPGA-based development platforms and the network function virtualization software. The results showed that the system, which can open up to $2 \times 40$ GbE traffic [7].

## 3 Overview and Drawbacks in Linux Packet Processing Network Stack

In Fig. 1, it depicts the smart packet processing. Instead of completely depending on tradition way of processing the packets using Linux kernel stack. With the help of

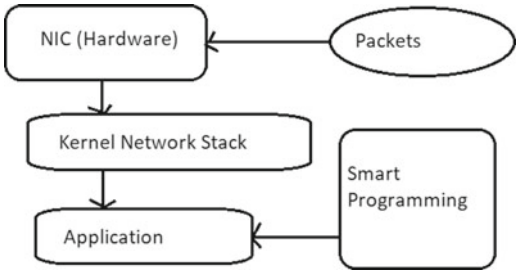



**Fig. 1** Packet processing with smart programming

smart programming and techniques, when an application running in user space, wants to receive the packets, it is first received by the network interface card, later process the packets using libraries and APIs to make best use of hardware and improve rate of packet processing capabilities.

## *3.1 Packet Processing in Linux Network Stack: Main Stages*

When a network interface card (NIC) first receives a data packet, it transmits it to a circular receiving queue (RX), which are also called rings. The data structure, the receiver descriptor, helps in holding those data packets, till the data packets are copied from NIC to the main memory. The data transfer from NIC to main memory is accomplished through the Direct Memory Access (DMA) operations, without involving the CPU. Thereafter, there needs a mechanism that can help in notifying the system about the new data packet received and pass every data packet onto a specially allocated buffer called the sk_buff struct. This data structure is assigned for each data packet and set off free when a packet enters from kernel space to user space. The disadvantage of this procedure is it consumes a lot of bus cycles. Another problem with the sk_buff structs is, it creates overhead as it is designed to contain metadata for all the protocols so that it can be compatible with as many protocols as possible. But that's simply not a requisite for processing specific data packets. Because of this additional struct, processing performance is brought down. Context switching is one more factor that negatively affects performance because it significantly consumes system resources [8].

# 4 A Brief Discussion of Techniques for Faster Packet Processing

## *4.1 Netmap*

Netmap is one such technology, based on zero-copy, which provides packet processing at higher speeds by reducing processing slow down costs. Netmap is built on existing OS features and characteristics, which are independent of hardware and few devices. In Fig. 2 depicts the Netmap framework. The data structures present in Netmap are Netmap ring and Netmap packet buffers.

**Netmap Packet Buffers.** The Netmap packet buffers are pre-allocated and with predefined size. This helps in reducing the cost per packet acquisition and distribution. These data structures are present in shared memory regions to reduce the zero-copy between kernel to user-space.

**Netmap Ring.** The Netmap rings are circular queues that possess buffer-related metadata, which is alike to Network card Rings or NIC rings. The metadata carries information related to available buffers, count of slots, that a ring can hold.

**Netmap_if.** It is a data structure that possesses all the interface-related info, like the number of rings present.

When a network card runs in this Netmap specified approach, the network card will uncouple from the protocol stack provided by the host. A Netmap ring also known as the NIC ring is created by the Netmap, which helps in copying. Netmap will create two pairs of NIC rings or Netmap rings to communicate between the host protocol stack. The network card directly moves the data packets onto rings referred to cache use, in the shared space. The application running in user space

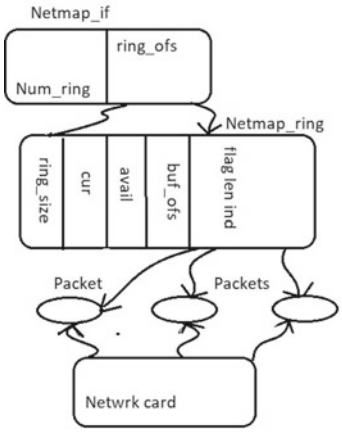



**Fig. 2** Netmap data structure and its components include Netmap_if, Netmap_ring and NIC

will call a Netmap API, to access the contents of data present in the NIC ring or Netmap ring can straightaway read the cache and start writing the data packet, in turn, helps in less copy between kernel space to user space and provides zero-copy functionality. Key memory regions of the network card register and the kernel to applications are not shared, hence there is no kernel crash caused by the application running in user space. Remarkable packet rate can be achieved based on the Netmap applications, which incorporate huge pages. With modifications to global variables netmap.nr_huge_page, huge__page, the Netmap application gets pool of memory with memory of huge pages. With 64B packet size, the throughput increases by 37%, from 15.714 to 21.557 Mpps on two ports which support up to 10 Gbits/s. Whereas with on four ports, throughput increases from 17.672 Mpps to 20.571, making it to 16.3% [9].

### *4.2 NetSlice*

Contrary to the fast packet solutions introduced in this survey, this method, NetSlice is not taking advantage of zero-copy methods. It is an OS abstraction, which provides faster package performance and works in the user-land. It attempts to synchronize the advantages of package processing systems operating in the user-land, e.g. the separation of error and configuration. NetSlice deploys on the advanced integration of components of hardware and software associated with packet processing. It is based on local architecture, rather than the temporary fragmentation of components of computer soft- ware and hardware suitable for package processing, namely memory, CPU cores, and NICs. By making such a distinction, NetSlice reduces the conflict of resources shared. The working idea of NetSlice is its state of action, called NetSlice. NetSlice uses several multi-line NICs. To support the same practice for many believers, multi-line NICs maintain more than one transmission and receipt queue.

The context of making NetSlice consists of completely separated resources. Even the NICs and the CPU cores are treated as separate resources. At least two CPU cores contribute to one NetSlice. The kernel and user mode operations are done by dividing CPU cores into u-peer and k-peer, which are NetSlice output states. The u-peer CPU helps perform the user mode functionality. Whereas k-peer CPU is used for the in- kernel network stack. The k-peer helps detect contextual cables and also summaries the required NetSlices. The NetSlice helps decide which data packets to flow from NICs to user-land and from user-land to NICs. The packets received from NICs are lightly processed on k-peer cores and forwarded to user-land for application use and later using pipe, packets are processed. NetSlice uses socket's standard APIs to read, write and vote different streams of data for other NetSlices. Using ioctl format NetSlice expands the API. NetSlice uses encryption and hence reduced delays in system calls and the number of system calls. Extended APIs can be used for setting up the system calls. Batch processing helps reduce

the overhead caused by each packet. NetSlice is not taking the privilege of zero-copy operations. It copies from kernel-space to user-space. Establishing zero-copy also improves NetSlice performance. However, it is not included in the frame, as it will limit the load. On Dell PowerEdge R900, scaling the number of CPU cores and with application Ipsec, NetSlice can reduce shared resources and effectively improve the 10 Gbps network link speeds. Kernel throughput reaches 7.591 Gbps, whereas NetSlice attains 76% more than Kernel throughput [10].

### *4.3 PF_RING*

PF_RING is one more such framework that implements a zero-copy method between the user-space and the kernel-space. It was invented by Luca. It provides a fast network packet acquisition and distribution framework. It attains fast packet processing with the help of PF_RING buffers, which are present in the shared memory region which is common to user space and kernel space. These PF_RING buffers are pre-allocated, which helps in reducing the cost per network packet memory assignment and un-assignment.

The main elements present in PF_Ring's framework are:

1. *PF_Ring user-space library.* This helps in accessing the PF_RING module present in the kernel from applications running in user-space.
2. *kernel module.* This helps in copying data from the network card to PF_Ring circular queue.
3. *PF_Ring aware drivers.* This helps in increasing the performance by using exclusive drivers to access the network cards.

The PF_RING uses exclusive device drivers, called Direct NIC access or DNA, for gaining fast network data packet processing without the involvement of a central processing unit or CPU, by using system calls. PF_RING uses mapping from NIC memory to user-space memory and helps transmission between network card and applications running in user space. The network data packets are sent from the network card to the userland, without the operations of the PF_RING module or Linux kernel. Then zero-copy happens by doing Direct memory access from the NIC Network process unit and kernel data packet buffer is copied. With using PF_Ring a high-performance VPN based SSL shows up to 30% improvement in throughput up to 500 Mbps with 1500 b packet size, network latency to 9 ms and download rate to 1.5 Mb/s for a file size of 6 Gb [11].

### *4.4 DPDK*

This framework in Fig. 3, provides a set of libraries, developed by Intel under the open source to accelerate packet processing workloads. The DPDK framework is

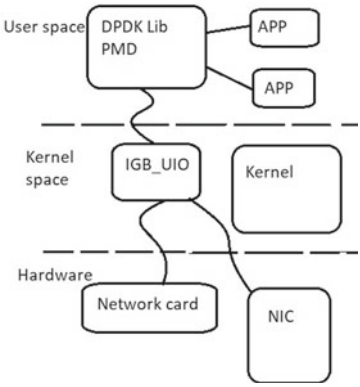



**Fig. 3** DPDK architecture which help bypass kernel stack with the help of IGB_UIO module and DPDK libraries

different from the Linux kernel network stack. The DPDK framework works excellently in user space with the help of IGB_UIO by not entering into the Linux kernel network protocol stack to decrease information copying and make full dominance over provided hardware. The applications run at user space and they use DPDK set of libraries and poll mode drivers at user space to acquire and distribute data packets, when new data packets enter network card in the system, the new data packets are transferred directly from the network card to the applications running in user space without using Linux network stack.

DPDK implements either in the run to completion model or pipeline model. With the system with many processing cores. The idea in the run to completion model is to poll the network card and process packets and transmit from and to the same network ports, attached to the cores. In the run to completion model, all the cores work to complete a common application code. Whereas in the pipeline model, only one core is used for receiving and transmitting the data packets from the network card. It makes use of the ring queue from Librte_ring to pass this data to other cores to process the packets and in the pipeline model, cores can work on different application codes. With 2 CPU cores to application running on host and unsupported 1 Gb/s network card, the line rate can be achieved with packet size of 128 B [12].

**DPDK Strategies for High Performance Packet Processing**

1. It uses pre-supplied memory buffers called mbuf. It stores both meta-data as well as the actual data. And it requires only one allocation per packet.
2. Provides cores with their cache memory and helps reduce the access to the shared pool ring-memory, which adds more efficiency with the CPU.
3. It uses circular rings, used as a queue. It has more advantages than the linked list. Besides, it reduces the time required to do more time-consuming tasks.
4. Reduces further disruption by using Poll Mode Drivers [13].

## 5 RSS Algorithm on DPDK Based Packet Processing

RSS (Receive Side Scaling) is a driver-based strategy that can productively pass on data across multi-core platforms after acquiring messages under a multi-processor environment. In Fig. 4, the NIC examines the data to find five tuples of IP add, protocol, and associate port. Then the line is regulated using the combined estimate from these 5 instances utilizing the RSS Hashing functions. Thereafter the output will be put in mbuf data structures to eliminate duplicate computation in performance such as queue line computation. The RSS algorithm utilizes the Toeplitz-hash based algorithm, where the 5 fields are Sender IP, Destination IP, Sender port, Destination port, and key to calculate the hash effect. The algorithm inputs the 5 Tuples with a random k value. It traverses each bit and XOR the result if the found bit is 1 else result is unchanged. Then the key is moved to the left. Then the algorithm will separate the incoming traffic into specific lines. The pseudo-code of the algorithm is as follows.

**Algorithm 1. RSS pseudo-code**

```
For hash-input input[] of length N bytes (8N bits) and a random secret key
K of 320 bits
Result = 0;
For each bit b in input[]{
if (b==1) then Result = (left-most 32 bits of K);
shift K left 1 bit position; }
```

### *5.1 Flow Distribution Module*

1. Call rte_eal_init() to initialize ports, huge-pages and build environmental-specific libraries.
2. Using rte-en-rx-burst() function, NIC acquires the packets, hash operations are ap- plied based on the 5 tuple values and cached in MBUF structure.

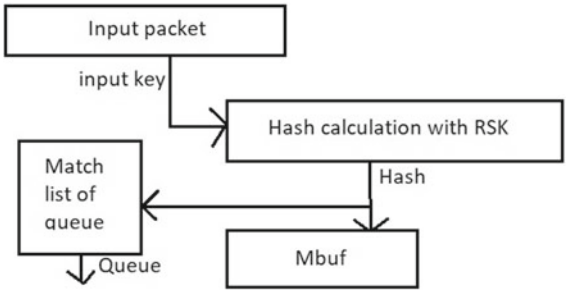



**Fig. 4** Depicts the RSS-based hash calculation

3. Using hash valued cached in MBUF structure. Distributer process is assigned to Workers using rte-distributor-create(name,s-id,n,w,algorithm type) function.
4. Then Distributer and worker periodically checks for new distributer elements and new packets

## 6 Comparative Analysis of the Frameworks

This section compares the frameworks based on throughputs, latency, CPU cycles and loss rate. The comparison of frameworks is based on their performances concerning packet length. Test was run on Ubuntu Linux, with DPDK version 17.05.2. The incoming traffic is 1–10 Gb/s with the same number of packets. It depicts DPDK packet loss with a packet length of 64, 128, 256, 512, 1024 B is relatively stable and high performance. While in Linux and PF_RING the packet loss rate is high and decreases with an increase in packet rate. 64 B packet size had a Loss rate of 7, 70, 84% with respect to DPDK, PF_ring and Linux. Whereas, with 1500 B loss rate were 0.8, 42, 58% similarly [14].

Based on the packet forwarding of the frameworks. i.e. DPDK (1.6.0), Netmap, PF_Ring(6.0.0.2). The latency and throughput were studied with 10 Gb line rate. The platform was equipped with 2 port of Intel X520SR2 NIC and generator with Intel Xeon E3-1230 V2 CPU. The latency was captured using IEEE 1588 hardware time stamping. DPDK and PF_Ring had a latency of 9 and 10 μs, whereas Netmap reached more than 1000 μs for 16 batched size. Per packet, Netmap adds more 50 cycles to CPU, Pfring used 150 cycles and DPDK used 100 cycles [15].

The insights processing capability of PF_Ring and DPDK were made based on type of distribution threads on a NUMA based on load balancing technique with respect to Pipeline Model. DPDK version 1.8.0, PF_Ring version 6.2.0, Centos based Linux platform with Intel® Xeon® CPU and Intel 10 Gbps SFI/SFP + NIC was used. With a packet rate of 64 and 1500 B, the line rate achieved by DPDK were 8, 10 Gb/s. Whereas PF_ring achieved 0.3 and 9 Gb/s [16]. The resources usage and packet rate are addressed with frameworks libpcap, DPDK, snort and PF_Ring. The experiment was conducted on an Ubuntu 18.04 with an incoming traffic rate of 1000Mbps using hping3. The CPU usage with PF_Ring was 4 and 3 with respect to packet size of 64 and 1500 B. Whereas DPDK achieved 5 and 4 with respect to 64 and 1500 B packet size. The packet loss rate with DPDK and PF_Ring with respect to 64 and 1500 B were 14, 1 and 15, 70 [17].

Figure 5 depicts the comparison based on their packet size, loss rate, throughput, and CPU cycles. If the only requirement is latency, then PF_Ring with zero copy takes less latency but with more CPU cycles than DPDK for processing the packets. If the requirement is throughput and less loss rate, then DPDK outperforms others. Netmap has good performance but lacks API's and libraries.

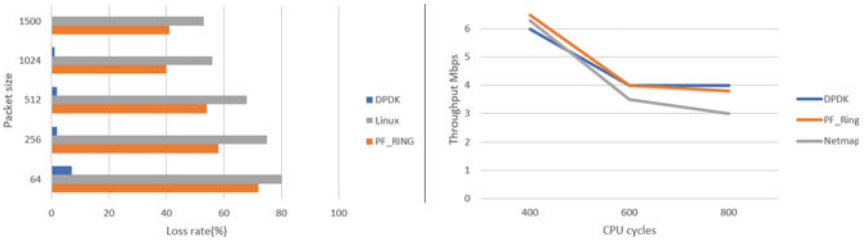



**Fig. 5** Represents the effect of packet length on packet loss, throughput and CPU cycles between Linux, PF_Ring, DPDK

## 7 Conclusion and Future Work

In this paper, a thorough walk-over on the internal structure and principles of software solutions for high-performance packet processing was done. The main stages of the Linux network protocol stack in-network packet processing were discussed. The bottlenecks of data packet processing under traditional network packet processing were discussed. Considering the inherent bottlenecks, DPDK, a software solution to a high- speed and high-performance framework is studied, and significant data processing capabilities were improved with the RSS algorithm. This study and comparison between network data packet processing between Netmap, NetSlice, and PF_RING techniques demonstrate that the Data plane development kit is magnificent, which can be used in network packet processing. In coming time, we plan to work on improvised DPDK based packet processing on NUMA and which can enhance more capabilities and optimizations.

## References

1. Van Tu N, Yoo J-H, Hong JW-K (2020) Real-time monitoring of packet processing time for virtual network functions. In: 2020 21st Asia-pacific network operations and management symposium (APNOMS), pp 138–143. https://doi.org/10.23919/APNOMS50412.2020.9237057
2. Kim S, Lee G, Woo J, Jeong J (2021) Zero-Copying I/O stack for low-latency SSDs. IEEE Comput Archit Lett 20(1):50–53. https://doi.org/10.1109/LCA.2021.3064876
3. Pandey A et al (2020) DPDK-FQM: framework for queue management algorithms in DPDK. In: 2020 IEEE conference on network function virtualization and software defined networks (NFV-SDN), pp 1–6. https://doi.org/10.1109/NFV-SDN50289.2020.9289914
4. Gong Y, Xi W, Han H, Yan J (2020) Research on a reliable method of receiving data based on zero-copy non-drop packet network card. In: 2020 22nd international conference on advanced communication technology (ICACT), pp 472–477. https://doi.org/10.23919/ICACT48636.2020.9061402
5. Emmerich P, Pudelko M, Bauer S, Huber S, Zwickl T, Carle G (2019) User space network drivers. In: 2019 ACM/IEEE symposium on architectures for networking and communications systems (ANCS), pp 1–12. https://doi.org/10.1109/ANCS.2019.8901894

6. Kim J, Na J (2017) A study on one-way communication using PF_RING ZC. In: 2017 19th international conference on advanced communication technology (ICACT), pp 301–304. https://doi.org/10.23919/ICACT.2017.7890102
7. Watanabe Y, Kobayashi Y, Takenaka T, Hosomi T, Nakamura Y (2017) Accelerating NFV application using CPU-FPGA tightly coupled architecture. In: 2017 international conference on field programmable technology (ICFPT), pp 136–143. https://doi.org/10.1109/FPT.2017.8280131
8. Ramneek S, Cha SH, Jeon YJ, Jeong JM Kim Jung S (2018) Analysis of Linux Kernel packet processing on manycore systems. In: TENCON 2018 - 2018 IEEE region 10 conference, pp 2276–2280. https://doi.org/10.1109/TENCON.2018.8650173
9. Aksić M, Redžović H, Smiljanić A (2017) Application of huge pages to the NetMap platform. In: 2017 25th telecommunication forum (TELFOR), pp 1–4. https://doi.org/10.1109/TELFOR.2017.8249300
10. Marian T, Lee KS, Weatherspoon H (2012) NetSlices: scalable multi-core packet processing in user-space. In: 2012 ACM/IEEE symposium on architectures for networking and communications systems (ANCS), pp 27–38
11. Ke He W, Peng Z, Wen Chao C, Xian Kang Z, Jun CA (2019) Tunneling SSL VPN based on PF_RING. In: 2019 IEEE 10th international conference on software engineering and service science (ICSESS), pp 1–4. https://doi.org/10.1109/ICSESS47205.2019.9040747
12. Askari L, Majidzadeh P, Ayoub O, Tornatore M (2020) Exploiting DPDK in containerized environment with unsupported hardware. In: 2020 IEEE conference on network function virtualization and software defined networks (NFV-SDN), pp 7–12. https://doi.org/10.1109/NFV-SDN50289.2020.9289904
13. https://www.dpdk.org/
14. Zhu W, Li P, Luo B, Xu H, Zhang Y (2018) Research and implementation of high performance traffic processing based on Intel DPDK. In: 2018 9th international symposium on parallel architectures, algorithms and programming (PAAP), pp 62–68. https://doi.org/10.1109/PAAP.2018.00018
15. Gallenmüller S, Emmerich P, Wohlfart F, Raumer D, Carle G (2015) Comparison of frameworks for high-performance packet IO. In: 2015 ACM/IEEE symposium on architectures for networking and communications systems (ANCS), pp 29–38. https://doi.org/10.1109/ANCS.2015.7110118
16. Wang H, He D, Wang H (2016) Comparison of high-performance packet processing frameworks on NUMA. In: 2016 7th IEEE international conference on software engineering and service science (ICSESS), pp 54–58. https://doi.org/10.1109/ICSESS.2016.7883014
17. Li J, Wu C, Ye J, Ding J, Fu Q, Huang J (2019) The comparison and verification of some efficient packet capture and processing technologies. In: 2019 IEEE international conference on dependable, autonomic and secure computing, international conference on pervasive intelligence and computing, international conference on cloud and big data computing, international conference on cyber science and technology congress (DASC/PiCom/CBDCom/CyberSciTech), pp 967–973. https://doi.org/10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00177

# Telemedicine IoT Prototype "Doctor Pi" for Measuring Elders Vital Signs in Rural Areas of Ecuador


**Carlos Bosquez and Wilson Valencia**

**Abstract** This document is focused on the designing of a prototype to be used in telemedicine called "Doctor PI" capable of receiving, transmitting, viewing, alerting, and storing data acquired by nine biomedical sensors, such as glucometer, pulse and oxygen sensor, blood pressure sensor, EMG, ECG, patient position sensor, galvanic response sensor, body temperature sensor and breathing airflow sensor, for elder people around rural areas of Ecuador. The prototype consists of an electronic health system, which in conjunction with Arduino, GSM, and Raspberry Pi will process and store the data for the medical sheet and patient analysis. The physiological parameters obtained are stored in a Raspberry Pi database and displayed via web or mobile APP. SMS alerts will be obtained in case the measured parameters exceed or decrease the values parameterized by the doctor. This prototype aims to help expand telemedicine in rural areas of Ecuador through IoT technologies with low-cost hardware and internet connectivity.

**Keywords** APP · ECG · EMG · GSM · IoT · Raspberry Pi · Telemedicine

## 1 Introduction

Telemedicine is essentially a remote doctor-patient interaction. That is why many believe telemedicine can help small hospitals that are struggling to provide quality healthcare to patients, even more, if it is done directly from home. Telemedicine can solve many problems giving access to clinical medicine by focusing on rural areas away from main medical facilities or individuals, especially for adults with low economic resources and reduced mobility, which allows access and improve the quality of medical care [1].

C. Bosquez (✉) · W. Valencia
GISTEL (Telecommunications Systems Research Group), Salesian Polytechnic University, Robles 107 and Chambers, Guayaquil, Ecuador
e-mail: cbosquez@ups.edu.ec

W. Valencia
e-mail: wvalenciaz@est.ups.edu.ec



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_76

Currently, in Ecuador there is no such implementation in the use of telemedicine, specifically in rural areas [2], this due to the little healthcare coverage that the authorities have for these areas, however in the context of a worldwide pandemic, has increased the use of the internet in rural areas of Ecuador in the last year [3], opening up the use of internet technologies for applications such as rural telemedicine.

This document emphasizes the description of the prototype "Doctor Pi", then goes through the experimentation and results of the telemedicine prototype and ends with the conclusions of the tests carried out and the technical feasibility of the prototype for commercial application in telemedicine for rural areas in Ecuador.

## 2 Related Works

Previous research in the field of telemedicine using IoT devices such as Raspberry Pi and sensors to obtain physiological patient data has used low-cost web programming and hardware for monitoring patient data [4], however, no complete prototypes are observed for monitoring with more than five sensors [5]. With the integration of GSM modules for SMS alerts and local monitoring in the prototype [6], the prototype can work with more emphasis on the development of telemedicine prototypes only with EGC sensors [7], and web monitoring [8]. Other developments in telemedicine prototypes with an emphasis on pulsometers propose the integration of a Raspberry Pi with an oximeter sensor [9]. The most approximate research to the proposal in this document is telemedicine prototypes such as web monitoring system for patient healthcare [10], where five sensors are integrated [11], Raspberry Pi and GSM module, however, do not propose the creation of an APP as a monitoring tool [12].

In this context, the research focuses on the design of a prototype IoT telemedicine with nine sensors that integrates local monitoring, via web, via GSM, and via mobile APP, to be used in telemedicine around all rural areas of Ecuador. The telemedicine prototype has been called "Doctor Pi" and consists in the integration of an e-health PCB of cooking-hacks [13], Raspberry Pi, Arduino, GSM module, nine biomedical sensors that together with web programming and a mobile App can visualize, store and treat information from biomedical sensors (glucometer, pulse and oxygen sensor, blood pressure sensor, EMG, ECG, patient position sensor, galvanic response sensor, body temperature sensor, and airflow sensor inbreathing) which are processed and stored on a Raspberry Pi web server.

With this friendly and easy-to-use prototype, you can keep basic control of the health status of elder people thanks to its storage on a web server. This will help doctors to monitor the patient's condition through an APP and web access, thus carrying a clinical history of easy use and access that would be reviewed by the doctor for future treatments.

# 3 Methods

## *3.1 Description of the Prototype "Doctor Pi"*

Figure 1 shows the block diagram of the prototype "Doctor Pi", where you can visualize the different processes from the taking of the information made by the e-health board using the sensors obtaining the analog data, that is sent to the Arduino board so that the information is digitized and attached to the Raspberry Pi. In this way, the measurements obtained from the patient can be reflected in the web application or an APP through the cloud thanks to the internet. With the data obtained from the prototype, you can view information from the nine biomedical sensors and anywhere in the world [14].

Figure 2 shows the e-health sensor platform is a complete measurement system that allows monitoring the human body through nine sensors connected to a central board, the e-health shield, generating signals that can be interpreted through an Arduino board or a Raspberry Pi microcomputer [15].

The Arduino UNO is a board based on the ATmega328P microcontroller [16]. It has 14 digital input/output pins (of which 6 can be used with PWM), 6 analog inputs, a 16 MHz glass, USB connection, power jack connector, ICSP connection terminals, and a reset button, which will be a bridge between analog and digital signals that will be transmitted to the local server. It also has a low power consumption which is equal to 3.3 W at rest [15].

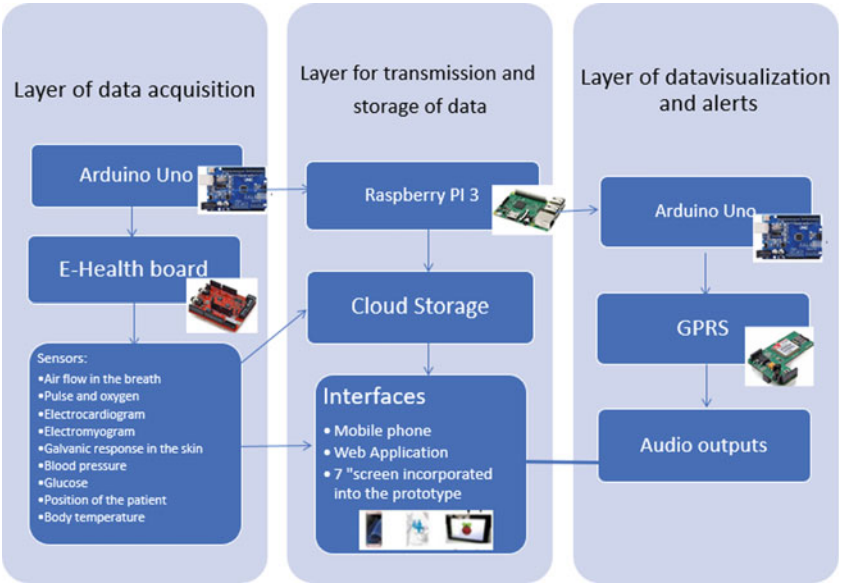



**Fig. 1** Block diagram of the "Doctor Pi" prototype

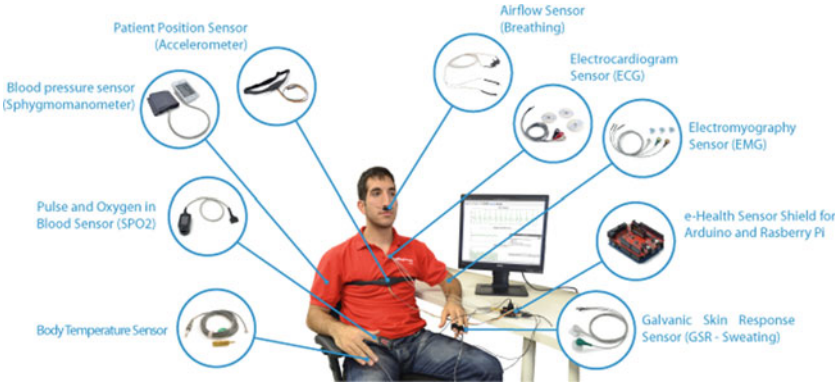



**Fig. 2** Sensor connectivity [13]

The Raspberry Pi is a low-cost, small-sized computer. It works with low power, consumes fewer resources compared to a desktop PC or laptop. Also, it has a low power consumption that is equal to 4 W [7]. The card is capable of supporting three video cards, facilitating the user interface. When using Arduino one for the implementation of the prototype, the consumption around energy is minimal, this being 7 to 12 V for the system to work [15]. Its coding will be carried in Ide Arduino being this a free software.

Table 1 shows the components that make up the prototype and the instruments used by a doctor, along with their price respectively, making a comparison of the monetary value. Doctor Pi reduces the cost of implementation, execution, and development to monitor vital signs.

**Table 1** Comparative table of components between doctor pi and medical instruments

| Doctor PI | | Medical equipment | |
|---|---|---|---|
| Components | Cost ($) | Components | Cost ($) |
| E-Health and sensors v1 | 350 | 3-Channel Electrocardiogram | 700 |
| Raspberry Pi 4/Touch screen 7”/Mouse/Keyboard/Speakers/Speech | 120 | Vital sign monitor | 750 |
| Shield GPRS/GSM | 40 | Digital arterial tensiometer | 95 |
| Pulse controller | 21 | Glucometer | 45 |
| Arduino Uno | 20 | SpO2 | 35 |
| Glucometer | 19 | Body temperature | 30 |
| Case | 60 | Oxímetro | 20 |
| Web and APP design | 100 | CPU/Monitor 19”/Mouse/Keyboard/Speakers | 500 |
| Total | 730 | Total | 2175 |

**Table 2** Applicability of sensors in the selection of 4 common categories of diseases [13]

| Sensor | Cardiovascular disease | Chronic pulmonary obstructive disease | Parkinson's and Huntington's diseases | Diabetes |
|---|---|---|---|---|
| Electrocardiogram | XX | XX | X | X |
| Airflow in breathing | XX | XX | X | - |
| Body temperature | X | X | X | X |
| Electromyography | X | X | X | – |
| Galvanic response | – | – | X | – |
| Blood pressure | XX | X | X | X |
| Glucose | – | – | – | XX |
| Oxygen | XX | XX | – | – |
| Movements (position) | XX | XX | XX | – |

XX Indicates high applicability
X Indicates average applicability
– Indicates indeterminate applicability

When using all these implements the efficiency of having an accurate diagnosis is effective, without complications taking into account that the results will be sent to the doctor [16].

Table 2 shows the application of sensors in the arrest of various diseases. The data obtained by the sensors used have applicability for the diagnosis of four common diseases. The sensors used are light, small and try not to impede the free mobility of patients.

Figure 3 shows the block diagram of the Doctor Pi prototype and the tools to be used, also has a web interface made in HTML and CSS, from which the doctor can click to perform a measurement using the sensors [11]. When performing this action, an HTTP request is made to the webserver which has a program written in PHP using the Laravel framework which runs a program in Python [12]. The Python program asks the Arduino to send the data of a certain sensor through the serial port, then the sensor data is displayed in the web interface (HTML and CSS) [16].

The Arduino has a program written in its language, which has the function of waiting for some instruction through the serial port, when this instruction arrives, the sensor data is read, and it's returned through the serial port [16].

In the web interface, the deployed data is sent to the remote server using JavaScript as these must be uploaded asynchronously [11]. The data on the server is received by a script program in PHP and gets stored in a MySQL database so that later the Android application written in Java consults the data and displays it in the mobile interface [12].

After the Arduino collects measurements of the sensors obtained by the e-Health board, they move to the Raspberry Pi board to store all the information on the webserver [14].

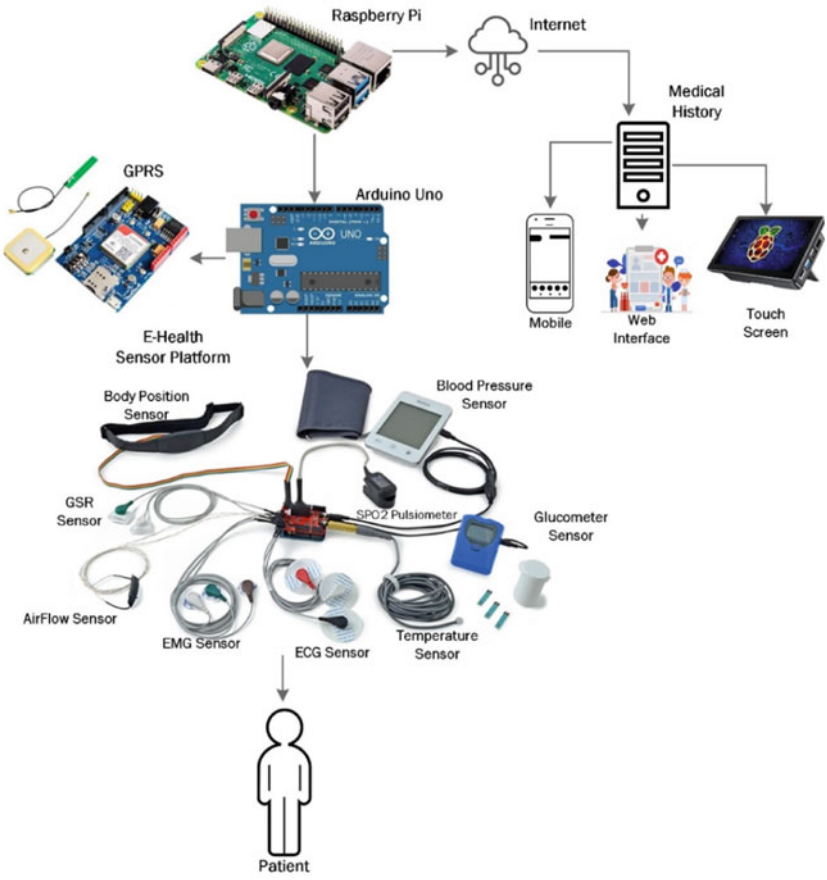



**Fig. 3** Schematic of the prototype "Doctor Pi"

Figure 4 shows the main screen of the APP on a mobile, specifically the sections to take measurements and view statistics.

Figure 5 shows the final prototype "Doctor Pi" consisting of a telemedicine device equipped with a 7 inches screen, keyboard, mouse, speakers as peripherals, and a built-in development, with low-cost hardware such as an e-health card, GSM shield, Raspberry Pi 3, and Arduino.

Data visualization plays an extremely important role because medical equipment needs to have an easy way to view and analyze the hundreds and even thousands of data collected and transmitted. Electromyography and electrocardiogram should be visualized using a rendering graph, of all the values obtained from the measurements. Visualization methods enable an accessible way to manage and analyze data avoiding traditional methods, such as medical records or a large number of archived documents. Since the patient's personal information is fully private, only authorized


**Fig. 4** Android APP interface

**Fig. 5** Final prototype Doctor PI

persons will be able to access this database, who will have to authenticate their user and password when logging in from the built-in means for data visualization.

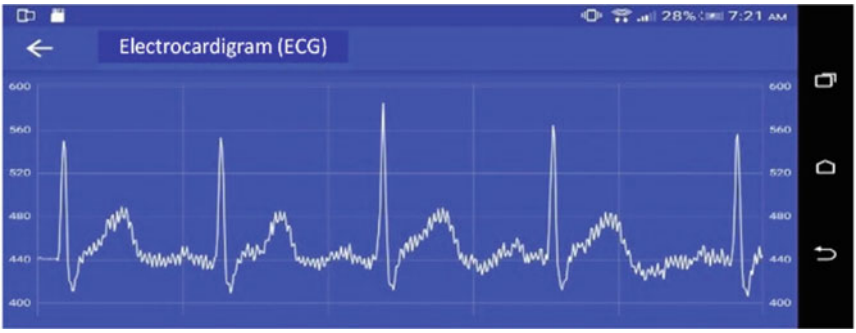



**Fig. 6** Results of electrocardiogram sensor (APP interface)

## 4 Results

Several tests using the prototype were carried out in older adult patients in rural areas of the coast of Ecuador, obtaining results helps to technically validate the prototype. Samples were taken in elder people with different problems and elder people who are in good health. 100 samples were taken from 50 people between 65 to 75 years and 50 people between 76 and 80 years.

During experimental samples, the prototype is set up to back up each sample on the server, which stores the clinical history of each measurement and each patient also. When a sample is out of range an alarm occurs, a missed call is made, and a text message is sent to the previously programmed number. Figure 6 shows the APP interfaces of sensor ECG, from the interface of a smartphone.

Tests were performed using the prototype with an out-of-range measurement of the glucose sensor to show the missed call and a text message, alerting the family doctor or caretaker of the eldest adult.

## 5 Discussion and Conclusions

The Doctor PI prototype consists of different low-cost hardware components and free software that granted the possibility to potentialize the device for future work in an improved version, as well as making it useful for different types of patients with different needs. This telemedicine prototype covers the clinical need for low-income and rural adult patients and opens the gap in the use of telemedicine in vulnerable sectors, especially in Ecuador. For its widespread use, it is recommended that government authorities agreed to globalize the use of telemedicine, especially in the most vulnerable areas.

Data obtained during patient samples ensure an approximation of data compared to higher-cost medical equipment. Thanks to its easy interpretation of clinical history data and consultations through an APP, the medical professional can evaluate and

track online the samples acquired in patients. Access to the mobile network must be available to use the GPRS/GSM module of the prototype without problems because this will be responsible for sending the text messages and making the missed call to the scheduled number and alert that some measurement is outside its allowed range. This would help quickly and effectively to keep informed the family about the patient's situation in case the prototype detects any anomalies in the measurement.

Another advantage of "Doctor Pi" is the communication with the user, since an audio system was implemented that allows to reproduce of the measurement made during the sample taking or data in an audible way. Audio will also be heard if the ranges are out of range. The obtained results show the prototype technical test. It also indicates that in most cases with disabled patient measurement. The complexity to make the prototype is not so high because anyone with knowledge of electronics and basic programming along with the indications and instructions indicated on web pages or in this paper can be used to apply or replicate the prototype. Therefore, the complexity of use is very reduced because a responsive web-type design was made in which only data is selected, similar to a mobile application to take the measurements, however, it is recommended to previously reviewing the instructions for the use of the equipment and have basic knowledge of computing and medicine.

Finally, it is concluded that the prototype is technically feasible due to the obtained results and thanks to the acceptance of the clinical staff that has used the prototype. Doctor Pi was evaluated at the Hospital of the Guayaquil Institute of Social Security by doctors and professional nurses who determined that it is a device that approximates values compared to professional measuring instruments and is recommended Install in rural health centers where resources are limited. In Ecuador there are no implementations of these prototypes for telemedicine in rural areas therefore it would be the first attempt in the country with the prototype "Doctor Pi".

## References

1. Naranjo D et al (2019) Wearable telemedicine system for real-time monitoring of electro cardiographic signals. In: Sixth international conference on eDemocracy and eGovernment (ICEDEG), pp 69–75. IEEE, Quito, Ecuador
2. noticias.usfq.edu.ec Homepage. https://noticias.usfq.edu.ec/2020/08/proyecto-usfq-de-telemedicina-apoya.html. Accessed 09 June 2021
3. Elcomercio.com Homepage. http://www.elcomercio.com/actualidad/negocios/ecuador-internet-tecnologia-fibra-optica.html. Accessed 09 June 2021
4. Ahsanuzzaman SM, Ahmed T, Rahman MA (2020) Low cost, portable ECG monitoring and alarming system based on deep learning. In: IEEE Region 10 Symposium, TENSYMP 2020, pp 316–319. Dhaka, Bangladesh
5. Orbe Y, Alulema D, Trivino RD, Guaman AV, Andres Arcentales V (2020) Design of a mHealth device for peak expiratory flow measurement. In: IEEE ANDESCON, ANDESCON 2020. Quito, Ecuador
6. Hodrob R, Obaid M, Abdulsalam Mansour AM, Sawahreh A, Naghnagheah M, Abushanab SS (2020) An IoT based healthcare using ECG. In: 21st international Arab conference on information technology, ACIT 2020, pp 1–4. Giza, Egypt

7. Khan I et al (2019) Healthcare monitoring system and transforming monitored data into real-time clinical feedback based on IoT using Raspberry Pi. In: 2nd International Conference on Computing, Mathematics and Engineering Technologies, iCoMET 2019. Sukkur, Pakistan
8. Rahman A, Rahman T, Ghani NH, Hossain S, Uddin J (2019) IoT based patient monitoring system using ECG sensor. In: 1st international conference on robotics, electrical and signal processing techniques, ICREST 2019. Dhaka, Bangladesh
9. Aileni RM, Pasca S, Florescu A (2019) E-health monitoring by smart pulse oximeter systems integrated in SDU. In: 11th international symposium on advanced topics in electrical engineering, ATEE 2019. Bucharest, Romania
10. Chand RD, Kumar A, Kumar A, Tiwari P, Rajnish R, Mishra SK (2016) Advanced communication technologies for collaborative learning in telemedicine and telecare. In: Proceedings of the 9th international conference on cloud computing, data science and engineering, Confluence 2019, pp 601–605. Noida, India
11. Kassem A, Tamazin M, Aly MH (2021) A context-aware IoT-based smart wearable health monitoring system. In: 2020 international conference on communications, signal processing, and their applications (ICCSPA), pp 1–6. Sharjah, United Arab Emirates
12. Ramani JG, Madhusudan S, Nila AL, Pradeep A, Manibharathi S (2020) IOT based employee health monitoring system. In: 2020 6th international conference on advanced computing and communication systems (ICACCS), pp 298–301
13. Cooking-hacks.com Homepage. https://www.cooking-hacks.com. Accessed09 June 2021
14. Sharma HK, Taneja S, Ahmed E, Patni JC (2019) I-Doctor: An IoT based self patient's health monitoring system. In: 2019 international conference on innovative sustainable computational technologies, CISCT 2019, pp 1–6, Dehradun, India
15. Lu D et al (2021) XTSeH: a trusted platform module sharing scheme towards smart IoT-eHealth devices. IEEE J Sel Areas Commun 39(2):370–383
16. Pendurthi HK, Kanneganti SS, Godavarthi J, Kavitha S, Gokarakonda HS (2021) Heart Pulse Monitoring and Notification System using Arduino. In: 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), pp 1271–1278 Coimbatore, India

# Virtual Machine Consolidation Using Enhanced Crow Search Optimization Algorithm in Cloud Computing Environment


**Kethavath Prem Kumar, Thirumalaisamy Ragunathan, and Devara Vasumathi**

**Abstract** The widespread usage of cloud computing technologies benefits the services, providers and investors in constructing the large scale data centers. However, an increase in energy consumption from Physical machines showed significant impact on environment due to emissions of carbon dioxide. The Virtual Machines (VMs) used the minimal number of Physical Machines (PMs) for performing Dynamic Consolidation obtained magic solutions for managing the power consumption. The present research uses Enhanced Crow Search Optimization Algorithm (ECSOA) technique for reducing the usage of energy by mapping the VMs migrating to hypothesis showed an aggressive consolidation. The ECSOA consolidates the VMs by migrating the number of PMs which decreases the number of hosts as they were overloaded in an environment. The number of migrations are drastically reduced as it considered VMs as a main source of migration which overloaded and under loaded the hosts. The results obtained from the research showed that the proposed ECSOA obtained SLA violation of 0.000267 better when compared to the existing Osmotic Hybrid Artificial Bee-Ant Colony Optimization (OH-BAC) and Particle Swarm Optimization (PSO)-Decimal Encoding techniques that obtained 0.5 and 0.0002 SLA violation.

**Keywords** Energy consumption · Enhanced crow search algorithm · Service level agreement · Virtual machines · VM migrations

K. P. Kumar (✉)
Department of Computer Science and Engineering, ACE Engineering College, Hyderabad, India
e-mail: kpremkumarcs@gmail.com

T. Ragunathan
Department of Computer Science and Engineering, SRM University, Amaravathi, India
e-mail: ragunathan.t@srmap.edu.in

D. Vasumathi
Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Hyderabad, India
e-mail: asukumar_devara@jntuh.ac.in



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_77

## 1 Introduction

The cloud computing system is created and allows an appropriate required network for accessing the common store for customizing the resources such as storage, applications and networks [1]. The cloud computing allows people all over the globe for their respective clients working in companies such as Salesforce, Microsoft, Amazon, IBM office for storing and computing data tries to build the latest data center [2]. The VM live migration will be a major advantage for virtualization as it is helpful in managing the cloud environment resources [3]. The VMs are migrated seamlessly and shows transparency from physical server of one to another [4]. However, there occurs challenges when VM consolidation shows optimization problems for time based approaches utilized resource consuming [5]. The major challenge on system performance is that it performs consolidation with other VMs which increased workload dynamically [6]. The huge amount of electricity is consumed and greenhouse gas like CO2 are emitted on the environment due to the data centers which contains plenty of physical machines such as hosts or servers [7]. The problem will lead to hold the Quality of Service (QoS) of applications which in turn results in high reaction time, non-success or down time [8]. In order to overcome such an issue, Enhanced Crow Search Optimization algorithm is proposed and it reduces the consumption of power by mapping the VMs, which shows an aggressive consolidation approach in VM [9]. The present approach will minimize the number of servers that reduces the migration thrashing and will not minimize the sever numbers as it reduced the thrashing migration [10].

The structure of the paper is followed as: Sect. 2 describes the surveys related work on VM consolidation for the cloud data centers. Section 3 describes about the proposed ECSO algorithm and Sect. 4 describes about the results and discussion of the proposed ECSO algorithm. The conclusion and future work for the present research work is presented in Sect. 5.

## 2 Literature Review

The review on the existing researches have undergone for the VM consolidation with respect to cloud data centers are as explained as follows:

Ibrahim et al. [10] developed a power aware technique that was used for determining an optimal placement of the migrated VMs. The Particle Swarm Optimization (PSO) algorithm was used for decimal encoding that mapped the migrated VMs based on the best appropriate PMs. However, the number of overloaded servers necessitated for migrating VMs overloaded hosts which violated SLA. Mosa et al. [11] presented an approach that used utility functions for self-management of VM placement that created a solution in cloud data centers. The utility based approaches were used for applications that created VM placement problem estimated the profit from the adaptive functions. However, the developed model consumed more CPU time of the

server's power as every network resources and hosts were involved in an adaptive VM-PM assignment.

Malekloo et al. [12] developed energy-aware and QoS-aware Multi-Objective Ant Colony Optimization (MACO) approach that performed consolidation and VM placement. The developed MACO VMs used the consolidation algorithm that build the results based on the VM migration and algorithm placements. However, the objective functions used resulted time complexity and degraded the performance. Malekloo et al. [13] developed Multi-Objective Ant Colony Optimization (MACO) that minimized the SLA violations and energy consumption, which improved the QoS for the VM allocation. The developed model showed degradation in performance as it was difficult for determining best solutions that simultaneously satisfied the requirements. Gamal et al. [14] combined the Artificial Bee Colony and Ant Colony Optimization (ABC + CO) formed an algorithm based on the behavior of osmotic metaheuristics hybridized Osmotic Hybrid Artificial Bee and Ant Colony Optimization (OH-BAC) for effective load balancing. Though, the OH-BAC method achieved better performance and its SLA Violations Time per Active Host (SLATAH) was higher than existing algorithms.

## 3 Proposed Methodology

The block diagram of the proposed method is shown in the Fig. 1 that includes the process of placing VMs for detecting the overloaded and under loaded hosts. The proposed approach involves 3 process such as initialization, mapping and fitness function.

### *3.1 Initialization*

The present research work used CSO algorithm that started to initialize the parameters. The factors such as population size, number of iterations, number of particles are utilized for mapping the solutions. The CPU utilization is the main factor used with respect to power consumption, migrates VMs for obtaining the solution.

### *3.2 Mapping*

The particle movements are controlled for each particle and each of them are updated with the speed. The initial value is set with a dimension that represents the particle speed and their positions which will lead to new position of available host list based on the representation of index.

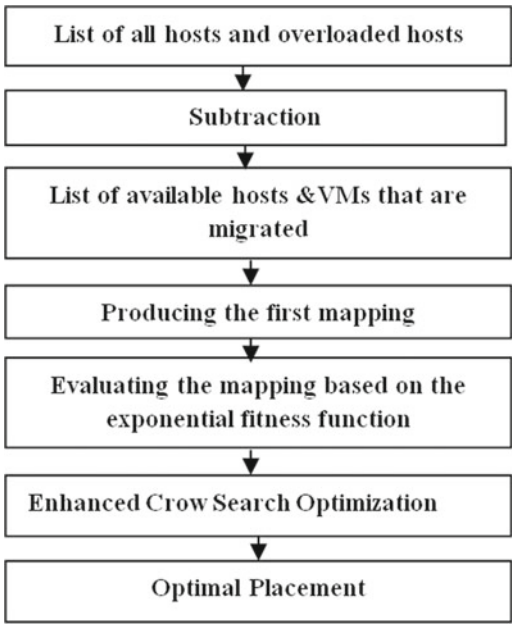



**Fig. 1** Flow diagram of proposed ECSOA

### *3.3 Fitness Function*

The present research goal is achieved with the benefits of fitness function as it is capable of weighing the total cost function for finding the composite objective function. The developed model uses fitness function for finding the best global function that is having optimal mapping finds the minimal VMs with PMs the consumption of power without increase in the overloaded servers.

### *3.4 Steps Involved in the Enhanced Crow Search Optimization Algorithm*

The present research paper is based on the meta heuristic CSA algorithm includes the intelligent population behaviors. The flock size is also referred to as flock size having crows size $N$ with the position $i$ at the time $iter$ is identified with the vector $x^{i,iter}$ ($i = 1, 2, \ldots, N; iter = 1, 2, \ldots, iter_{max}$) where $x^{i,iter} = [x_1^{i,iter}, x_2^{i,iter}, .., x_d^{i,iter}]$ and $iter_{max}$ defines the maximum number of iteration. The $iter$ has the hiding place for the crow $i$ and the positions are updated as shown $m^{i,iter}$ in Eq. (1).

The best crow position is updated as $i$ and thus the memory for each row position will be based on crow $j$ position in the hiding place. The two states of the algorithm happen based on the following.

**State 1:** The crow $i$ will be not having an idea of crow $j$, that crow $j$ is following crow $i$. The crow $i$ has an approach for crow $j$ position hiding and the new position $i$ is obtained as shown below:

**Step 1:** The CSA mimics the crow birds' behavior as shown in the following and the group of crows represents the search agent that obtains the solution for population initializing based on representing the environment.

**Step 2:** The population of $m$ crows are initialized with the dimension $t$ and the Memory $M$ for all crows are initialized using the Eq. (1).

$$x^{i,iter+1} = x^{i,iter} + r_i \times fl^{i,iter} \times (m^{j,iter} - x^{i,iter}) \tag{1}$$

where, $r_i$ is known as the random number obtained based on the uniform distribution of 0 and 1.

$fl^{i,iter}$ is the flight length for the $i^{th}$ crow at the iteration $iter$.

**State 2:** Crow $j$ knows that $i$ number of crows are following and as a result the cache is protected from crow $j$. It fools the crow $i$ reaches another crow search space position.

**Step 1:** The crow $j$ knows that the crow $i$ has another position and that position will be watched by crow $j$. The crow discovers the food's hiding place and the crow $j$ will randomly fool the crow $i$. There are two states based on the awareness probability $AP^i$ as each of the crows follows in population.

Totally 1 and 2 states are expressed by using the Eq. (2)

$$\begin{aligned} x^{iter+1} = \{ & x^{i,\,iter} + r_i \times fl^{i,\,iter} \times \left(m^{j,\,iter} - x^{i,\,iiter}\right) r \\ & \geq AP^{j,\;iter}\; a\;\; random\;\; position\, otherwise \} \end{aligned} \tag{2}$$

From the Eq. (2) $r_j$ is known as the random number uniformly distributed lies between 0 and 1 and the awareness probability $AP^{j,iter}$ at the iteration $iter$ for crow $j$

**Step 2:** The crow has a solution that updates the memory based on the fitness value and the new crow's position is based on the current memory value using $f(x) = a^x$. Where "$x$" is a variable and "$a$" is a constant for base function that should be greater than 0. The proposed ECSOA technique reduces the consumed power thereby maps the migrated VMs up to a number of minimum PMs and also avoids for producing the overloaded servers kept the SLA. Figure 2 shows the ECSOA flow chart and implementation that is required to optimize the parameters using ECSOA.

**Pseudo-code for ECSO**

**Step 1:** Initialize problem and the parameters such as flock size, maximum iterations ($itermax$), Awareness probability as ($AP$) and flight length ($fl$).

**Step 2:** The position of the crows and memory are expressed using the Eq. (3) & (4).

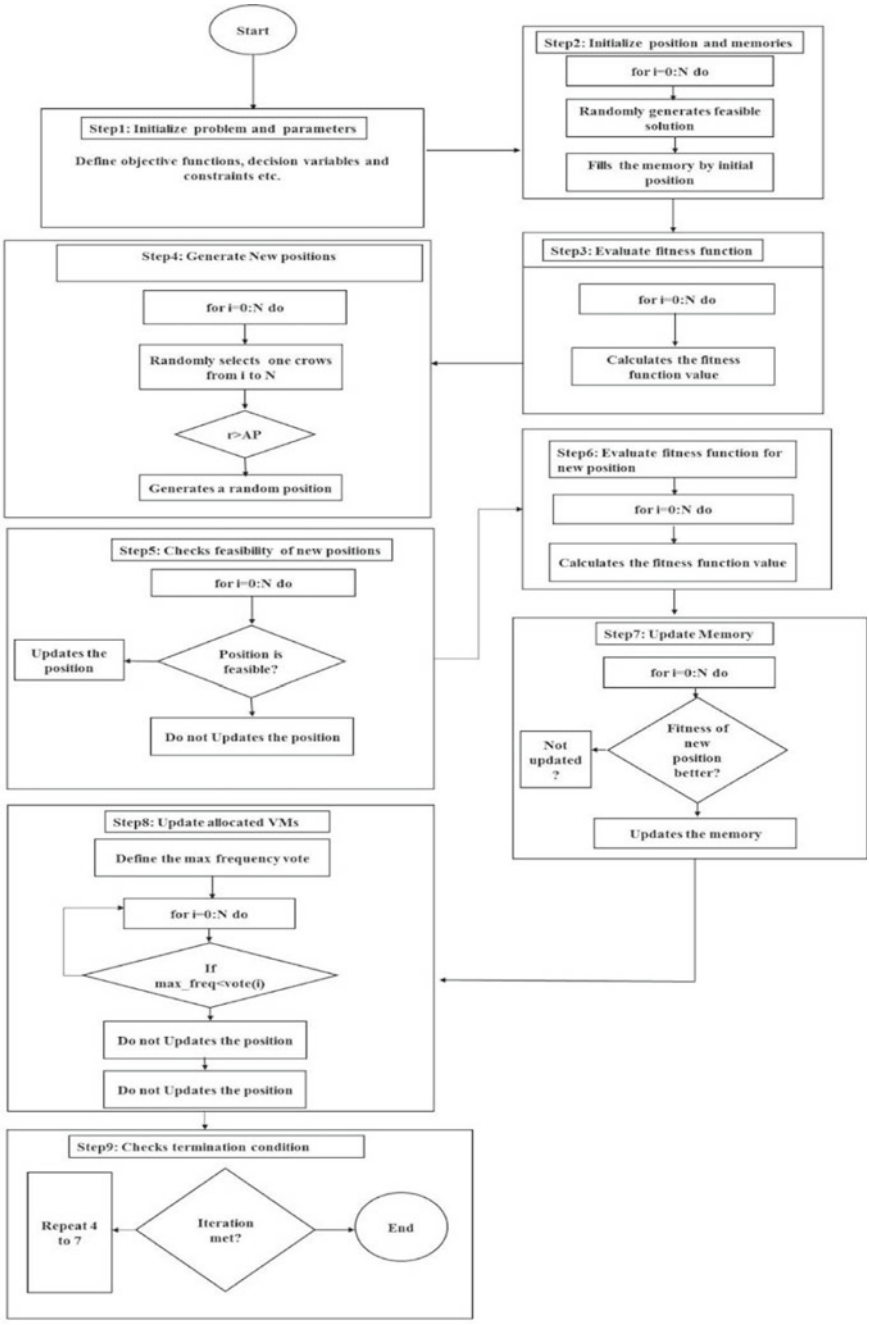



**Fig. 2** The flowchart of ECSOA implementation

$$Crows = \left[ x_1^1 x_2^1 \cdots x_d^1 x_1^2 x_2^2 \cdots x_d^2 \vdots\vdots\vdots x_1^N x_2^N \cdots x_d^N \right] \tag{3}$$

$$Memory(M) = \left[ m_1^1 m_2^1 \cdots m_d^1 m_1^2 m_2^2 \cdots m_d^2 \vdots\vdots\vdots m_1^N m_2^N \cdots m_d^N \right] \tag{4}$$

**Step 3:** The fitness function for each crow is evaluated for position $d$.

**Step 4:** The new positions of the crows are generated in the search space. The crow's new position is updated based on the Eq. (2). The position of the food hidden and are discovered based on the $m$ crow and repeat for all those crows.

**Step 5:** The feasibility for the new positions of the crows are checked. If crow position is suitable then the crow position is updated or else the crow current position will be stayed and will not move for other new position.

**Step 6:** The fitness function is evaluated for the new positions and fitness value for each of the updated crow positions are computed.

**Step 7:** The memory of the of crow's position is updated using Eq. (5).

$$m^{i,iter+1} = \left\{ x^{i,iter+1} f\left(x^{i,iter+1}\right) is\ better\ than\ f\ (m^{i,iter}) m^{i,iter} \right\} \tag{5}$$

where $f(\cdot)$ represents an objective function and crow's new position is updated and memorized with the position.

**Step 8:** The termination process includes the steps from 4 to 7 that will be repeated until the max $iter$ is satisfied. Once the criteria for the termination is followed, then the best position for the objective function is reported overcomes the problem of optimization.

## 4 Results and Discussion

The proposed ECSOA method is evaluated in the system that possess Intel i7 processor with 8 GB of RAM, 500 GB hard disk and total number of user requests (i.e. 250 tasks) are used to evaluate the performance of the proposed method. The cloud setup configuration of the CloudSim environment will be given in the Table 1.

**Table 1** Cloud setup configuration of the CloudSim environment

| Configuration | C1 | C2 | C3 |
|---|---|---|---|
| Number of hosts | 50 | 75 | 100 |
| Number of VMs | 50 | 75 | 100 |
| Bandwidth | 100 | 100 | 100 |
| MIPS | 2500 | 2000 | 1500 |
| No. of cores | 2 | 2 | 2 |

### 4.1 Service Level Agreement Violation

SLA is defined as the level of service quality agreed between the cloud service providers and the users. ESV is mathematically defined in Eq. (6).

$$ESV = Energy\left(\frac{kw}{h}\right) \times SLA \tag{6}$$

**Energy Consumption.** Energy Consumption is the usage of data from the shared pool of computing resources. Then, the mathematical Eq. (7) is used to compute the total energy consumption of a cloud datacenter with $n$ nodes as follows.

$$EC = \sum_{i=1}^{n} x_i EC_i \tag{7}$$

where, $x_i = \{0\, if\ the\ H_i\ is\ shutdown\ 1\, other$

**Number of VM Migration.** VM migration is the movement of virtual machines from particular environment to other environment which is evaluated by using Eq. (8).

$$T_1 = \frac{N_m}{N_c} \tag{8}$$

$N_m$ is the number of the VMs migrations occurred.
$N_c$ is the number of consolidation steps during the whole execution.

### 4.2 Quantitative Analysis

Table 2 shows the results obtained by the proposed ECSOA in terms of SLA violation, ESLV, Number of Host shutdowns, energy consumption and number of VM migrations is shown in Table 2.

The amount of reduction is averagely estimated up to 50% that considered the consumed energy in data centers as an addition for SLA violation. Similarly, ECSOA will decrease as the number of active hosts migrates and consolidates the VMs

**Table 2** The proposed ECSOA in terms of SLA violation, ESLV, Number of Host shutdowns, energy consumption and number of VM migrations

| Workload | SLA violation | ESLV | Energy consumption | Number of host shutdowns | Number of VM migrations |
|---|---|---|---|---|---|
| C1 | 0.000119 | 0.00595 | 33.4 | 204 | 1635 |
| C2 | 0.000238 | 0.0119 | 42.2 | 413 | 3274 |
| C3 | 0.000267 | 0.01335 | 52.75 | 687 | 6845 |

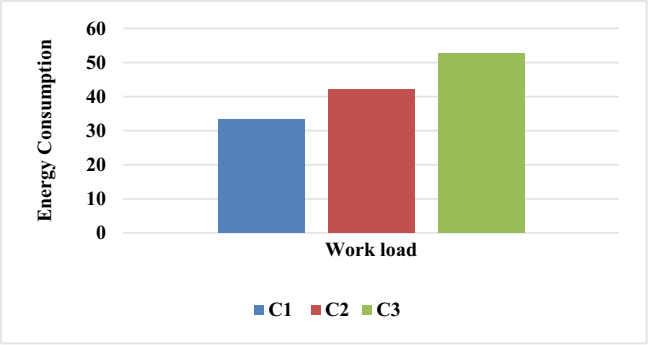



**Fig. 3** Energy consumption

position into minimal hosts number. Thus, as the CPU utilization is covering the servers of running mode, the other hosts will be switched to the sleep mode. The CPU utilization in a server is having a specific limit and the VMs are migrated from one server to the other.

Thus, the under loaded server will be switched off as they put into low power mode. The proposed ECSOA technique reduces the number of VM migrations from 1635 to 6845 that considered categories C1, C2, and C3. ECSOA achieves a remarkable reduction as the number of VM migrations are increased in terms of energy consumption as shown in Fig. 3.

## *4.3 Comparative Analysis*

Table 3 is the comparative analysis done for the proposed ECSOA method with other existing techniques MACO [13], OH-BAC [14], PSO [9] techniques in terms of energy consumption and SLA violations. The existing PAPSO was not applicable with respect to real environment for improving the efficiency. The developed model failed to considered memory for optimization and the network factors. Similarly, the OH_BAC has more SLA that affect the performance of the considered cloud system. Lastly, MACO increased the number of objective functions in the developed model showed system complexity and degraded the performance and therefore

**Table 3** Comparative analysis for the proposed ECSOA with the existing methods

| Methodology | SLA violation | Energy consumption (J) |
|---|---|---|
| PSO based on the decimal encoding [9] | 0.0002 | 64 |
| MACO [13] | 1.5 | 78 |
| OH-BAC [14] | 0.5 | 82 |
| Proposed ECSOA | 0.000267 | 52.75 |

finding solution to meet these requirements simultaneously was difficult. Whereas the proposed ECSOA considered Memory for the optimization process that reduced the complexity during VM migrations which lowered the SLA violation and also Energy consumption when compared to the existing PAPSO, OH-BAC, and MACO. From the Table 3, the proposed ECSOA has less SLA violation when compared to other existing techniques. For instance, ECSOA has only 0.000267 SLA violation, where OH-BAC has 0.5 SLA Violation and MACO has 1.5 SLA Violation.

## 5 Conclusion

The present research work discussed the dynamic consolidation of VMs as the number of servers solves the significant issues with respect to the power consumption. The present research uses ECSOA power aware VM placement technique using Exponential function, which reduces the power consumption with no violation of SLA. The exponential function has an effective objective function and fitness function is used reduce the consumption of power. Also, finding an appropriate host will help the hosts to receive the migrated VMs for vital mission. The exponential function is an effective fitness function that was employed for reducing the number of overloaded ones in the active servers. In the future work, ECSOA efficiency can be applied for real time environment through implementation.

## References

1. Saadi Y, El Kafhali S (2020) Energy-efficient strategy for virtual machine consolidation in cloud environment. Soft Comput 24:1–15
2. Fard SYZ, Ahmadi MR, Adabi S (2017) A dynamic VM consolidation technique for QoS and energy consumption in cloud environment. J Supercomput 73(10):4347–4368
3. Tarahomi M, Izadi M, Ghobaei-Arani M (2020) An efficient power-aware VM allocation mechanism in cloud data centers: a micro genetic-based approach. Cluster Comput 24:1–16
4. Tarafdar A, Debnath M, Khatua S, Das RK (2020) Energy and quality of service-aware virtual machine consolidation in a cloud data center. J Supercomput 76:1–32
5. Han G, Que W, Jia G, Shu L (2016) An efficient virtual machine consolidation scheme for multimedia cloud computing. Sensors 16(2):246
6. Sharma O, Saini H (2019) Energy and SLA efficient virtual machine placement in cloud environment using non-dominated sorting genetic algorithm. Int J Inf Secur Priv (IJISP) 13(1):1–16
7. Gomez-Rodriguez MA, Sosa-Sosa VJ, Carretero J, Gonzalez JL (2020) CloudBench, an integrated evaluation of VM placement algorithms in clouds. J Supercomput 76:1–34
8. Wang H, Tianfield H (2018) Energy-aware dynamic virtual machine consolidation for cloud datacenters. IEEE Access 6:15259–15273
9. Ibrahim A, Noshy M, Ali HA, Badawy M (2020) PAPSO: a power-aware VM placement technique based on particle swarm optimization. IEEE Access 8:81747–81764
10. Mosa A, Paton NW (2016) Optimizing virtual machine placement for energy and SLA in clouds using utility functions. J Cloud Comput 5(1):17

11. Li L, Dong J, Zuo D, Wu J (2019) SLA-aware and energy-efficient VM consolidation in cloud data centers using robust linear regression prediction model. IEEE Access 7:9490–9500
12. Moges FF, Abebe SL (2019) Energy-aware VM placement algorithms for the OpenStack Neat consolidation framework. J Cloud Comput 8(1):1–14
13. Malekloo MH, Kara N, El Barachi M (2018) An energy efficient and SLA compliant approach for resource allocation and consolidation in cloud computing environments. Sustain Comput Inf Syst 17:9–24
14. Gamal M, Rizk R, Mahdi H, Elnaghi BE (2019) Osmotic bio-inspired load balancing algorithm in cloud computing. IEEE Access 7:42735–42744

# Web Vulnerability Detection: The Case of Cross-Site Request Forgery Using Classification and Regression Trees


**Rajendra Gurram, P. Dhanunjaya Babu, Adusumalli Sai Tejaswi, Chattu Sai Ganesh, and Karlaputi Narendra**

**Abstract** In the world of hacking, one of the simple and human-understandable web attacks is Cross-Site request forgery, but this is the most top priority attack that needed to handle with keen observation because of the very small difference in their appearance. A hacker develops a cross-site application that looks similar to the original site and embeds this URL in either back transactions or E-commerce applications, it misleads the user to deviate from the original website without his notice by disabling all the authentication mechanisms. The hackers develop this type of cross website by embedding the SQL injection queries either in HTML navigation pages or while executing the external Java Scripts. To develop a solution for this type of attack, the system needs a tool that balances both scalability and usability because of millions of users working with the internet to performs various tasks on social media platforms. This tool can also isolate web attacks by designing robust web applications which include advanced encryption algorithms, which require a lot of effort to decrypt the source code. The developers of the website have to work a lot to take care of the security functionalities, negligence of which may cost the website source code to be hacked. Accidentally, Cross-Site request forgery (CSRF) attacks can be left behind, which motivated recent research on asynchronous Cross-Site request forgery detection. The proposed paper uses the Classification and Regression Trees (CART) Algorithm to detect phishing websites based on the URL posted.

**Keywords** Cross-Site request forgery · Classification and Regression · Security attacks · Website phishing

R. Gurram (✉)
Department of Information Technology, Lakireddy Bali Reddy College of Engineering, Mylavaram, India
e-mail: rajendra.raju@gmail.com

P. Dhanunjaya Babu · A. S. Tejaswi · C. S. Ganesh · K. Narendra
MCA Department, Lakireddy Bali Reddy College of Engineering, Mylavaram 521230, India



S. Majhi et al. (eds.), *Distributed Computing and Optimization Techniques*, Lecture Notes in Electrical Engineering 903, https://doi.org/10.1007/978-981-19-2281-7_78

## 1 Introduction

The proposed model is designed to detect the website phishing attacks using the Hyper-parameterized classification and regression tree algorithm by computing the number of cross fold validations using the Cross Validated Search (CVsearch) algorithm to identify the best iterations count to divide the data into training and testing data set, which plays an important role in improving the accuracy of the model. All the previous models have implemented the traditional algorithms by passing normal estimator values due to which the accuracy of the models is not up-to the mark. In this cross validated search, basing on the number of attributes and records the algorithm automatically finds the number of folds to be passed as input to the CART algorithm. XSS [1] is one type of attack that generates because of integrating the API within the website. Demon is the security tool to identify Cross-Site request forgery (CSRF) [2] security problems automatically i.e., a PHP-based framework to test the models safely during their runtime. It is one of the efficient open-source web applications for performing functionality testing among all the GUI-based testing tools [3]. The users need to first manually identify the sensitive Hypertext Transfer Protocol requests through the process of manual testing or SSL certificate understanding and then they should depend on automation techniques[4] that deals with security, and testing simultaneously like Open Web Application Security Project Testing Guidance[5], which reliably notifies them that this attack by is generated by using CSRF concept proof and the designed web browser supports them to visually check their results.

## 2 Related Works

In [7] Sundeep et al. proposed a web application in which they implemented a BTB prototype using the Integrated Development Environment (IDE) of visual studio 2015 using c#, whose object-oriented concepts help in handling the security issues and the backend is developed using the database that can handle the relations and network issues effectively. The proposed model implements a linear model SVM algorithm to determine the algorithm that can efficiently identify the features iteratively in their central server. The SVM algorithm has achieved an accuracy of "88.2%".

In [8] Alves et al. introduced CSRF detection approaches, i.e., CSRF text processing instruments, to handle the serious issues related to the text in the GUI forms. It also discusses issues related to CSRF advocacy and the measurements to be considered for providing safety to the designed forms. It also mentions a recent review of other major web security threats. The model has implemented Random Forest Algorithm which has resulted in accuracy of 88.7%.

In [10] Barreno et al. mentioned that the security risk in the software is susceptibility before releasing the software, they identify functionality using different techniques. Few loopholes neither can be fully deleted nor can it can avoid. Since these vulnerabilities are known as zero-day vulnerabilities, they are not correctly documented. Not only did competition and demand raise today's technology standards exponentially, but the number of threats and failures has moved. It somewhat tested the security of the software to fulfill both the deadline and customer's requirements. There are several tools [11] available to know the characteristics of threats within the code.

In [12] Ndichu et al. identified Cyber-attacks using manipulative JS codes always which involve increasing numbers of attacks and traditional security techniques like signal lentils, matching models, and intelligent search approaches that result in many false negatives and slow detection in terms of zero-day attacks. It is well known that the prevention and detection of attacks using alone SQL injections is quite inappropriate to handle in efficient manner. The drawbacks with the machine learning algorithms lie in training the module with accurate records. All these models have split the training and testing data using the traditional approaches like 70–30 and 80–20 ratios without inducing any randomness in the data.

## 3 Methods

The proposed paper works with the phishing dataset that contains 31 attributes and 11, 055 instances of phishing records. The system accepts the URL given by the user and passes it as input to the prediction system to check whether it is a phishing website or not. The proposed algorithm handles the missing values using the mean replacement technique and transforms the data into min–max standard normalization to scale all the values to the same range. Then Gini index is computed to identify the impurity nodes using below algorithm.

### *3.1 Algorithm for CART*

Input: Phishing Dataset, PD
Output: Tree-based on Gini Index
Begin:
1. Compute the Gini Index for all the attributes and select the attribute with the least value
2. Split the tree based on the best attribute
   a. if the class label can be determined then print the result
   b. else continue the computation of the Gini Index based on the best attribute
   c. K-folded_value←CV_Search(train,test,n)
   d. for i←1 to K-folded_value:
      compute the tree generation based on entropy
3. Continue step 2 until all the leaf nodes are marked as terminal nodes
End

The overall procedure is represented in Fig. 1.

Among all the machine learning algorithms decision trees are simple and easy to understand. A flavor of a decision tree is "Classification and Regression Trees" (CART), which are constructed based on the Gini Index. The Gini Index measures the impurity of the node, which means if the node belongs to a single class, then it is known as "Pure Node" otherwise it is known as "Impure Node". The Gini Index is computed as shown in Eq. 1.

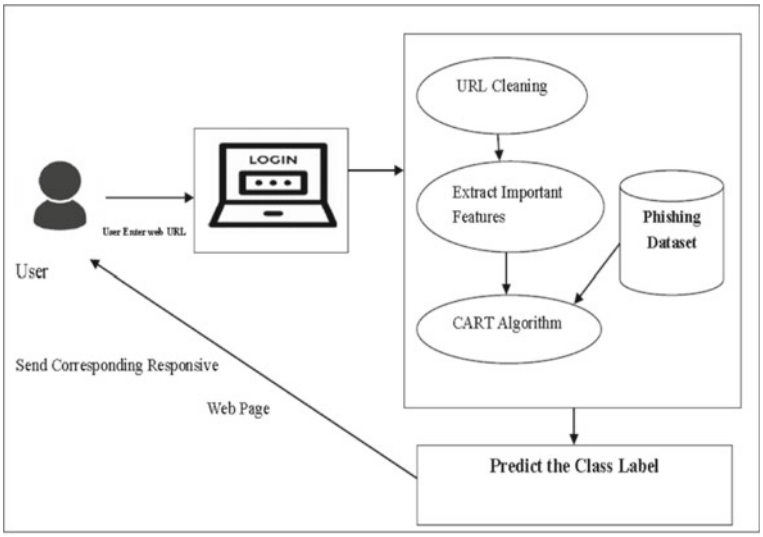



**Fig. 1** Block diagram for web site phishing response to the user

| SSLfinal_State | -1 | 1 | Count |
|---|---|---|---|
| -1 | 1836 | 1721 | 3557 |
| 1 | 810 | 5521 | 6331 |
| 0 | 1167 | 0 | 1167 |

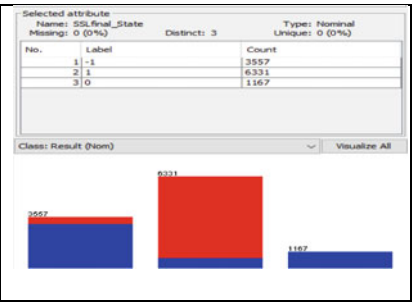



**Fig. 2** Number of Instances of SSLfinal_State with respect to –1, 1

$$G.I(C) = 1 - \sum_{i=1}^{n} P(C_i)^2 \tag{1}$$

where $P(C_i)$ represents the probability of each class label, which calculates the distribution of attacks over the different possibilities available among the given data.

Let us compute the Gini Index for the SSLfinal_State by considering the instances from the given dataset as shown in Fig. 2.

Therefore, the Gini Index of SSLfinal_State is computed using the weighted average of each feature value Gini Index and it can be computed as shown in Eqs. (2), (3) and (4) for the class labels –1, 0 and 1 respectively.

$$\text{G.I}(-1) = 1 - \left(\frac{1836^2}{3557}\right) - \left(\frac{1721^2}{3557}\right) = 1 - (0.51)^2 - (0.48)^2 = 1 - 0.26 - 0.23 = 0.51 \tag{2}$$

The above equations describe about the distribution of SSLfinal_state with possible values –1 and 1. So that the Gini index for the class is computed as "0.51"

$$\text{G.I}(1) = 1 - \left(\frac{810^2}{6331}\right) - \left(\frac{5521^2}{6331}\right) = 1 - (0.12)^2 - (0.87)^2 = 1 - .0.01 - 0.75 = 0.24 \tag{3}$$

The Gini index value for the feature SSLFinal_state with value "0" is computed as "0", which means that the node is most purified node which can be considered for taking the decisions about the tree split process.

$$\text{G.I(SSLfinal\_State)} = \frac{3557}{11055} * 0.51 + \frac{6331}{11055} * 0.24 + \frac{1167}{11055} * 0 => 0.32 * 0.51 + 0.57 * 0.24 + 0 => 0.163 + 0.136 => 0.29 \tag{4}$$

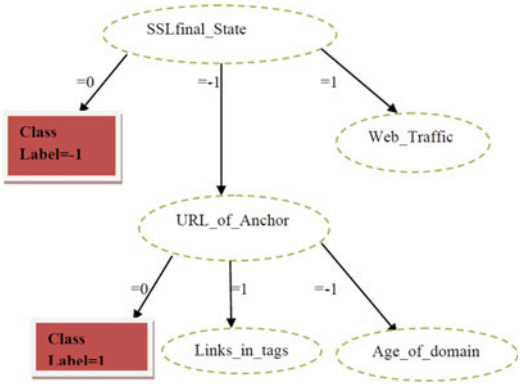



**Fig. 3** Tree Construction of CART

From the above equations, the model generates the decision rules as follows.

a. If (SSLfina_state = 0) then it is –1, which represents it is a phishing website
b. else tree should be further expanded for decision making.

Among the 31 attributes, the SSLfinal_State has obtained the Gini Index as 0.29, which is the least value. So, it is considered as the pure node and made as to the root node for the CART. The given dataset is a multi-classification problem hence it uses the heuristic search procedure for performing the binary split at each step of the tree construction. One of the hyper parameters that have to be taken care of in this algorithm is "Number of Objects to be Pruned", which represents the reduction of tree size by computing the cross-validation accuracy.

If the K-Folded accuracy is not improved, then it stops expanding the tree further. In this paper, the K value is considered as 5. The novelty of this CART algorithm is, the model identifies the K value by performing the CVSEARCH, hyper-parameterization approach and this K-folded cross validation is applied for the traditional CART algorithm. Based on the Gini Index information a sample screenshot of the tree is constructed as shown in Fig. 3.

The resultant tree generates 244 leaf nodes and the size of the tree is 487. In the above Fig. 3, the leaf nodes are represented with the rectangle symbol, and internal nodes are represented using the ellipse. To evaluate the performance, the system has performed 5-cross fold validation i.e., cross-validation is the measurement to assess the quality of the model so that it holds out some data at every iteration and uses the remaining data for assessing the model.

## 4 Results and Discussion

The model is executed in the popular data mining tool known "WEKA Tool". WEKA tool is GUI based tool which is simple in terms of performing the machine learning algorithms by using the drag and drop options. The WEKA tool is installed from the website: http://www.cs.waikato.ac.nz/ml/weka. The proposed algorithm splits the dataset into 80% of training data and 20% testing data. The test dataset contains 2211 instances, out of which the algorithm has classified 2114 instances correctly and has achieved an accuracy of 95.62%. Confusion Matrix, sometimes it is also known as Error Matrix which computes the statistical measurement over the test data to measure the worthiness of the predicted class labels. The Classification and Regression Trees algorithms detect whether a website is phishing or not. The detailed accuracy results are reported in Table 1.

The comparison between the positive class labels and negative class labels data are represented in Fig. 4 and it is found that dataset contains the balanced data. So, with individual traditional algorithms, only the model can obtain good accuracy rather than using the ensemble algorithm. The X-axis represents the metrics values required for computation of performance of the model and Y-axis represents the class labels possible values.

**Table 1** Accuracy results of CART algorithm

| | Precision | Recall | F-Measure | ROC area | TP rate | FP rate |
|---|---|---|---|---|---|---|
| −1 | 0.977 | 0.925 | 0.95 | 0.987 | 0.925 | 0.018 |
| 1 | 0.941 | 0.982 | 0.961 | 0.987 | 0.982 | 0.075 |
| Weighted avg | 0.957 | 0.956 | 0.956 | 0.987 | 0.956 | 0.049 |

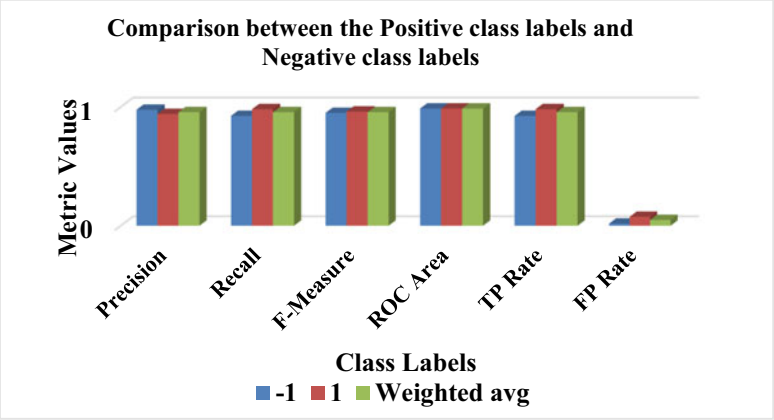



**Fig. 4** Performance metrics for positive and negative class labels

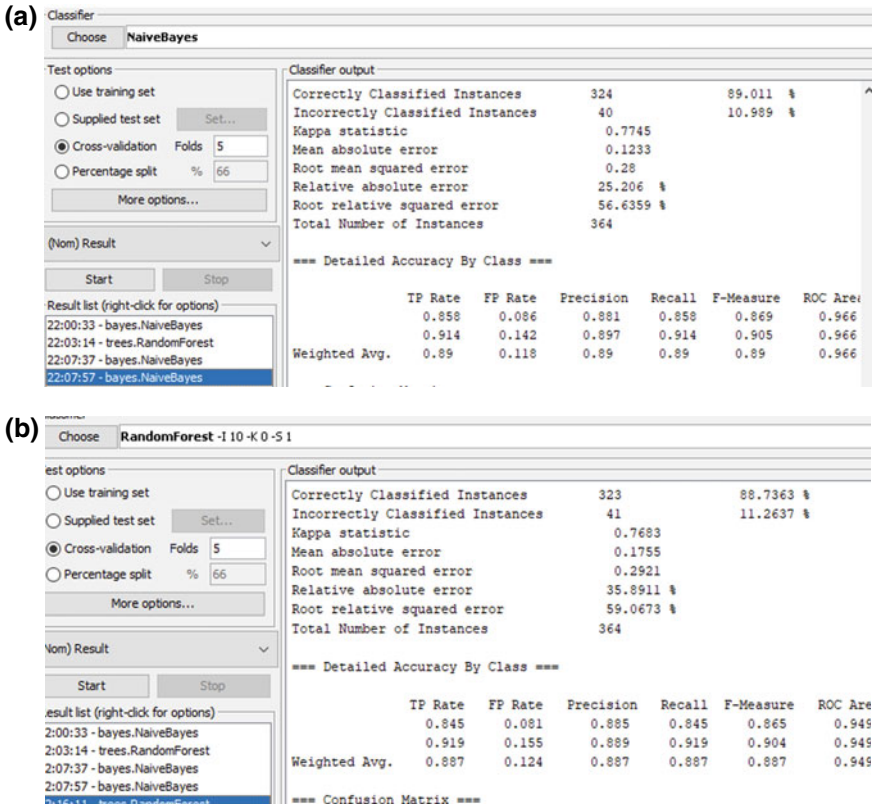



**Fig. 5** **a** Naïve Bayesian (Traditional Algorithm). **b** Random Forest (Ensemble Algorithm)

To evaluate the performance of the model it has compared the proposed model with both traditional and ensemble algorithms as shown in Fig. 5a and 5b.

The comparison between these methods namely Naïve Bayesian, Random Forest and, the proposed hyper-parameterized CART algorithm on various parameters is illustrated in Table 2.

**Table 2** Comparative study on evaluation metrics

| | Accuracy (%) | Recall (%) | Precision (%) | F-Measure (%) | Kappa (%) |
|---|---|---|---|---|---|
| Naïve Bayesian [11] | 89.01 | 89 | 89 | 89 | 77.4 |
| Decision Trees [6] | 85.33 | 82.1 | 84 | 83.1 | 73.8 |
| Logistic Regression [5] | 91.1 | 89.54 | 89.54 | 89.54 | 80.22 |
| SVM [7] | 88.2 | 80.1 | 80.1 | 80.1 | 73.2 |
| Random Forest [8] | 88.7 | 88.7 | 88.7 | 88.7 | 76.8 |
| Proposed Algorithm | 95.62 | 95.6 | 95.7 | 95.6 | 88.78 |

## 5 Conclusion

Website Phishing is a very simple attack but manipulates the person with very small manipulations in the original URL. The CART algorithm is good at handling the multi-classification problem and it can also handle the non-linear relationship with less time complexity notations. The major goal of this paper is to identify the URL passed by the user is a fake website or not. If the website is fake, then it sends a response as a security alerts otherwise it opens the corresponding URL. The same problem can be handled by many machine learning algorithms but the proposed algorithm can handle both numerical and categorical data. At the same time, when the system has compared the accuracy performed by the ensemble algorithms, the proposed algorithm has achieved 95.62% accuracy whereas the Random Forest, ensemble algorithm has achieved 88.7% accuracy. So, the proposed algorithm helps the users who want to protect their data from cyberattacks while browsing the information on the internet or performing e-commerce transactions regularly.

## References

1. Kumar A, Gupta A, Mittal P, Gupta PK, Varghese S (2021) Prevention of XSS attack using Cryptography and API integration with Web Security, 25 April 2021. https://ssrn.com/abstract=3833910, http://dx.doi.org/10.2139/ssrn.3833910
2. Wibowo R, Sulaksono A (2021) Web vulnerability through cross site scripting (XSS) detection with OWASP security Shepherd. Indonesian J Inf Syst 3(2):149–159. https://doi.org/10.24002/ijis.v3i2.4192
3. Arora N, Singh P, Sahu S, Keshari VK, Vinoth Kumar M (2021) Preventing SSRF (Server-Side Request Forgery) and CSRF (Cross-Site Request Forgery) using extended visual cryptography and QR Code. In: Goyal D, Chaturvedi P, Nagar AK, Purohit S (eds) Proceedings of Second International Conference on Smart Energy and Communication. AIS. Springer, Singapore. https://doi.org/10.1007/978-981-15-6707-0_20

4. Hashim A, Medani R, Attia TA (2021) Defences against web application attacks and detecting phishing links using machine learning. In: 2020 international conference on computer, control, electrical, and electronics engineering (ICCCEEE), pp 1–6. https://doi.org/10.1109/ICCCEEE49695.2021.9429609
5. Gogoi B, Ahmed T, Saikia, HK (2021) Detection of XSS attacks in web applications: a machine learning approach. Int J Innov Res Comput Sci Technol (IJIRCST) 9(1):1–10. https://ssrn.com/abstract=3778063. ISSN 2347-5552
6. Munonye K, Péter M (2021) Machine learning approach to vulnerability detection in OAuth 2.0 authentication and authorization flow. Int J Inf Secur 21(2):223–237. https://doi.org/10.1007/s10207-021-00551-w
7. Tommy R, Sundeep G, Jose H (2017) Automatic detection and correction of vulnerabilities using machine learning. In: 2017 international conference on current trends in computer, electrical, electronics and communication (CTCEEC). https://doi.org/10.1109/ctceec.2017.8454995
8. Alves H, Fonseca B, Antunes N (2016) Experimenting machine learning techniques to predict vulnerabilities. In: 2016 Seventh Latin-American symposium on dependable computing (LADC), pp 151–156. IEEE
9. Monika TV (2021) DPLOOP: detection and prevention of loopholes in web application security. In: Gao XZ, Tiwari S, Trivedi M, Mishra K (eds) Advances in Computational Intelligence and Communication Technology. AISC, vol 1086. Springer, Singapore. https://doi.org/10.1007/978-981-15-1275-9_14
10. Barreno M, Nelson B, Joseph AD, Tygar JD (2010) The security of machine learning. Mach Learn 81(2):121–148
11. Calzavara S, Conti M, Focardi R, Rabitti A, Tolomei G (2020) Machine learning for web vulnerability detection: the case of cross-site request forgery. IEEE Secur. Priv. 18(3):8–16. https://doi.org/10.1109/MSEC.2019.2961649
12. Ndichu S, Kim S, Ozawa S, Misu T, Makishima K (2019) A machine learning approach to detection of JavaScript-based attacks using AST features and paragraph vectors. Appl Soft Comput 84:105721. https://doi.org/10.1016/j.asoc.2019.105721
13. Khalid MN, Iqbal M, Rasheed K, Abid MM (2020) Web vulnerability finder (WVF): automated black- box web vulnerability scanner. Int J Inf Technol Comput Sci 12(4):38–46. https://doi.org/10.5815/ijitcs.2020.04.05
14. Chen H, Chen J, Chen J, Yin S, Wu Y, Xu J (2020) An automatic vulnerability scanner for web applications. In: 2020 IEEE 19th international conference on trust, security and privacy in computing and communications (TrustCom), pp 1519–1524. https://doi.org/10.1109/TrustCom50675.2020.00207
15. Sirisha G, Reddy AM (2018) Smart healthcare analysis and therapy for voice disorder using cloud and edge computing. In: 2018 4th international conference on applied and theoretical computing and communication technology (iCATccT), Mangalore, India, pp 103–106. https://doi.org/10.1109/iCATccT44854.2018.9001280
16. Mallikarjuna A, KarunaSree B (2019) Security towards flooding attacks in inter domain routing object using ad hoc network. Int J Eng Adv Technol (IJEAT) 8(3)